



**Universidad de Valladolid**



**ESCUELA DE INGENIERÍAS  
INDUSTRIALES**

**UNIVERSIDAD DE VALLADOLID**

**ESCUELA DE INGENIERÍAS INDUSTRIALES**

**Grado en Ingeniería Electrónica Industrial y Automática**

# **Integración de un sistema de visión 3D en un entorno robotizado de cirugía**

**Autor:**

**Loidi Yarza, Ander**

**Tutor:**

**Fuente López, Eusebio de la**

**Ingeniería en Sistemas y  
Automática**

**Valladolid, junio de 2019**



Universidad de Valladolid

# Integración de un sistema de visión 3D en un entorno robotizado de cirugía Ander Loidi Yarza



ESCUELA DE INGENIERÍAS  
INDUSTRIALES



## RESUMEN Y PALABRAS CLAVES

---

### RESUMEN

Hoy en día la cirugía endoscópica endonasal transesfenoidal es una técnica mínimamente invasiva que presenta numerosas ventajas al permitir a los cirujanos un acceso rápido y directo a determinadas lesiones cerebrales a través de las fosas nasales. Sin embargo, este tipo de cirugía requiere de la participación de dos cirujanos que deben operar con elevada precisión en una zona de trabajo con muchas restricciones de espacio. Por lo tanto, en el presente Trabajo Fin de Grado se expone un sistema de visión artificial tridimensional que ayuda al cirujano y al guiado de un robot quirúrgico basándose en la información proporcionada por una cámara ToF (Time of Flight) y la información preoperatoria obtenida mediante una serie de imágenes de resonancia magnética (IRM). Este sistema formará parte de un asistente robótico que gestione el manejo del endoscopio de forma autónoma.

### PALABRAS CLAVE

- Robótica médica.
- Imágenes de Resonancia Magnética (IRM).
- Visión artificial.
- Cirugía endoscópica endonasal transesfenoidal.
- PCL – Point Cloud Library.



Universidad de Valladolid

# Integración de un sistema de visión 3D en un entorno robotizado de cirugía Ander Loidi Yarza



ESCUELA DE INGENIERÍAS  
INDUSTRIALES



## ABSTRACT AND KEY WORDS

---

### ABSTRACT

Nowadays, endoscopic transsphenoidal endonasal surgery is a minimally invasive technique with numerous advantages allowing surgeons quick and direct access to certain brain injuries through the nostrils. However, this type of surgery requires the participation of two surgeons who must operate with high precision in a very restricted working space. Therefore, the following final degree project is exposed by a three-dimensional artificial vision system that supports the surgeon and the guide of a surgical robot based on the information provided by de ToF (Time of Flight) camera and the preoperative information obtained through a series of magnetic resonance images (MRI). This system Will be part of a robotic assistant that manages the handling of the endoscope autonomously.

### KEY WORDS

- Medical robotics.
- Magnetic Resonance Imaging (MRI).
- Artificial vision.
- Endoscopic transsphenoidal endonasal surgery.
- PCL – Point Cloud Library



Universidad de Valladolid

# Integración de un sistema de visión 3D en un entorno robotizado de cirugía Ander Loidi Yarza



ESCUELA DE INGENIERÍAS  
INDUSTRIALES



## ÍNDICE

---

1	Introducción.....	13
1.1	Introducción al capítulo.....	13
1.2	Contextualización del proyecto.....	13
1.3	La robótica en la medicina .....	14
1.4	Objetivos del proyecto .....	15
1.5	Índice desarrollado.....	16
1.5.1	Introducción.....	16
1.5.2	Entorno de desarrollo y bases teóricas .....	16
1.5.3	Software y sistema de visión artificial utilizados .....	16
1.5.4	Obtención nube tridimensional por captura .....	17
1.5.5	Procesamiento de IRMs.....	17
1.5.6	Conclusiones y líneas futuras .....	17
2	Entorno de desarrollo y bases teóricas .....	19
2.1	Introducción al capítulo.....	19
2.2	Entorno de desarrollo.....	19
2.3	Visión artificial .....	21
2.3.1	Visión humana y visión artificial.....	22
2.3.2	Aplicaciones.....	24
2.3.3	Visión Artificial en la medicina .....	25
2.3.4	Componentes de un sistema de visión artificial.....	26
2.3.5	Funcionamiento .....	28
2.4	Operaciones de cirugía transesfenoidal .....	29
2.4.1	Procedimiento quirúrgico .....	31
2.4.2	Ventajas .....	33
2.5	Imágenes de resonancia magnética.....	33
3	Software y sistema de visión artificial utilizados .....	37
3.1	Introducción al capítulo.....	37
3.2	Obtención de modelos tridimensionales .....	37
3.2.1	Cámara estereoscópica.....	37
3.2.2	Cámara ToF .....	40
3.3	Software utilizado.....	44
3.3.1	Ubuntu - Linux.....	44



3.3.2	OpenCV .....	45
3.3.3	PCL .....	46
4	Obtención nube tridimensional por captura .....	49
4.1	Introducción al capítulo.....	49
4.2	Implementación de cámaras estéreo .....	49
4.3	Implementación de cámaras ToF.....	57
5	Procesamiento de IRMs.....	75
5.1	Introducción al capítulo.....	75
5.2	Generación de los tres planos anatómicos .....	76
5.2.1	Plano sagital .....	76
5.2.2	Plano coronal.....	78
5.3	Obtención tridimensional de las narinas.....	80
5.4	Obtención tridimensional de la cabeza .....	95
5.4.1	Nube contorno.....	95
5.4.2	Nube completa .....	98
5.5	Obtención tridimensional del cerebro.....	102
5.6	Sistema de navegación quirúrgico .....	109
6	Conclusiones y líneas futuras .....	113
6.1	Introducción al capítulo.....	113
6.2	Conclusión del proyecto.....	113
6.2.1	Resultados cámara ToF y registro de nubes de puntos.....	114
6.2.2	Resultados análisis de IRM .....	114
6.3	Líneas futuras.....	116
6.3.1	Modulo con acelerómetro/giroscopio.....	116
6.3.2	Colorear modelos tridimensionales .....	116
7	Bibliografía.....	119



## ÍNDICE DE FIGURAS

<i>Figura 1: Aplicaciones de la robótica en la medicina.</i>	15
<i>Figura 2: Esquema proyecto CRANEEAL.</i>	20
<i>Figura 3: Sistema de visión humano frente a la visión artificial.</i>	23
<i>Figura 4: Previsión económica en el mercado de la visión artificial.</i>	25
<i>Figura 5: Representación de superficie y salida de sensores.</i>	27
<i>Figura 6: Representación del proceso de digitalización y cuantización.</i>	29
<i>Figura 7: IRM del plano sagital de un paciente sano (izquierda) frente a uno con una lesión tumoral (derecha).</i>	30
<i>Figura 8: Abordaje nasal (izquierda) y trayectoria para acceso transesfenoidal a la hipófisis (derecha).</i>	31
<i>Figura 9: Cavidad nasal.</i>	32
<i>Figura 10: Scanner de resonancia magnética y representación 3D del proceso de obtención de IRMs.</i>	34
<i>Figura 11: Mejora por aporte de gadolinio.</i>	35
<i>Figura 12: IRM secuencia T1 (izquierda) y T2 (derecha).</i>	36
<i>Figura 13: Esquema de funcionamiento de cámaras estereoscópicas.</i>	39
<i>Figura 14: Principio de funcionamiento de cámara ToF.</i>	41
<i>Figura 15: Profundidad calculada estimando la diferencia de fase entre la señal IR radiada (a) y la reflejada (b).</i>	43
<i>Figura 16: Librerías PCL.</i>	47
<i>Figura 17: Cámaras estereoscópicas.</i>	49
<i>Figura 18: Cráneo artificial y espacio de trabajo.</i>	50
<i>Figura 19: Tablero utilizado para la calibración.</i>	51
<i>Figura 20: Eliminación de valores atípicos aplicando SOR.</i>	53
<i>Figura 21: Eliminación basada en componentes conexas aplicando ROR.</i>	54
<i>Figura 22: Nubes de puntos inicial (arriba) y filtrada (abajo) obtenidas.</i>	55
<i>Figura 23: Visualización detallada de la superficie obtenida.</i>	56
<i>Figura 24: Cámara ToF utilizada en el proyecto.</i>	57
<i>Figura 25: Disposición de trabajo para la captura de la cámara ToF.</i>	58
<i>Figura 26: Resultado de eliminar puntos NaN.</i>	59
<i>Figura 27: Esquema de aplicación del filtro VoxelGrid.</i>	60
<i>Figura 28: Resultado de aplicar el filtro VoxelGrid sobre las nubes de puntos.</i>	61
<i>Figura 29: Representación de las nubes obtenidas.</i>	62
<i>Figura 30: Visualización detallada de la superficie obtenida.</i>	63
<i>Figura 31: Esquema de posicionamiento de la cámara para las capturas.</i>	65
<i>Figura 32: Robot colaborativo UR3 (izquierda) y modelo geométrico (derecha).</i>	66
<i>Figura 33: Resultado de llevar a cabo el registro de las nubes de puntos.</i>	68
<i>Figura 34: Nube de puntos de partida para aplicar RANSAC.</i>	70



<i>Figura 35: Selección aleatoria.</i>	70
<i>Figura 36: Calculo del modelo lineal.</i>	70
<i>Figura 37: Valoración de la desviación estándar.</i>	71
<i>Figura 38: Obtención de puntuación.</i>	71
<i>Figura 39: Resultado obtenido por RANSAC, mostrando inliers (izquierda) y outliers(derecha).</i>	72
<i>Figura 40: Nube de puntos de entrada que contiene un plano a eliminar.</i>	73
<i>Figura 41: Resultado obtenido por RANSAC, mostrando inliers (izquierda) y outliers (derecha).</i>	73
<i>Figura 42: Representación de los tres planos anatómicos principales.</i>	75
<i>Figura 43: IRM correspondientes al plano axial.</i>	76
<i>Figura 44: Esquema de la generación de las IRM del plano sagital a partir del axial.</i>	77
<i>Figura 45: IRM correspondientes al plano sagital.</i>	78
<i>Figura 46: Esquema de la generación de las IRM del plano coronal a partir del axial.</i>	79
<i>Figura 47: IRM correspondientes al plano coronal.</i>	80
<i>Figura 48: Localización del centro de la región de interés (ROI) sobre imagen axial.</i>	81
<i>Figura 49: Diferencia entre aplicar un umbral global (arriba) y uno adaptativo (abajo).</i>	83
<i>Figura 50: Resultado de binarizar con umbral adaptativo.</i>	83
<i>Figura 51: Caracterización de la región de interés (ROI).</i>	84
<i>Figura 52: Ejemplo esquemático de aplicar una erosión sobre A con un ee B.</i>	85
<i>Figura 53: Ejemplo esquemático de aplicar una dilatación sobre A con un ee B.</i>	86
<i>Figura 54: Resultado de aplicar la operación morfológica de cierre sobre la ROI.</i>	87
<i>Figura 55: Detección de contornos sobre la ROI.</i>	88
<i>Figura 56: Representación por colores de las diferentes superficies en función de su tamaño.</i>	89
<i>Figura 57: Segmentación de las narinas por superficie.</i>	89
<i>Figura 58: Proceso de clasificación de las fosas nasales individualmente.</i>	91
<i>Figura 59: Ejemplos de obtención de las narinas individualmente.</i>	92
<i>Figura 60: Vista frontal e inferior del modelo 3D de las fosas nasales.</i>	93
<i>Figura 61: Ejemplos de obtención del contorno externo de la nariz.</i>	93
<i>Figura 62: Vista frontal del modelo de las fosas nasales junto con el de la nariz.</i>	94
<i>Figura 63: Vista inferior del modelo de las fosas nasales junto con el de la nariz.</i>	94
<i>Figura 64: Diferentes binarizaciones.</i>	96
<i>Figura 65: Imagen original y su correspondiente binarización inversa.</i>	96
<i>Figura 66: Resultado de implementar el rellenado con floodfill.</i>	97
<i>Figura 67: Imagen rellena y su correspondiente inversa.</i>	97



<i>Figura 68: Ejemplos del resultado obtenido con las imágenes combinadas. ....</i>	<i>98</i>
<i>Figura 69: Imagen original y su correspondiente binarización. ....</i>	<i>99</i>
<i>Figura 70: Ejemplos de la obtención de los contornos en cada corte de las IRM. .....</i>	<i>100</i>
<i>Figura 71: Modelo tridimensional correspondiente al exterior de la cabeza. ....</i>	<i>101</i>
<i>Figura 72: Modelo tridimensional completo de la cabeza. ....</i>	<i>102</i>
<i>Figura 73: Imagen original y su correspondiente binarización global. ....</i>	<i>103</i>
<i>Figura 74: Resultado de aplicar la erosión sobre la imagen. ....</i>	<i>104</i>
<i>Figura 75: Resultado de aplicar el cierre y la dilatación de refuerzo sobre la imagen.....</i>	<i>105</i>
<i>Figura 76: Representación de todos los contornos encontrados en la imagen. .</i>	<i>105</i>
<i>Figura 77: Definición de la región de aceptación junto con su centro. ....</i>	<i>106</i>
<i>Figura 78: Representación de los elementos en función del cumplimiento de las condiciones. ....</i>	<i>107</i>
<i>Figura 79: Representación del contorno valido que cumple las condiciones.....</i>	<i>107</i>
<i>Figura 80: Representación del modelo tridimensional cerebral creado a partir de las IRM.....</i>	<i>108</i>
<i>Figura 81: Representación del modelo cerebral creado junto con el contorno exterior de la cabeza. ....</i>	<i>109</i>
<i>Figura 82: Diferentes menús creados para la navegación quirúrgica. ....</i>	<i>110</i>
<i>Figura 83: Menú auxiliar para la selección de sección.....</i>	<i>111</i>
<i>Figura 84: Sección individual del plano coronal (izquierda) y superficie externa axial hasta una sección determinada (derecha). ....</i>	<i>112</i>



Universidad de Valladolid

# Integración de un sistema de visión 3D en un entorno robotizado de cirugía Ander Loidi Yarza



ESCUELA DE INGENIERÍAS  
INDUSTRIALES



# 1 INTRODUCCIÓN

---

## 1.1 INTRODUCCIÓN AL CAPÍTULO

En este primer capítulo se realiza una descripción breve y general acerca del trabajo “Integración de un sistema de visión 3D en un entorno robotizado de cirugía”. Además, se añade una previa contextualización con la cual se introducen conceptos básicos acerca del proyecto.

## 1.2 CONTEXTUALIZACIÓN DEL PROYECTO

La visión es sin duda una de las principales y más complejas capacidades sensoriales del ser humano, con la cual el ser humano es capaz de interpretar el entorno que le rodea. Con ella se pueden percibir múltiples colores, detectar el movimiento, analizar y diferenciar la forma de objetos e incluso percibir distancias. Se puede pensar que las personas ven gracias a los ojos, pero estos no son más que sensores. Como Alberich, Gómez y Ferrer (2013) concluyen: “Aunque lo intuitivo sería creer que vemos por nuestros ojos, es mucho más correcto decir que el que ve es nuestro cerebro. La mayor parte del proceso de visión se produce en él y los ojos pueden considerarse como meros receptores de estímulos luminosos” (p.5). La visión es un aspecto de gran utilidad en la vida cotidiana y a pesar de poder considerarla casi como un sistema perfecto, puede sufrir deterioros físicos y, además, presentar numerosas limitaciones como la de ser incapaz de percibir ciertas frecuencias de luz.

Debido a la importancia de corregir las citadas imperfecciones del sistema de visión humano y a la exponencial y constante evolución de las tecnologías en la que estamos inmersos, se ha desarrollado hardware de visión artificial que satisface y excede las necesidades del sistema visual humano.

Se sabe que hoy en día la automatización de sistemas y procesos es una cuestión en auge e innegable. Si a esto se le suma la necesidad que tienen las labores de visión de varios ámbitos con intereses científicos o industriales en requerir un nivel de precisión muy elevado, es obvia la transición de trabajar con la vista humana a la implementación de los mencionados sistemas de visión artificial.

La automatización industrial se define como la implantación de elementos computarizados y electromecánicos para controlar y monitorizar un proceso o dispositivo, que por lo general cumple funciones o tareas repetitivas,



haciendo que opere automáticamente, reduciendo al máximo la intervención humana. Aplicando dicho concepto concretamente a este proyecto, se puede decir que el trabajo está centrado en la automatización de un proceso médico, el cual se explicará posteriormente, con el fin de aumentar la calidad y seguridad en su ejecución.

### 1.3 LA ROBÓTICA EN LA MEDICINA

Los campos de la robótica y la medicina, analizándolos individualmente, han evolucionado considerablemente en las últimas décadas. Cabe suponer que cada uno de estos campos es totalmente dispar entre sí, pero debido a los avances en cada uno de ellos ha podido llevarse a cabo su fusión, la cual es muy interesante. A pesar de todas sus diferencias, se han unido con un objetivo común, el cuidado del cuerpo y la salud del ser humano. Es, por todo ello, por lo que la robótica ha adquirido en estos últimos años un papel fundamental en el campo de la medicina.

Los denominados “robots de servicio” son los encargados de trabajar de forma autónoma o semiautónoma, es decir, asistida por una persona humana, para llevar a cabo una serie de tareas útiles para el bienestar de las personas. Estos “robots de servicio” se manifiestan en el campo de la medicina de diversas maneras:

- Robots quirúrgicos
- Robots de almacenamiento y distribución de medicamentos
- Robots Asistenciales
- Robots-Prótesis
- Exoesqueletos
- Robótica terapéutica



Figura 1: Aplicaciones de la robótica en la medicina.

(Fuente: (izquierda) «Da Vinci Surgery | Da Vinci Surgical System | Robotic Technology» [En línea]. Disponible: <https://www.davincisurgery.com/>. (derecha) «Motorika rehabilitation robotics» [En línea]. Disponible: <http://motorika.com/>.)

Por lo tanto, es indiscutible observar el gran abanico de ámbitos que la robótica abarca en la medicina, ya sea en cirugías de alto riesgo, en rehabilitaciones o en la ayuda a personas con enfermedades de movilidad reducida o discapacitadas.

#### 1.4 OBJETIVOS DEL PROYECTO

El trabajo forma parte del proyecto CRANEEAL, un proyecto de investigación del Instituto de Tecnologías Avanzadas de la Producción financiado por el Ministerio de Economía, Industria y Competitividad. el cual se expondrá en el próximo capítulo. A pesar de no conocer aún las metas de dicho programa de manera global, queda definir cuáles son los objetivos de este trabajo. Concretamente, lo principal será centrarse en el sistema de visión artificial que implementa el proyecto con el fin de investigar en el diseño y creación de nuevos elementos y sistemas de navegación quirúrgicos tridimensionales que faciliten, al cirujano y al robot asistencial durante la intervención endoscópica, datos adicionales para mejorar ciertas características como pueden ser la posición, orientación y precisión que debe mantener el robot asistencial, así como la trayectoria que este debe seguir o la seguridad del paciente a lo largo de la intervención.

Para ello se realizan dos desarrollos paralelos. Por un lado, la creación de una nube de puntos tridimensional a partir de la toma de imágenes de la



cabeza del paciente, con el fin de emparejar dicha nube con el modelo que se genere a través del procesado de las imágenes de resonancia magnética, siendo este último el otro desarrollo que se realiza.

Para poder entender mejor ambos procedimientos, se explicarán detalladamente cada uno de ellos de manera individual en los capítulos 4 y 5.

## 1.5 ÍNDICE DESARROLLADO

El documento se ha dividido en 6 capítulos, de los cuales se aporta información resumida a continuación con el objetivo de tener un conocimiento de la materia más relevante de cada uno de ellos.

### 1.5.1 Introducción

Descripción breve y general acerca del Trabajo Fin de Grado. Además, se añade una previa contextualización con la cual se introducen conceptos básicos acerca del proyecto.

### 1.5.2 Entorno de desarrollo y bases teóricas

Explicación del marco de trabajo bajo el cual se ha desarrollado el proyecto, además de una introducción a la visión artificial y sus aplicaciones, las operaciones de cirugía transesfenoidal y las imágenes de resonancia magnética.

### 1.5.3 Software y sistema de visión artificial utilizados

Análisis de los dos métodos de visión puestos en práctica y del software con el que se ha trabajado para el procesamiento de imágenes.



#### 1.5.4 Obtención nube tridimensional por captura

Generación de la nube de puntos tridimensional a partir de la toma de imágenes de la cabeza del paciente. Desarrollada con los dos métodos expuestos en el capítulo anterior

#### 1.5.5 Procesamiento de IRMs

Distintos análisis a las imágenes de resonancia magnética (IRM), con el objetivo obtener diferentes modelos tridimensionales de estructuras anatómicas humanas.

#### 1.5.6 Conclusiones y líneas futuras

Reflexión acerca de los resultados obtenidos, comprobando si se han cumplido los objetivos establecidos. Además, se citarán las conclusiones extraídas debido al desarrollo llevado a cabo junto con algunas de las líneas futuras aplicables al proyecto.



Universidad de Valladolid

# Integración de un sistema de visión 3D en un entorno robotizado de cirugía Ander Loidi Yarza



ESCUELA DE INGENIERÍAS  
INDUSTRIALES



## 2 ENTORNO DE DESARROLLO Y BASES TEÓRICAS

---

### 2.1 INTRODUCCIÓN AL CAPÍTULO

En este capítulo se explica el marco de trabajo bajo el cual se ha desarrollado el proyecto, de manera que lo contextualice.

Por otro lado, se hace una introducción de la visión artificial y sus aplicaciones, así como de las operaciones de cirugía transesfenoidal y las imágenes de resonancia magnética.

### 2.2 ENTORNO DE DESARROLLO

ITAP, siglas que corresponden al “*Instituto de las Tecnologías Avanzadas de la Producción*”, es un instituto universitario de investigación situado en la Universidad de Valladolid, en funcionamiento desde el 22 de septiembre de 2011.

Dicho instituto está formado por más de 20 profesores de la Universidad de Valladolid pertenecientes a la Escuela de Ingenierías Industriales y a la Facultad de Medicina. Dicha combinación es perfecta para poder llevar a cabo los proyectos que citaré posteriormente. Cabe destacar que ITAP colabora con más de 25 profesionales, destacando investigadores de la Fundación Cartif, la cual comparte varios objetivos comunes, y médicos del Hospital Clínico Universitario y el Hospital Universitario Río Hortega de Valladolid. Todos ellos colaboran en tareas del ámbito de las ingenierías en general con especial actividad en las ingenierías industriales y en la ingeniería médica, compaginando las labores de investigación con las transferencias de conocimiento y tecnología.

Debido a que los principales objetivos del instituto universitario son la generación, transmisión y difusión de nuevos conocimientos en el ámbito de la ingeniería, se siguen unas líneas de investigación bastante amplias en campos como la producción industrial y el control de procesos, visión artificial, medicina y cirugía, el medioambiente, simulación y modelado, mantenimiento de infraestructuras y patrimonio, etc.

Es por ello por lo que dicho instituto es un entorno ideal donde desarrollar el proyecto y lo que este trabajo va a llevar a cabo.

En el proyecto participan la Universidad de Málaga, la Universidad Miguel Hernández y la Universidad de Valladolid y está financiado por el Ministerio de Economía, Industria y Competitividad (MINECO). El proyecto toma el nombre de “CRANEEAL” y consiste en la creación de un robot de asistencia quirúrgica en cirugías endoscópicas endonasales guiadas por imágenes. Dicho sistema puede verse de manera esquemática en la figura 2. Estas, son operaciones en las que el cirujano hace uso de la tecnología informática en conjunto con imágenes radiológicas, concretamente en este caso imágenes de resonancia magnética (IRM), para proceder a la intervención con ayuda de un robot quirúrgico.

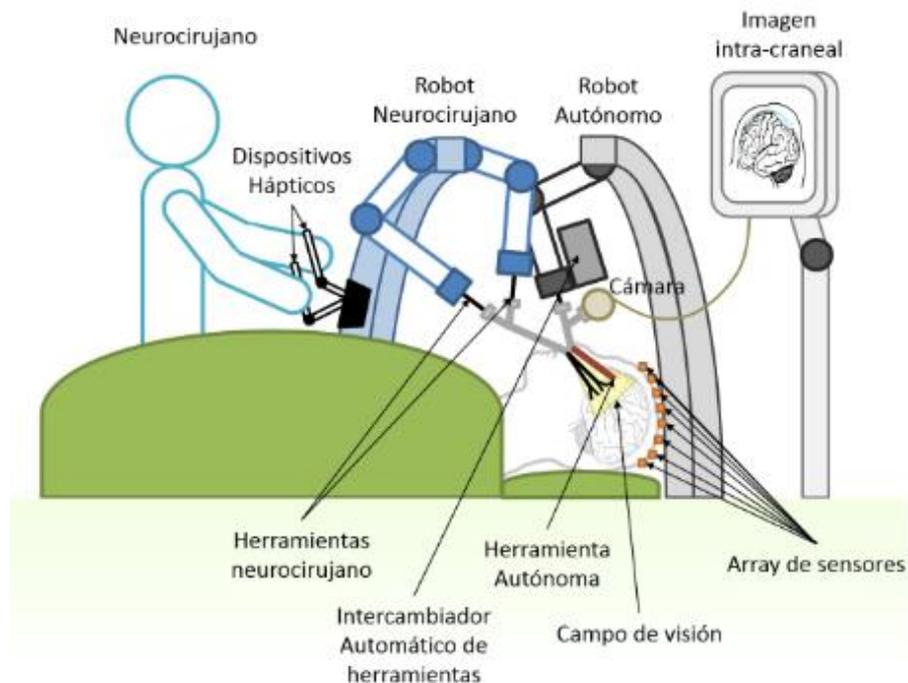


Figura 2: Esquema proyecto CRANEEAL.

(Fuente: «Proyecto CRANEEAL» [En línea]. Disponible: <https://www.uma.es/medical-robotics/info/107696/proyectedpi201680391-C3-1-R/>)

Para poder llevar a cabo dichas operaciones y lograr su propósito general, es necesario definir las pautas que deberán llevarse a cabo:

- 1- Diseño y desarrollo de un concepto de sistema robotizado especializado en el abordaje endonasal para la neurocirugía:  
Deberá controlar tanto los movimientos automáticos como los colaborativos de ambos brazos robóticos, adaptarse a las necesidades del cirujano y del quirófano y ser capaz de realizar un cambio de herramienta quirúrgica.



## 2-Diseño y desarrollo de un sistema de navegación neuroquirúrgica para abordaje endonasal:

Durante el preoperatorio se deberán proporcionar las herramientas necesarias para la planificación inicial de la colaboración y de los movimientos de los diferentes instrumentos. Así mismo, deberá actualizar el plan durante la operación debido al “brain-shift” y proporcionar información a través de realidad aumentada.

Para comprender mejor el término de “brain-shift”, se puede ver lo que Nimsky, Ganslandt, Hastreiter y Fahlbusch (2001) concluyen: “La extirpación del tumor y el drenaje de líquido cefalorraquídeo provocan una deformación cerebral intraoperatoria conocida como desplazamiento cerebral. Por lo tanto, los sistemas de navegación neuronal que se basan en datos de imágenes preoperatorias tienen una precisión decreciente durante el procedimiento quirúrgico” (p. 357-364).

## 3-Diseño y desarrollo de técnicas de cooperación entre el cirujano y el sistema robótico durante el abordaje endonasal:

Deberá proporcionar un enfoque cognitivo y estar fundamentada en el concepto de un robot colaborativo, facilitando una comunicación hombre-máquina.

## 4-Integración y evaluación de la eficiencia del sistema robótico completo: Se realizará a través de experimentos in-vitro, es decir en un entorno controlado fuera de un organismo vivo, con el propósito de obtener resultados contundentes para establecer las ventajas que representa el sistema creado frente a las técnicas utilizadas actualmente.

Desde la Universidad de Valladolid se contribuirá en el control del brazo del robot colaborativo y en el diseño y desarrollo de un modelo de visión durante el preoperatorio con el fin de planificar con la mayor exactitud posible las diferentes trayectorias seguidas por los brazos robóticos. Es decir, se centrará en una parte del primer y una parte del segundo de los objetivos.

## 2.3 VISIÓN ARTIFICIAL

Se puede dar una definición sencilla estableciendo que, básicamente, consiste en un campo científico que aborda la forma en la que distintos



dispositivos pueden ser construidos de manera que obtengan una comprensión de alto nivel a partir de una sola imagen o de una secuencia de imágenes. Desde el punto de vista de la ingeniería, busca automatizar las tareas que el sistema visual humano puede realizar.

Dentro de las tareas de las que se ocupa la visión artificial, se encuentran los métodos para adquirir, procesar, analizar y comprender imágenes digitales de manera automática con su correspondiente extracción de datos con el objetivo de producir información útil, ya sea numérica o simbólica.

Los datos de las imágenes pueden tener naturalezas muy diferentes, desde datos obtenidos a través de secuencias de video o vistas de múltiples cámaras, hasta datos multidimensionales de un escáner médico. Lo que si comparten todas ellas, es que forman un conjunto de datos numéricos que, basados en disciplinas como la geometría o la física, dan la opción de procesar la información de entrada como lo haría el ojo humano.

### 2.3.1 Visión humana y visión artificial

Desde sus inicios a finales de la década de 1960, los sistemas de visión artificial tienen el objetivo de imitar el sistema visual humano donde la cámara realiza las funciones del glóbulo ocular, mientras que el dispositivo de procesamiento, generalmente un ordenador, se encarga de las tareas que lleva a cabo el cerebro.

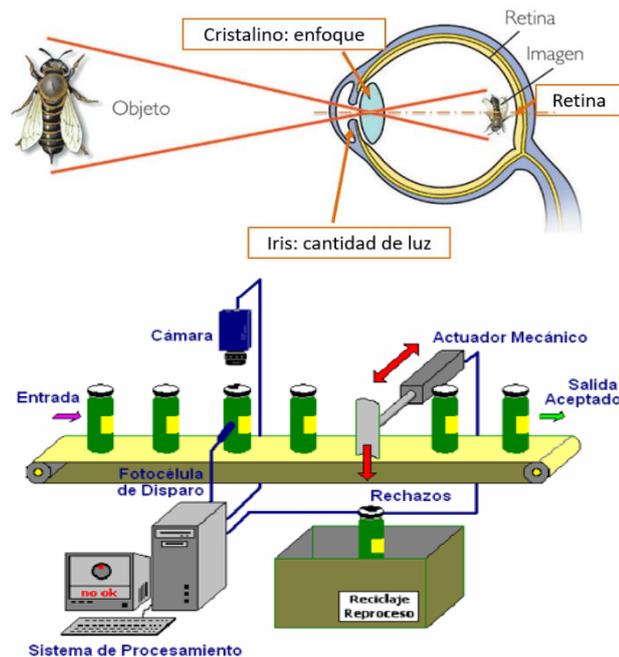


Figura 3: Sistema de visión humano frente a la visión artificial.

(Fuente: Apuntes Visión artificial – Universidad de Valladolid)

El ojo humano es uno de los órganos de los que más información se posee, en cambio, el cerebro es ese gran desconocido. Aún no se han descifrado sus sofisticados mecanismos de modo que permitan ser implantados de manera artificial. No obstante, los sistemas de visión artificial, a pesar de no tener las excelentes propiedades del ojo humano en cuanto a reconocimiento de objetos o a la adaptación ante situaciones imprevistas, pueden realizar ciertas tareas de forma más eficaz que la visión humana. Estos modernos sistemas poseen potentes características como una mayor velocidad de respuesta o la capacidad de percibir todo el espectro electromagnético. Además, las cámaras presentan ciertas ventajas como la de poder estar funcionando prácticamente sin interrupciones durante largos periodos de tiempo manteniendo en todo instante el máximo rendimiento, la eliminación de la subjetividad en el análisis o la capacidad de adaptarse y tener flexibilidad ante los procesos.

Por consiguiente, a pesar de que la visión artificial no alcance todas las características de la visión humana, es capaz de realizar múltiples tareas visuales con gran eficacia.



### 2.3.2 Aplicaciones

El rango de aplicaciones que tiene la visión artificial es inmenso y, actualmente, su implantación es cada vez más común. Las aplicaciones van desde tareas industriales, hasta tareas de investigación científica o robots que pueden comprender el mundo que les rodea.

- APLICACIONES INDUSTRIALES

El principal objetivo de estas aplicaciones es producir mejoras en el proceso productivo. Estas pueden ser reducir costes de producción o hacer que los procesos sean más rápidos y eficaces, con el objetivo de obtener un producto final de mejor calidad.

En este sector, la visión artificial afecta a aspectos tan importantes como la trazabilidad, el control de calidad, el control de posición y rotación de objetos, la seguridad industrial, el control de logística y procesos o el soporte a la producción.

- APLICACIONES CIENTÍFICAS

Otro gran sector que se beneficia de la visión artificial es el científico. Cada vez son más las aplicaciones de sistemas de visión en ámbitos como la biología o la medicina. El objetivo de estas aplicaciones es aprovechar las potentes características de estos sistemas para mejorar los distintos procedimientos científicos, los cuales requieren de elevada precisión.

A modo de ejemplo de aplicación podría plantearse el diagnóstico de enfermedades, aplicaciones de robótica quirúrgica, visualización tridimensional de órganos o aplicaciones microscópicas, como la de identificar células o contar microorganismos.

- APLICACIONES CON APRENDIZAJE

Generalmente, la mayoría de las aplicaciones de los sistemas de visión implementados en la actualidad tienen ordenadores preprogramados para resolver una tarea concreta, pero los métodos basados en el aprendizaje se están volviendo cada vez más comunes.

Estos, dan lugar a nuevas áreas de aplicación como son los vehículos autónomos. El nivel de autonomía oscila entre vehículos totalmente autónomos (no tripulados) y vehículos en los que los sistemas de visión artificial apoyan a un conductor.

Ejemplo de sistemas totalmente autónomos pueden ser los vehículos que utilizan la visión para navegación, saber dónde están o

para elaborar un mapa de su entorno. También se podrían tener vehículos aéreos no tripulados en busca de incendios forestales. Por otro lado, un ejemplo de un sistema de apoyo es el de los sistemas de alerta de obstáculos en los automóviles.

Según múltiples estudios, aseguran que el mercado de los sistemas de visión artificial va a continuar creciendo. De hecho, el último estudio realizado por *Tractica*, prevé que el mercado hardware, software y de servicios de dichos sistemas alcancen los 26.2\$ billones en el año 2025. Claramente se puede observar en la figura 4 la evolución que se espera de la visión artificial.

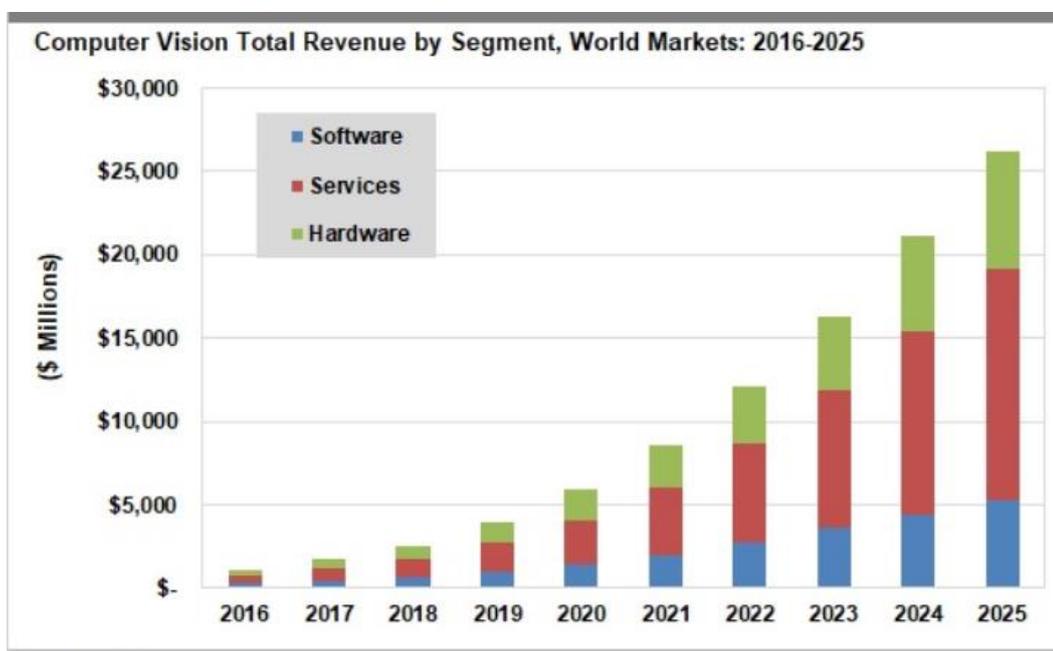


Figura 4: Previsión económica en el mercado de la visión artificial.

(Fuente: «Tratica computer vision research» [En línea]. Disponible: [https://www.tractica.com/.](https://www.tractica.com/))

### 2.3.3 Visión Artificial en la medicina

Los análisis de imágenes en medicina comienzan en la década de 1960. Comenzó con simples aplicaciones como la clasificación morfológica y el recuento de partículas o el apareamiento de cromosomas. Recientemente, la llegada de las nuevas tecnologías y su constante desarrollo, han llevado hasta métodos que exploran estructuras en tres dimensiones y cambios dinámicos de objetos.



Hoy en día es prácticamente imprescindible tener una estación multimedia de trabajo, con la que el médico tenga a su alcance herramientas para manipular imágenes, realizar mediciones y, en términos generales, extraer y recopilar información, debido a las altas exigencias que presenta la medicina actual.

Debido a esto, hoy en día, uno de los campos de aplicación más destacado de la visión artificial es la medicina o el procesamiento de imágenes médicas, caracterizada por la extracción de información de datos útiles. Las aplicaciones principales son para la visualización en dos y tres dimensiones.

Ejemplos típicos de ello son la detección de tumores, arteriosclerosis u otros cambios malignos; las mediciones de las dimensiones de los órganos, el flujo sanguíneo, aplicaciones que apoyan la investigación médica proporcionando nueva información, ya sea sobre la estructura del cerebro o sobre la calidad de los tratamientos médicos. Las aplicaciones de la visión artificial en el área médica también incluyen la mejora de las imágenes interpretadas por el ser humano para reducir la influencia del ruido.

#### 2.3.4 Componentes de un sistema de visión artificial

Un sistema de visión artificial es exitoso en función de los elementos que lo formen. Los principales componentes que lo incluyen son los siguientes:

- **ILUMINACIÓN:**

Es uno de los aspectos más críticos de las aplicaciones de visión artificial, ya que es el que proporciona las condiciones óptimas para la toma de la imagen. Si el objetivo no se ilumina debidamente puede sufrirse una pérdida de información o incluso malinterpretar los datos obtenidos. Dependiendo del objeto que se quiera captar se deberá utilizar una iluminación concreta, permitiendo que sus características sobresalgan y evitando sombras y posibles reflejos.

Por esta razón, existen numerosos modos de iluminación en función del entorno y la aplicación.

- **LENTE:**

Son las encargadas de capturar la imagen y entregarla al sensor de imagen de la cámara. Su utilización determina, en gran parte, la

calidad y la resolución de la imagen que se tome. Los sistemas de visión, principalmente, presentan lentes fijas o intercambiables.

- SENSOR DE IMAGEN:

El sensor de imagen de una cámara tiene la función de capturar la imagen que se proyecta a través de las lentes.

Un sensor es un dispositivo capaz de detectar magnitudes físicas o químicas como temperatura, intensidad lumínica, velocidad, fuerza o humedad y las transforma a variables eléctricas como la corriente o el voltaje. Concretamente los sensores de visión artificial son sensibles al espectro electromagnético.

A pesar de que fueron desarrollados antes de los 80, hoy en día los sensores de visión siguen presentando grandes retos y están en constante desarrollo.

Haciendo una analogía entre el sensor de visión y uno de proximidad, puede imaginarse una matriz de sensores de proximidad de 34x34, generando una superficie en 2 dimensiones, sobre la que se coloca una pelota. La salida del set de sensores sería algo parecido a lo que se puede ver en la siguiente figura.

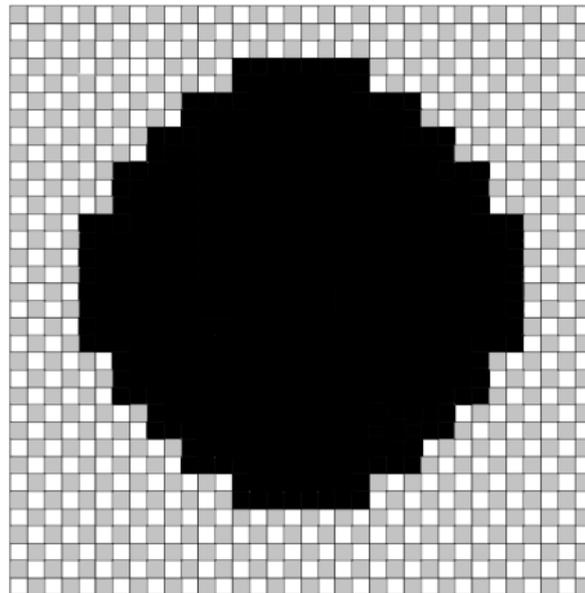


Figura 5: Representación de superficie y salida de sensores.

(Fuente: Elaboración propia)



El área oscura correspondería a los sensores activados, donde la pelota ha sido colocada. El resto de los sensores están inactivos (zona recuadrada).

A simple vista no es difícil para el cerebro humano clasificar la imagen y proporcionarnos una descripción detallada de ella. Sin embargo, todo lo que se tiene tras la medición de los 1156 sensores, es un conjunto de números con los estados de activación a los cuales se tendrá que aplicar algoritmos y métodos matemáticos, con sus respectivos tiempos de ejecución, para poder calcular y discernir que están captando en conjunto.

Una imagen pequeña, de 320x240 píxeles, es equivalente a tener 76.800 sensores. Si se habla de video, que no es más que una secuencia de imágenes representando movimiento, se tendría que trabajar con 76.800 sensores por fotograma, y suponiendo una cadencia de 30 fotogramas por segundo, consistiría en tener que leer y procesar el estado de más de dos millones de sensores por segundo.

Esto explica el enorme requerimiento computacional en términos de volumen y cantidad de datos con los que los sistemas de visión artificial tienen que lidiar. Es por ello por lo que existen multitud de sensores de imagen, cada uno especializado y optimizado para una serie de aplicaciones y requerimientos concretos. En función de estos, se utiliza el que más las satisfaga. Es por ello por lo que existen múltiples tipos de cámaras:

- Matriciales
- Lineales
- De alta velocidad
- 3D
- Infrarrojas/Térmicas
- Espectrales/Multiespectrales

### 2.3.5 Funcionamiento

Independientemente de la aplicación a la que estén destinados los sistemas de visión artificial, todos ellos tienen unas etapas de funcionamiento comunes. Estas etapas consisten en adquirir, procesar, analizar y comprender.

Una imagen digital consiste en un conjunto de números organizados en una matriz. Cada elemento de la matriz se denomina píxel y posee un valor discreto, como puede ser el nivel de gris, la luminosidad o la luminancia.

La obtención de esta se realiza a través de la digitalización, es decir, pasar de una imagen continua a una imagen digital. Para ello primero se realiza un muestreo de la imagen continua y posteriormente una cuantización en cada punto, obteniendo el valor de la característica deseada. Este proceso puede verse en la figura 6.

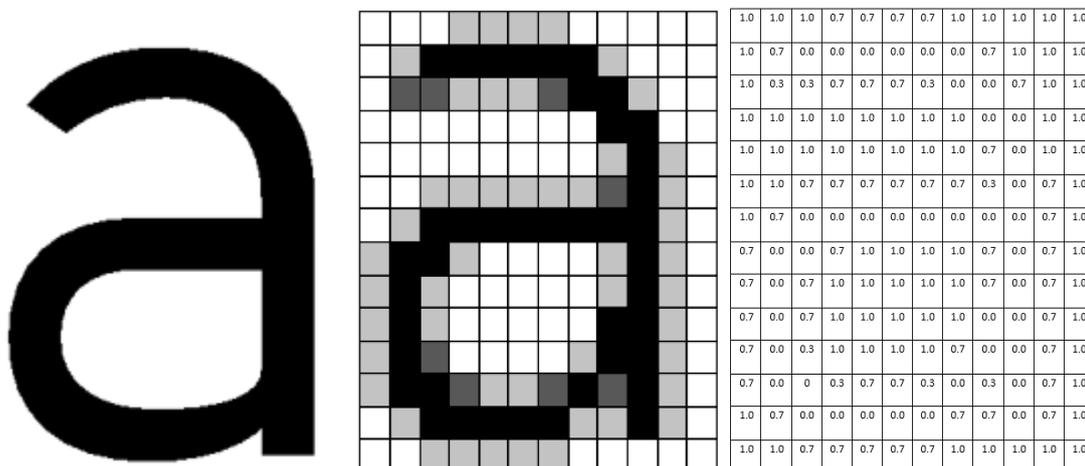


Figura 6: Representación del proceso de digitalización y cuantización.

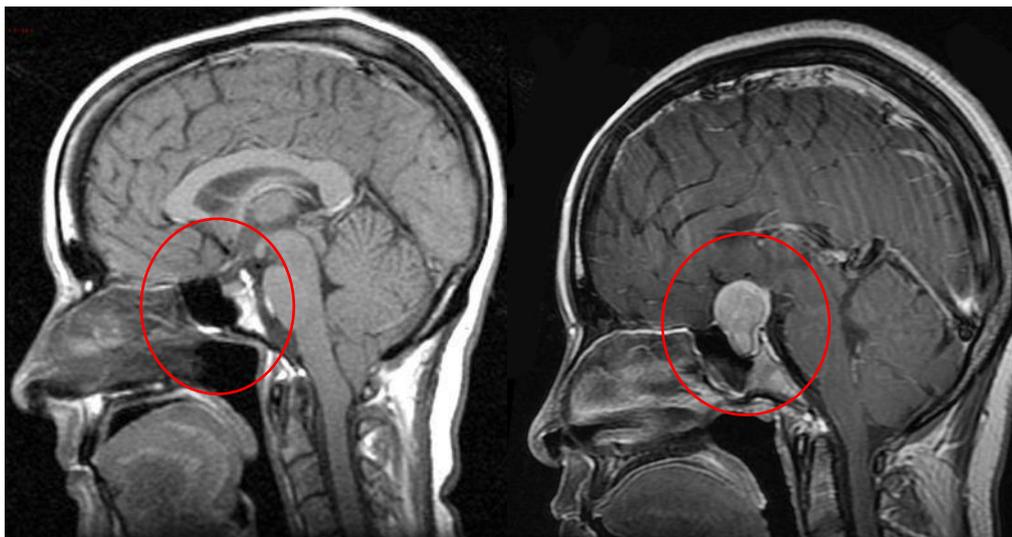
(Fuente: Elaboración propia)

Una vez obtenida la imagen digital comienza el procesado. Este lo componen los diferentes mecanismos por los cuales se extrae información de una imagen digital. En función de los diferentes procesados que se apliquen, que pueden ir desde transformaciones geométricas, segmentaciones o simples filtros, hasta análisis de formas y reconocimiento de objetos, se procede a analizar los datos. Una vez hecho el análisis, se procede a mostrar al usuario los resultados con el fin de que este los evalúe y los comprenda.

## 2.4 OPERACIONES DE CIRUGÍA TRANSESFENOIDAL

La región selar es una zona donde se producen diferentes lesiones, siendo las tumorales las más comunes, concretamente el tumor de hipófisis. Este último término agrupa una serie de patologías misceláneas en la silla turca, donde el adenoma hipofisario es el tumor más frecuente e importante, representando el 90% de las lesiones selares. Para tener una idea más general, hágase alusión a lo que Armengot et al. (2010) describieron: “Los adenomas pituitarios son los tumores intracraneales más comunes, representando aproximadamente el 15% del total” (p.26). Estos adenomas corresponden a

lesiones tumorales de la glándula pituitaria o hipófisis, la cual se sitúa en la base del cráneo, concretamente en la silla turca, con un tamaño como el de un guisante. Se manifiestan clínicamente por compresión de las estructuras vecinas y por hiper o hipofunción hormonal. En función de sus características, su manejo puede variar desde la observación o la manipulación médica, hasta la cirugía.



*Figura 7: IRM del plano sagital de un paciente sano (izquierda) frente a uno con una lesión tumoral (derecha).*

*(Fuente: Danqui C., et al. (2014). "Pituitary adenoma with gangliocytoma: Reporto of two cases". Oncology Letters, 8(2), 781-784.)*

La citada ubicación de la hipófisis se encuentra en una posición profunda y centrada de la base del cráneo con un difícil acceso. En vista de ello, inicialmente cuando la solución debía ser quirúrgica, el abordaje planteado era a través de la vía transcraneal, teniendo que manipular el cerebro y suponiendo unas intervenciones con altas tasas de mortalidad. Además, son intervenciones donde la precisión de los movimientos toma un papel muy importante debido a la cercanía de estructuras anatómicas críticas a la zona de trabajo, como el quiasma óptico y la arteria carótida. Otro factor muy importante a tener en cuenta es el desplazamiento que puede sufrir el cerebro debido al cambio de presiones que se provoca al perforar el cráneo. Por estos motivos, se han buscado otros caminos o métodos quirúrgicos de acercamiento extracraneal. Se han valorado vías sublabiales, a través del septum nasal o vías que atravesaban el foramen etmoidal, sin embargo, la vía por la que habitualmente se realizan, desde su implementación en 1987, es a través de los abordajes endoscópicos transesfenoidales endonasales, cruzando el seno esfenoidal de

la base del cráneo. Debido a la profundidad a la que se encuentra la silla turca, es imprescindible un buen sistema de visión e iluminación. En vista del gran desarrollo de la endoscopia, se ha logrado hacer de estos abordajes, antiguamente muy cruentos, unas intervenciones más delicadas con una mayor visualización y una disminución en las complicaciones.

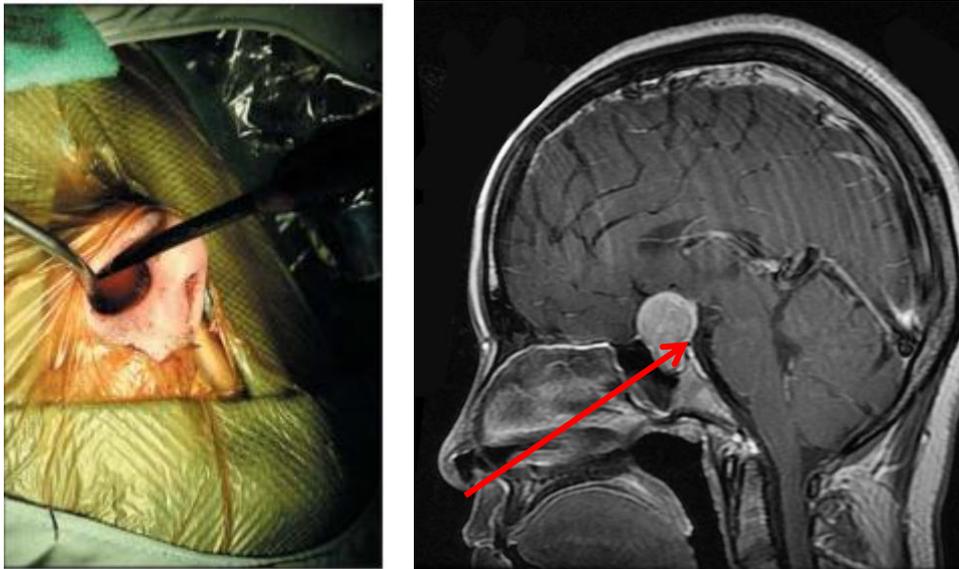
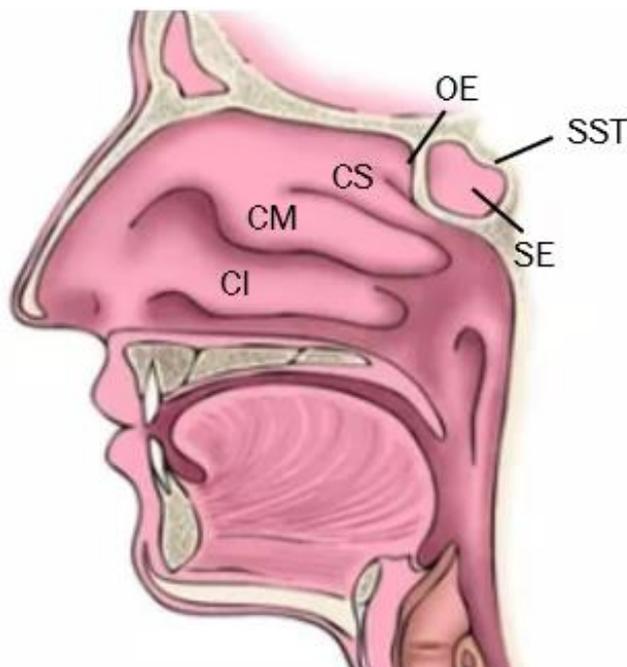


Figura 8: Abordaje nasal (izquierda) y trayectoria para acceso transesfenoidal a la hipófisis (derecha).

(Fuente: (izquierda) Kassam A. et al. (2007). *Endoscopic Expanded Endonasal Approaches to Caudo-Ventral Skull Base*. *Acta Otorrinolaringológica: Cirugía endoscópica avanzada de base de cráneo y espacios paranasales*, 58(1), 14-31. (derecha) Danqui C., et al. (2014). "Pituitary adenoma with gangliocytoma: Report of two cases". *Oncology Letters*, 8(2), 781-784.)

#### 2.4.1 Procedimiento quirúrgico

Explicando la intervención de manera sencilla se puede decir que, inicialmente se identifica cual es la fosa nasal de mayor amplitud. Una vez detectada, se introduce el endoscopio de manera gradual por ella y se procede a realizar un estudio de la cavidad nasal, dislocando lateralmente el cornete medio (CM) y el superior (CS) con el objetivo de localizar el ostium esfenoideal (OE), el cual presenta una abertura hacia el seno esfenoideal (SE).



LEYENDA: (CI) Cornete inferior, (CM) Cornete medio, (CS) Cornete superior, (OE) Ostium esfenoïdal, (SE) Seno esfenoïdal (SST) Suelo de la silla turca.

Figura 9: Cavidad nasal.

(Fuente: Elaboración propia).

Posteriormente con ayuda de un laminotomo se procede a la apertura de la pared anterior del seno esfenoïdal desde el ostium esfenoïdal (OE) hasta que quede expuesto su interior, con el objetivo de obtener una amplia apertura hacia él. Este proceso se repetirá desde la otra cavidad nasal. Una vez se acceda al interior del seno esfenoïdal, se podrá visualizar el suelo de la silla turca, a la cual con ayuda del laminotomo y el taladro se le extirpará un fragmento, permitiendo de esta forma, el acceso a su interior. Una vez se consiga acceder, con ayuda de diferentes pinzas, cucharillas, aspiradores, etc., se extraerá progresivamente el tumor. Además, la presión intracraneal facilitará su extirpación completa.

Una vez ha concluido la extracción del tumor, se comprueba a través de las lentes de 30 a 45 grados del endoscopio posibles residuos tumorales que aun residan en el interior y no hayan sido identificados. Posteriormente se realiza una hemostasia, es decir, se detienen o paran cualquier tipo de hemorragias que se hayan producido.

Una vez se asegure que el tumor ha sido extirpado en su totalidad, se procede al cierre de las aperturas selares. Para ello se hace uso de diferentes



materiales como pueden ser grasa, cola biológica, hemostáticos o parches entre otros.

## 2.4.2 Ventajas

Este tipo de abordaje quirúrgico presenta diversas ventajas frente a los demás métodos. La primera y más importante es facilitar un acceso directo y rápido a la silla sin turca sin tener que atravesar el cráneo en ningún momento. Esto hace que se reduzca drásticamente la alta tasa de mortalidad asociada a la craneotomía. Hoy en día, la mortalidad quirúrgica asociada a las intervenciones transesfenoidales está por debajo del 1%.

Por otro lado, el hacer uso del endoscopio, proporciona una mejor definición y una amplia perspectiva de las citadas estructuras anatómicas de la cavidad nasal, la región selar y la silla turca.

Actualmente se considera la técnica menos agresiva, la cual, además, no requiere taponamiento nasal. Permitiendo así, acortar la estancia de los pacientes en el hospital a unos pocos días.

## 2.5 IMÁGENES DE RESONANCIA MAGNÉTICA

El aumento de los estudios con imágenes médicas que se realizan en la actualidad, especialmente el trabajar con imágenes de resonancia magnética (IRM), ha facilitado enormemente la detección de infinidad de lesiones.

Una IRM consiste en una técnica de imagenología no invasiva y altamente versátil, utilizada en radiología dando lugar a la formación de imágenes de la anatomía y los procesos fisiológicos del cuerpo. Los escáneres de resonancia magnética, como el que se puede ver en la figura 10, utilizan fuertes campos magnéticos junto ondas de radio. El objetivo es que los protones (átomos de hidrogeno) excitados en los tejidos que contienen moléculas de agua creen una señal que se procese para posteriormente generar las imágenes médicas. El característico contraste entre los diferentes tejidos se debe a la velocidad a la que los átomos excitados regresan a su estado de equilibrio. En algunos casos, para obtener unos mejores resultados, se pueden administrar agentes de contraste exógenos al paciente.

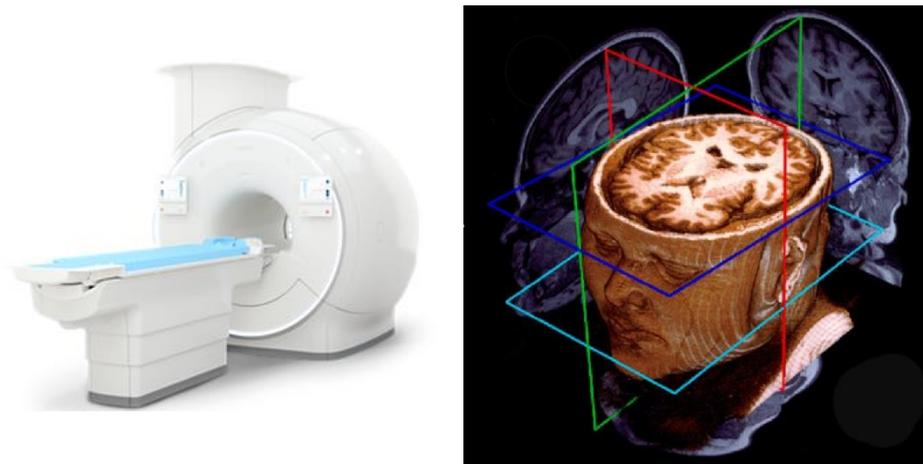


Figura 10: Scanner de resonancia magnética y representación 3D del proceso de obtención de IRMs.

(Fuente: (izquierda) «Philips Imaging System» [En línea]. Disponible: <https://www.philips.es/healthcare/>. (derecha) «Wolfson Brain Imaging Centre» [En línea]. Disponible: <https://www.wbic.cam.ac.uk/>.)

Existen dos tipos de secuencias de IRM, denominadas T1 y T2. La diferencia que existe entre ellos reside en la dirección en la que los tejidos vuelven a su estado de equilibrio.

La secuencia T1 recibe el nombre de tiempo de relajación spin-red o longitudinal. En esta secuencia se mide el tiempo, en ms, que tarda la magnetización longitudinal en recuperar el 63% de su estado de equilibrio después de interrumpir el pulso de radiofrecuencia. Este tipo de ponderación es útil para obtener información morfológica. En este tipo de secuencias, actualmente, se utilizan sustancias como el gadolinio para mejorar el contraste y aportar una mayor sensibilidad a las imágenes obtenidas.

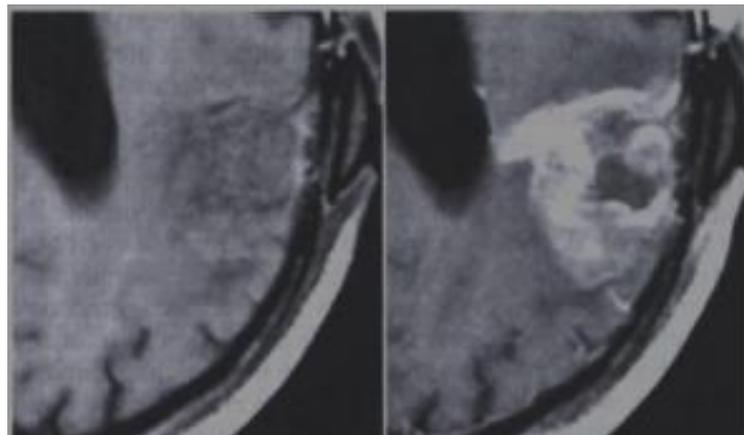


Figura 11: Mejora por aporte de gadolinio.

(Fuente: Pons L.M., Cruz A., Álvarez I., Daudinot Y. y Salomón J. (2012). Caracterización del uso de gadolinio en pacientes con indicación de resonancia magnética. MEDISAN 2012, 16(4): 498-503.)

La secuencia T2, denominada también tiempo de relajación spin-spin o transversal, mide el tiempo, en ms, que tarda la magnetización transversal en descender en un 63% de su fuerza máxima. Este tipo de secuencia es útil para detectar cuadros patológicos.

En función de la ponderación que se utilice en la toma de las IRM, la comprensión de la escala de grises obtenida en cada imagen es diferente. En la siguiente tabla se presentan de manera resumida estas diferencias.

<b>ELEMENTO</b>	<b>SECUENCIA T1</b>	<b>SECUENCIA T2</b>
Agua	Gris muy oscuro	Blanco
Grasa	Blanco	Gris oscuro
Aire	Negro	Negro
Tejido colágeno denso	Negro	Negro
Hueso cortical	Negro	Negro
Sangre en periodo subagudo	Blanco	Blanco
Sangre en periodo remoto	Negro	Negro
Sustancia gris	Gris oscuro	Gris claro
Sustancia blanca	Gris claro	Gris oscuro
Interior de vasos con alto flujo	Negro	Negro
Calcificaciones	Negro	Negro

Tabla 1: Diferenciación del nivel de gris en cada secuencia

A modo de ejemplo, se pueden ver también en la figura 12, las claras diferencias entre una IRM con secuencia T1 y T2.

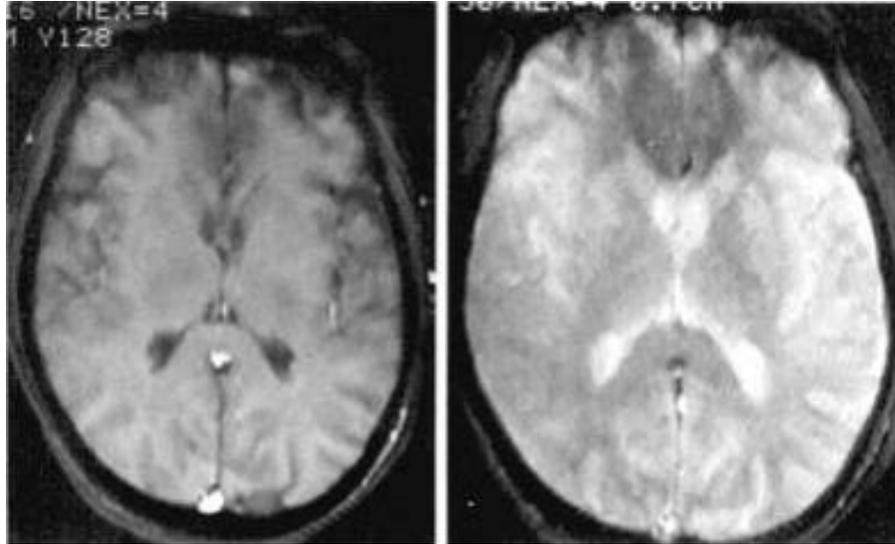


Figura 12: IRM secuencia T1 (izquierda) y T2 (derecha).

(Fuente: Buxton R., Edelman R., Rosen B., Wismer G. y Brady T. (1987). "Constrast in Rapid MR Imaging: T1 and T2 Weighted Imaging". *Journal of Computer Assisted Tomography*, 11(1): 7-16.)

La IRM presenta grandes ventajas frente a otros métodos de captura de imágenes. Entre ellas su capacidad de generar imágenes de cortes o planos en cualquier dirección del espacio, su elevada resolución de contraste y el control sobre ella o la ausencia de efectos nocivos.



## 3 SOFTWARE Y SISTEMA DE VISIÓN ARTIFICIAL UTILIZADOS

---

### 3.1 INTRODUCCIÓN AL CAPÍTULO

El procesamiento de imágenes en tres dimensiones ha sido un mercado de rápido crecimiento debido al amplio abanico de posibilidades que ofrece. Por tanto, en este capítulo se explican los dos métodos puestos en práctica, haciendo especial hincapié en el de las cámaras ToF, ya que han sido las que finalmente se han utilizado.

Además, se analiza y explica el software con el que se ha trabajado para toda la parte correspondiente al procesamiento de imágenes.

### 3.2 OBTENCIÓN DE MODELOS TRIDIMENSIONALES

Parte del trabajo pretende obtener un modelo tridimensional de la cabeza del paciente a través de un sistema de visión artificial. Para llevarlo a cabo se han utilizado dos métodos diferentes, los cuales implican la utilización de dos tipos de cámaras, las cuales se exponen a continuación.

#### 3.2.1 Cámara estereoscópica

Este tipo de cámaras pretenden simular la visión binocular humana, y, por lo tanto, pretenden tener la capacidad de generar imágenes tridimensionales. El ser humano genera una imagen con cada uno de los ojos, que después el cerebro se encarga de mezclar y procesar creando una visión en tres dimensiones. Estas cámaras pretenden realizar la misma función, sustituyendo los ojos por dos objetivos y el cerebro por un dispositivo electrónico capaz de procesar la información y generar la imagen 3D. Este proceso es conocido como fotografía estereoscópica.

La separación intraocular humana corresponde aproximadamente a una distancia promedio de 62 milímetros. Sin embargo, con las cámaras estereoscópicas se puede colocar ambos objetivos en una posición más estratégica en función de la aplicación, con el fin de conseguir las metas buscadas de la mejor manera posible.



Hay tres tipos de sistemas estereoscópicos principales que se pueden utilizar para la obtención de imágenes tridimensionales:

- Cámara estéreo con dos objetivos
- Dos cámaras iguales con la captura sincronizada
- Una sola cámara con desplazamiento

La primera de las opciones puede llevarse a cabo de manera más económica que con cualquiera de las dos últimas. El único inconveniente es que se pierde la capacidad de obtener imágenes 3D en movimiento, pero funcionan correctamente con elementos estáticos.

Concretamente en este trabajo se pone en funcionamiento un sistema de visión estéreo utilizando dos cámaras iguales.

En este tipo de sistema estereoscópico implementado es muy importante que la captura de imágenes de ambos objetivos sea sincronizada, si se pretenden obtener resultados válidos. Con cada cámara se genera una imagen en dos dimensiones que, tras su procesado, se obtiene información sobre la profundidad de cada punto, de este modo se puede decir que se obtiene una imagen 3D ya que se conocen las coordenadas tridimensionales de cada uno de los puntos.

#### *Principios de funcionamiento de cámara estéreo*

La información tridimensional se calcula a partir de un cálculo geométrico. Reduciendo la complejidad y bajo la suposición de que, en este caso, las imágenes se formen delante del objetivo, va a explicarse de manera esquemática como se calcularía la distancia a un punto a partir de la visión estereoscópica. Esta proporciona dos imágenes desde dos posiciones diferentes de la misma escena, donde la distancia entre las cámaras viene definida por  $D_{12}$  como puede apreciarse en la figura 13. También puede verse como la luz reflejada por un elemento (E) en el espacio se transmite y captura por las dos cámaras, representadas por cubos en la parte inferior de la figura.

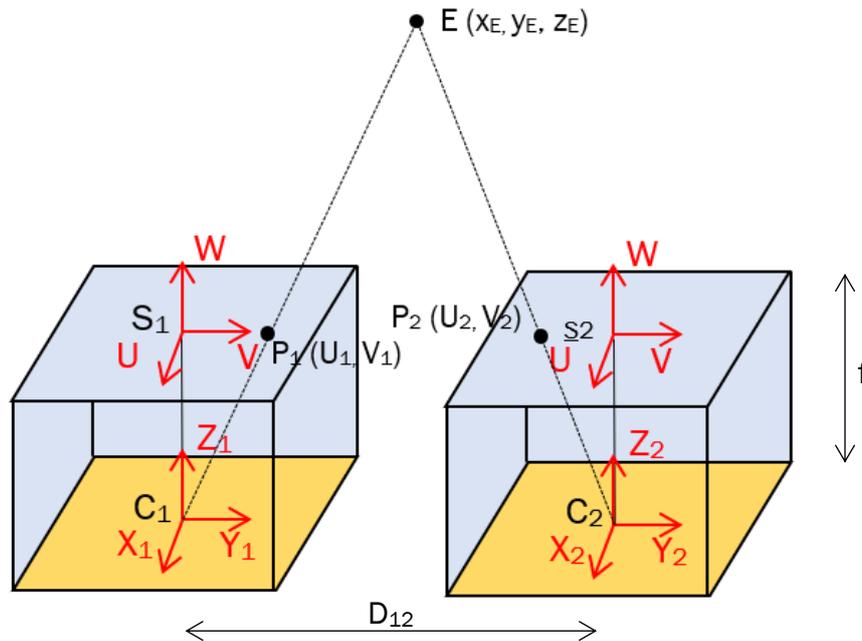


Figura 13: Esquema de funcionamiento de cámaras estereoscópicas.

(Fuente: Elaboración propia).

El origen del sistema de coordenadas de la cámara se sitúa en el centro óptico del objetivo de esta, representado con un  $C_1$  en la figura. Como se ha comentado previamente, para simplificar los cálculos, a diferencia de lo que ocurre en las cámaras convencionales, las imágenes se imprimen delante del centro óptico del objetivo. Los ejes  $U$  y  $V$  del sistema de coordenadas de la imagen están en la misma dirección que los ejes  $X$  e  $Y$  del sistema de coordenadas de la cámara, respectivamente.

Siendo el punto  $E$  un punto en el espacio, sus puntos de imagen correspondientes son  $P_1(U_1, V_1)$  y  $P_2(U_2, V_2)$  respectivamente en el plano izquierdo y derecho. Por otro lado, suponiendo que las dos cámaras se encuentren posicionadas en el mismo plano y de acuerdo a las relaciones trigonométricas, se puede definir lo siguiente:

$$V_1 = f \frac{Y_E}{Z_E}, \quad V_2 = f \frac{Y_E - D_{12}}{Z_E}, \quad U_1 = U_2 = f \frac{X_E}{Z_E}$$

Donde las coordenadas de  $P$  están referidas al sistema de coordenadas de la cámara izquierda,  $D_{12}$  es la distancia entre los centros de las cámaras y  $f$  es la distancia focal.

La disparidad visual se define como la diferencia en la ubicación del punto en la imagen de un determinado punto del mundo adquirido por dos cámaras. Numéricamente corresponde a lo siguiente:

$$d = V_1 - V_2 = f \frac{D_{12}}{Z_E}$$

Por lo tanto, una vez conocidas las coordenadas de los puntos en la imagen, además de los parámetros de dos cámaras, se puede determinar la coordenada 3D del punto como:

$$X_E = \frac{D_{12} \cdot U_1}{d}$$

$$Y_E = \frac{D_{12} \cdot V_1}{d}$$

$$Z_E = \frac{D_{12} \cdot f}{d}$$

### 3.2.2 Cámara ToF

El acrónimo ToF, viene del inglés “Time of Flight”, cuya traducción significa “tiempo de vuelo”.

Una cámara ToF, es un sistema que trabaja con imágenes de rango. Implementa técnicas de tiempo de vuelo para resolver la distancia que existe entre la cámara y la escena para cada punto de la imagen. Su funcionamiento es sencillo, una fuente de luz integrada en la cámara emite una onda infrarroja (IR) que es reflejada por la escena y es capturado por un sensor. Este procedimiento se puede ver en la figura 14. Dependiendo de la distancia a la que estén situados los elementos, la onda de luz capturada se retrasa en fase en comparación con la onda de luz original emitida. De esta manera midiendo la diferencia de fase, se puede estimar la distancia a cada píxel. La fuente de luz que utilizan estos sistemas puede estar basada en láser o en luces LED.

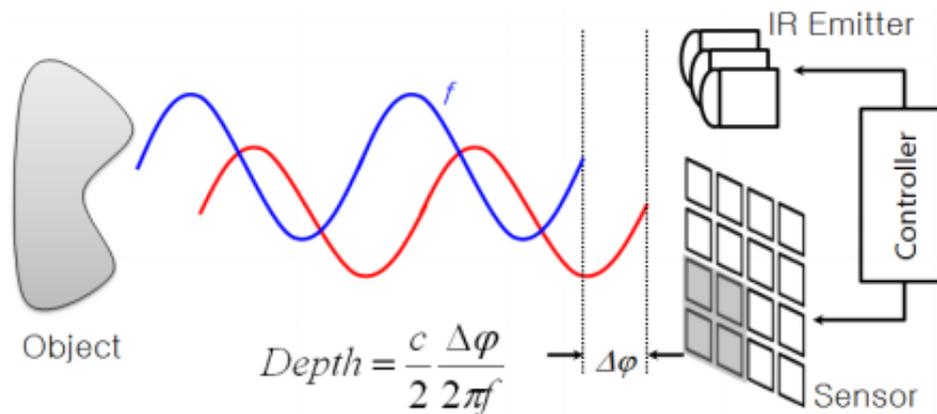


Figura 14: Principio de funcionamiento de cámara ToF.

(Fuente: Hansard M., Lee S., Choi O. y Horaud R. (2012). "Time of Flight Cameras: Principles, Methods, and Applications". Grenoble: Springer.)

### Principios de funcionamiento de cámara ToF

Ya se ha mencionado que la fuente de luz de la cámara ilumina por un periodo muy corto la escena y se refleja en los objetos, siendo el objetivo de la cámara recoger la luz reflejada. Dependiendo de la distancia a la que estén situados los diferentes elementos de la escena, la luz entrante experimentara un retraso u otro. Como la luz viaja a una velocidad aproximada de  $c = 300.000.000.000$  metros por segundo, estos retrasos serán muy cortos. Pueden ser calculados con la siguiente ecuación:

$$t_{retraso} = 2 \cdot \frac{D}{c}$$

Siendo D la distancia al objeto y c la velocidad de la luz.

Sin embargo, suele ser el ancho de pulso de la unidad de iluminación IR de la cámara la que determina el rango máximo de distancias que la cámara puede manejar.

$$D_{max} = \frac{1}{2f} \cdot c$$



Se puede comprobar de la ecuación anterior que, a mayor frecuencia de modulación, mayores distancias serán capaces de detectar las cámaras. El hecho de que la luz deba ser modulada a altas frecuencias, hace que solo se puedan implementar unidades de iluminación de LEDs o diodos laser.

El sensor de imagen, que puede ser analógico o digital, está compuesto por una cantidad determinada de elementos fotosensibles, capaces de convertir la luz entrante en un valor de corriente. Para reducir el ruido que este pueda recibir, se implementa en la óptica de la cámara un filtro pasa-banda óptico que solo deja pasar la luz con la misma longitud de onda que la de la unidad de iluminación IR. Una vez transcurrida la exposición, se procede a leer el valor generado por cada píxel. El cálculo de la diferencia de fase es posible gracias a la relación entre cuatro valores de carga eléctrica diferentes, como se puede ver en la figura 15. Estas cuatro señales de control tienen un retardo de fase de 90 grados entre sí que determinan la colección de electrones IR capturados. De esta manera, con los 4 valores se estima la diferencia de fase,  $t_d$ , según la siguiente ecuación.

$$t_d = \arctan\left(\frac{Q_3 - Q_4}{Q_1 - Q_2}\right)$$

Donde desde  $Q_1$  a  $Q_4$  se representa la cantidad de carga eléctrica para las señales de control desde  $C_1$  a  $C_4$ . Por tanto, implementando esta diferencia de fase en la ecuación donde se calculaba la distancia máxima que era capaz de medir la cámara, se puede calcular la distancia correspondiente a cualquier objeto siguiendo la siguiente ecuación.

$$d = \frac{c}{2f} \frac{t_d}{2\pi}$$

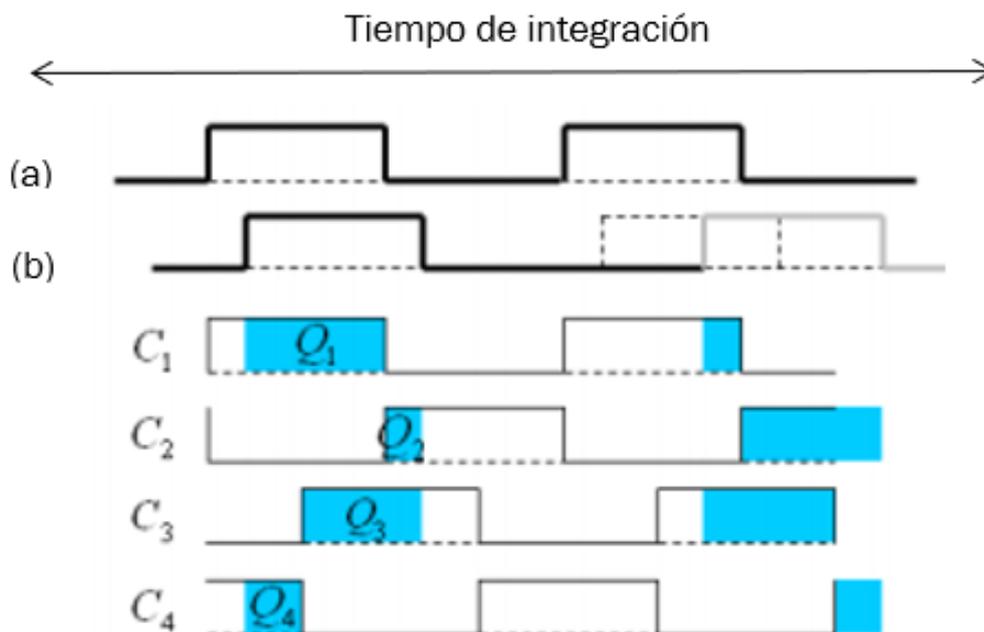


Figura 15: Profundidad calculada estimando la diferencia de fase entre la señal IR radiada (a) y la reflejada (b).

(Fuente: Hansard M., Lee S., Choi O. y Horaud R. (2012). "Time of Flight Cameras: Principles, Methods, and Applications". Grenoble: Springer)

Es verdad que la resolución del mapa de profundidad generado por las cámaras ToF es relativamente baja en comparación con las cámaras de dos dimensiones estándar. Además, los mapas de profundidad capturados a menudo se ven corrompidos por cantidades significantes de ruido. Sin embargo, en comparación con otros métodos de escaneo tridimensional, las cámaras de tiempo de vuelo funcionan mucho más rápido, proporcionando hasta 160 imágenes por segundo.

Por otro lado, son sistemas extremadamente simples sin necesidad de elementos adicionales para obtener su máximo rendimiento y, además, son sistemas provistos de un eficiente algoritmo de distancias que no tiene la necesidad de un alto coste de procesamiento.



### 3.3 SOFTWARE UTILIZADO

#### 3.3.1 Ubuntu - Linux

Ubuntu pertenece a una distribución de Linux basada en la arquitectura Debian GNU/Linux. Es un sistema operativo de código abierto compuesto de múltiple software bajo una licencia de código libre, haciendo excepciones en el caso de varios controladores privativos. Se mantiene gracias a la empresa británica *Canonical* y la comunidad de desarrolladores. El sistema puede ofrecerse de manera gratuita por medio de servicios vinculados al sistema operativo (SO) y la venta de soporte técnico que hace la citada empresa patrocinadora. Está orientado de manera que su uso sea sencillo y proporcione una mejor experiencia al usuario, además, el hecho de ser un SO libre y gratuito hace que los desarrolladores de la comunidad aporten mejoras al sistema.

El primer lanzamiento de Ubuntu fue en 2004. Surgió como una alternativa al proyecto Debian, con el principal objetivo de realizar una distribución más fácil de usar y entender por el usuario final, corrigiendo varios errores presentes en Debian. Desde entonces, Ubuntu ha dado un enfoque muy importante a la importancia del software y su mejora continua. Tal es el caso que cada seis meses se publica una nueva versión de Ubuntu, al cual incluye un soporte por parte de Canonical de nueve meses. Cabe destacar que cada dos años se liberan versiones LTS (Long Term Support), las cuales reciben soporte durante un periodo de cinco años.

Tales han sido los desarrollos de software que Canonical y la comunidad han realizado que, Ubuntu ha creado su propia interfaz de usuario, denominada *Unity*, y ha conseguido extenderse más allá de los SO de escritorio o servidores. Actualmente este SO puede ser integrado en dispositivos tales como tabletas, televisores y teléfonos.

La versión utilizada en el proyecto es la 18.04, la cual tiene asociado el nombre clave de "*Bionic Beaver*", que se traduce como "castor biónico" y este se muestra de manera predeterminada sobre el escritorio de dicho SO. Es una versión LTS lanzada en 2018 y, por lo tanto, con soporte de cinco años hasta el 2023. Haciendo referencia a lo que Larabel, M. (2018) escribió: "El cambio de Ubuntu 18.04 no es espectacular si se migra desde la 17.10, pero sí supone un gran cambio si se actualiza desde Ubuntu 16.04 LTS. Esta versión cuenta con el compilador GCC 7, descarta Unity 7 en favor de GNOME como shell por defecto, además de diversas actualizaciones de paquetes".



Para su instalación hace falta cumplir una serie de requisitos mínimos recomendados para permitir ejecutarlo. Estos requerimientos se muestran a continuación:

- Procesador x86 a 700 MHz.
- Memoria RAM de 512 Mb.
- Disco duro de 5 GB.
- Tarjeta gráfica y monitor capaz de soportar una resolución de 1024x768.
- Lector de DVD o puerto USB.
- Conexión a internet (No necesaria, pero de gran utilidad).

Estas condiciones hoy en día las cumplen la mayoría de las maquinas, por lo tanto, se facilita su implantación en gran cantidad de dispositivos.

### 3.3.2 OpenCV

OpenCV, que hace referencia a “Open Source Computer Vision Library, es una biblioteca de código abierto con licencia BSD que incluye varios cientos de algoritmos de visión computarizada y aprendizaje por ordenador. Además, es multiplataforma, existiendo versiones para GNU/Linux, MAC OS X, Windows y Android. OpenCV fue construido para proporcionar una infraestructura común para las aplicaciones de visión artificial y para acelerar el uso de la percepción artificial en los productos comerciales.

La biblioteca cuenta con más de 2500 algoritmos optimizados. Estos algoritmos pueden utilizarse para detectar y reconocer rostros, identificar objetos, clasificar acciones humanas en vídeos, rastrear movimientos de cámara, rastrear objetos en movimiento, extraer modelos 3D de objetos, producir nubes de puntos 3D, etc. OpenCV se inclina principalmente hacia aplicaciones de visión en tiempo real y aprovecha las instrucciones MMX y SSE cuando están disponibles. Tiene interfaces C++, Python, Java y MATLAB, aunque OpenCV fue escrito inicialmente en C++ y tiene una interfaz de plantilla que funciona perfectamente con los contenedores STL.

Cabe destacar que OpenCV tiene una estructura modular, lo que significa que se incluyen varias bibliotecas compartidas o estáticas. A continuación, se muestran varios de los módulos disponibles.

- Funcionalidad básica (core): módulo compacto que define las estructuras y funciones básicas.



- Procesamiento de imágenes (imgproc): módulo que incluye filtrado, transformaciones geométricas, conversión de color, histogramas, etc.
- Análisis de video (video): incorpora estimación de movimiento y algoritmos de seguimiento de objetos.
- Calibración y reconstrucción 3D (calib3d): introduce algoritmos básicos de geometría de múltiples vistas, calibraciones de cámaras y elementos de reconstrucción tridimensional.
- Funcionalidades 2D (features2d): proporciona detectores de características destacadas, descriptores y comparadores.
- Detección de objetos (objdetect): proporciona detección de instancias sobre las clases predefinidas.
- Interfaz de alto nivel (highgui): incluye una interfaz fácil de usar para funcionalidades de usuario simples.

### 3.3.3 PCL

“Point Cloud Library”, o PCL, se puede traducir como “Biblioteca de Nube Puntos”. Este, es un proyecto abierto a gran escala para el procesamiento de imágenes en dos y tres dimensiones y las nubes de puntos. El marco de trabajo de las PCL contiene numerosos algoritmos de última generación que incluyen filtrado, estimación de características, reconstrucción de superficies, registro, ajuste del modelo y segmentación. Estos algoritmos se pueden utilizar, por ejemplo, para filtrar valores atípicos de datos ruidosos, unir nubes de puntos 3D, segmentar partes relevantes de una escena, extraer puntos clave y calcular descriptores, y crear superficies a partir de nubes de puntos y visualizarlas, por nombrar algunos.

PCL figura bajo los términos de la licencia BSD y es software de código abierto. Por lo tanto, tiene una gran presencia en el uso comercial y de investigación. Además, es multiplataforma, y ha sido compilado e implementado con éxito en Linux, MacOS, Windows y Android/iOS.

Para simplificar el desarrollo, las PCL se dividen en una serie de librerías de código más pequeñas, que se pueden compilar por separado. Esta modularidad es importante para la distribución de PCL en plataformas con restricciones computacionales o de tamaño reducidas. Si se piensa en las PCL como un gráfico de librerías de código, el resultado sería el siguiente.

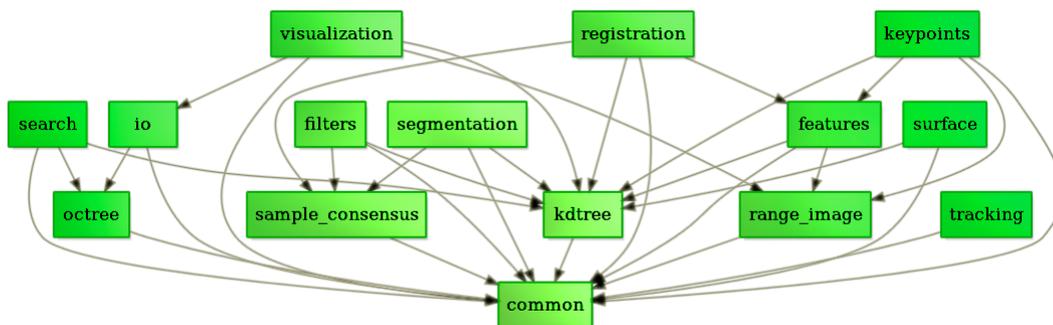


Figura 16: Librerías PCL.

(Fuente: «PCL - Point Cloud Documentation» [En línea]. Disponible: <http://www.pointclouds.org/documentation/>)

Un aspecto característico de las PCL es el modo en el que estas almacenan los datos, en las denominadas PCD. Este acrónimo hace referencia a “Point Cloud Data” y puede traducirse como “datos de la nube de puntos”. Una PCD es una estructura de datos utilizada para representar una colección de puntos multidimensionales y se utiliza comúnmente para representar datos tridimensionales. En una nube de puntos 3D, los puntos suelen representar las coordenadas geométricas X, Y y Z de una superficie muestreada. Cuando, además, la información de color está presente, la nube de puntos se convierte en 4D. Las nubes de puntos pueden ser adquiridas por sensores tales como cámaras estéreo, escáneres 3D o cámaras de tiempo de vuelo, o generadas desde un programa de forma sintética.

La estructura de los archivos con extensión “.pcd” que contienen los datos de las nubes de puntos siguen una estructura constante. Esta, consiste en dos partes claramente diferenciadas, las cuales se muestran a continuación.

1. Cabecera: Se encuentra en primer lugar e incluye información acerca de la versión del PCD (Point Cloud Data), del tamaño de la nube, del tipo de nube de puntos, del valor que puede tener cada punto y de si se trata de una nube ordenada o no. Cuando el parámetro HEIGHT sea igual a 1 se trata de una nube que se considera ordenada, mientras este parámetro sea distinto de 1, la nube se considerará desordenada. La estructura que sigue la cabecera de todos los PCD es la siguiente:

```
#. PCD v0.7 - Point Cloud Data file format
VERSION 0.7
FIELDS x y z
```



```
SIZE 4 4 4
TYPE F F F
COUNT 1 1 1
WIDTH 102087
HEIGHT 1
VIEWPOINT 0 0 0 1 0 0 0
POINTS 102087
DATA ascii
```

2. Conjunto de datos: Estos vienen inmediatamente después de la cabecera e incluyen todo el conjunto de puntos que constituyen la nube de puntos. A modo de ejemplo y complementando los datos que proporciona la cabecera descrita arriba, estos podrían ser algunos datos de la nube, los cuales proporcionan información de los campos x, y y z:

```
...
-129.18274 177.76834 -299.6243
-128.1591 177.28528 -298.81012
-127.48652 177.28528 -298.81012
-127.15948 177.76834 -299.6243
-126.83064 178.25404 -300.44296
-126.15438 178.25404 -300.44296
...
```

## 4 OBTENCIÓN NUBE TRIDIMENSIONAL POR CAPTURA

---

### 4.1 INTRODUCCIÓN AL CAPÍTULO

Esta parte del desarrollo, como del título del capítulo se puede intuir, corresponde a la generación de la nube de puntos tridimensional a partir de la toma de imágenes de la cabeza del paciente. Se ha desarrollado a través de 2 métodos diferentes con el objetivo de mejorar el resultado.

La principal diferencia entre uno y otro ha sido el tipo de sistema de cámara con el que se realizan las tomas. Debido a esto, se explicarán de manera individual cada uno de ellos.

### 4.2 IMPLEMENTACIÓN DE CÁMARAS ESTÉREO

Ya se ha citado en el capítulo anterior que el sistema estéreo está implementado por dos cámaras iguales situadas a una distancia determinada. Este sistema puede verse en la figura 17.

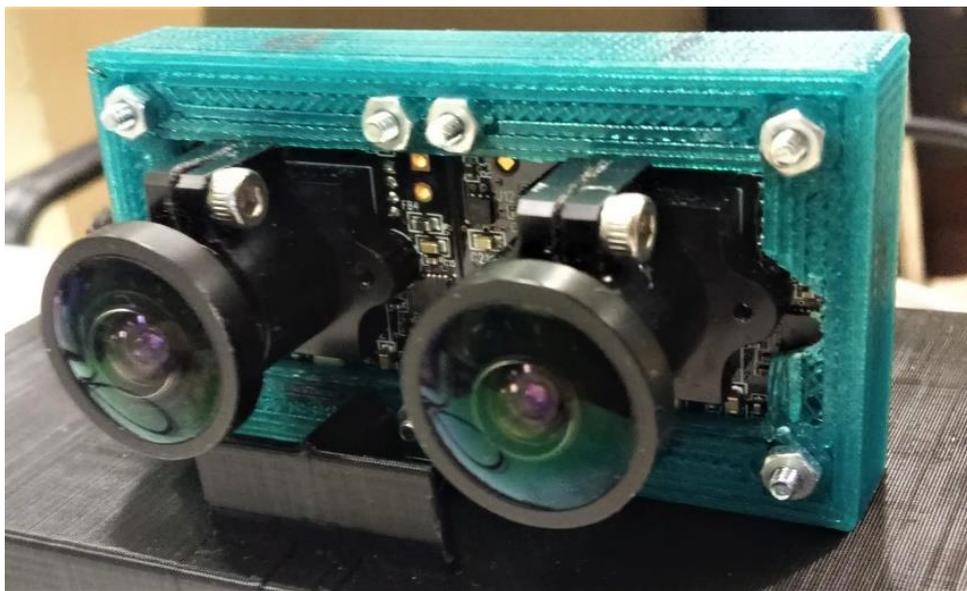


Figura 17: Cámaras estereoscópicas.

(Fuente: Elaboración propia)

Estas han sido unidas a través de un marco obtenido por impresión 3D de manera que mantengan en todo momento la misma distancia entre ellas.

Antes de comenzar a explicar el desarrollo, cabe detallar la zona de trabajo. Las primeras pruebas se han realizado sobre una imitación de cráneo humano, el cual contiene superficies más características con las que poder trabajar.

Por tanto, lo primero de todo se coloca el cráneo sobre un soporte con el fin de mantenerlo quieto en una misma posición y así, evitar posibles perturbaciones o datos no deseados a lo largo de la toma de imágenes. Una vez completado este paso, se coloca la cámara estéreo en una posición estable a una distancia aproximada de 20 centímetros, ya que se trata de la distancia óptima para obtener un objeto con definición para los ajustes de calibración que realiza el sistema de cámaras estéreo para la obtención de una imagen final a partir de las dos capturas iniciales. El citado soporte, sobre el cual se apoya el cráneo artificial, como la zona de trabajo pueden verse en la figura 18.

Una vez el cráneo y las cámaras están en posición, tiene lugar el proceso de obtención de las imágenes.

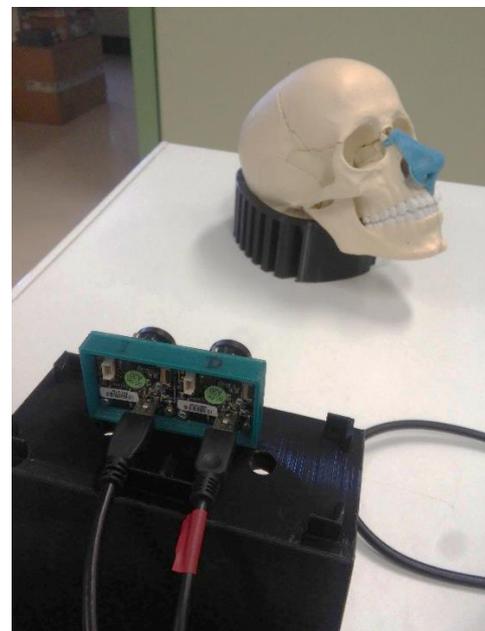


Figura 18: Cráneo artificial y espacio de trabajo.

(Fuente: Elaboración propia)

Al disponer de dos cámaras, será necesario un cable de conexión por cada una de ellas. Es importante que, para un correcto funcionamiento de los programas que se ejecuten posteriormente, se conecte primero al ordenador el cable correspondiente a la cámara de la izquierda y posteriormente el de la derecha. Esto es necesario debido a que los programas desarrollados tienen configurada la calibración de tal modo que, la primera cámara que detecte la asociará con la cámara izquierda. Es decir, la cámara izquierda se colocará en el puesto primero, mientras que la cámara derecha ocupará el segundo lugar.

Por otro lado, las cámaras han de ser calibradas de manera que, se obtengan unos resultados coherentes. Para ello se determinan los parámetros necesarios realizando capturas sobre un tablero como el que puede verse en la figura 19. Lo que se consigue con estas tomas, es calcular las matrices de calibración de cada cámara y sus coeficientes de distorsión. Para la calibración se pueden tomar tantas imágenes como se deseen.

Para llevar este proceso a cabo, en este caso, se proyecta en la pantalla del ordenador dicho tablero y se realizan diversas tomas desde ángulos y posiciones distintas, de manera que el sistema de cámaras estéreo quede calibrado para la aplicación que posteriormente implementaran.

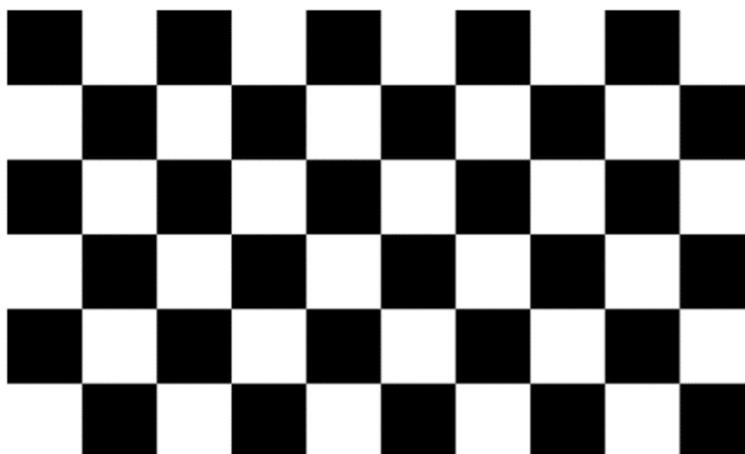


Figura 19: Tablero utilizado para la calibración.

(Fuente: Elaboración propia)

Por defecto el tablero de la figura 19 es el que está definido para hacer la calibración, el cual contiene 9x5 esquinas (10x9 cuadrados de 18 milímetros de lado).



Una vez estén todas las configuraciones cargadas, el programa se dispone a realizar la adquisición de imágenes. Saca por pantalla una ventana en la cual puede visualizarse la imagen formada por el programa, la cual se denomina “Mapa de profundidad” ya que cada punto de la imagen contiene información respecto a la distancia a ese punto. Hasta que no se pulse la tecla “ESC” del teclado, el programa irá adquiriendo continuamente imágenes, las cuales combinará y mostrará a través de la ventana. En el momento en el que se pulse sobre la imagen mostrada en la ventana, el programa almacenará la última imagen que haya combinado para su posterior procesamiento.

Es ahora, con la ayuda de las librerías PCL y con la información que proporciona la imagen creada a partir del sistema estéreo en cuanto a la profundidad de cada punto, cuando se procede a la creación de la nube de puntos con el fin de generar una imagen en tres dimensiones.

Para facilitar y agilizar el procesamiento de la nube de puntos creada, se aplicarán distintos filtros con el objetivo de reducir su tamaño, es decir, su número total de puntos y quedarnos con datos que realmente aporten información relevante.

Debido a que pueden existir posibles errores de medición en los escaneos, uno de los filtros que se aplica es el “*Statical Outlier Removal*” (SOR), el cual se encarga de eliminar mediciones ruidosas, como por ejemplo valores atípicos, de un conjunto de datos de la nube de puntos utilizando técnicas de análisis estadístico. Concretamente, con la aplicación de este filtro se pretenden eliminar muestras estratificadas, es decir, puntos generados en la nube que disten numéricamente del resto de datos.

El análisis estadístico, en este caso, consiste en el cómputo de la distribución de distancias del punto a sus vecinos. Por lo tanto, en esta observación, para cada punto, se hace un estudio de su vecindad y se calcula la distancia media desde él a todos sus vecinos. Suponiendo que la distribución resultante es gaussiana con una media y una desviación estándar, todos los puntos cuyas distancias medias estén fuera de un intervalo definido por la media de las distancias globales y la desviación estándar se pueden considerar como valores atípicos o *outliers* y se suprimirán del conjunto de datos.

Por consiguiente, para poder aplicar dicho filtro es necesario establecer el número de vecinos a analizar por cada uno de los puntos de la nube y determinar el valor del multiplicador de la desviación estándar.

Por otro lado, con el objetivo de reducir el número total de puntos de la nube, se aplica también el filtro “*Radius Outlier Removal*” (ROR), el cual está basado en componentes conexas. Este filtro excluye todos los índices de la nube de entrada que no tienen al menos un determinado número de vecinos

dentro de un cierto rango. Teniendo esto en cuenta, se define la cantidad de vecinos que debe contener cada punto dentro de un radio específico.

Para entender mejor el funcionamiento de los citados filtros se muestra en las siguientes figuras, de manera gráfica, cuál sería el resultado de su aplicación. Por ejemplo, en la aplicación del filtro ROR en la figura 21, si se especifica un vecino para un radio  $d$ , solo será eliminado de la nube el punto. Sin embargo, si se determina la necesidad de contener al menos 2 puntos dentro de la región delimitada por el radio  $d$ , el punto amarillo y el punto verde serán suprimidos de la nube.

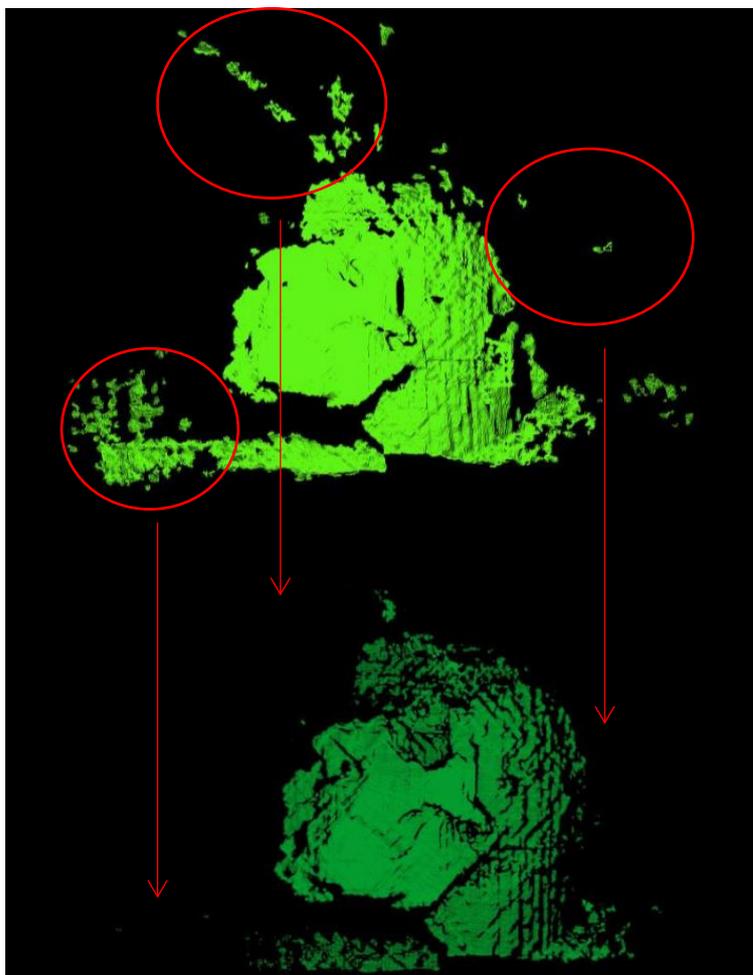


Figura 20: Eliminación de valores atípicos aplicando SOR.

(Fuente: Elaboración propia)

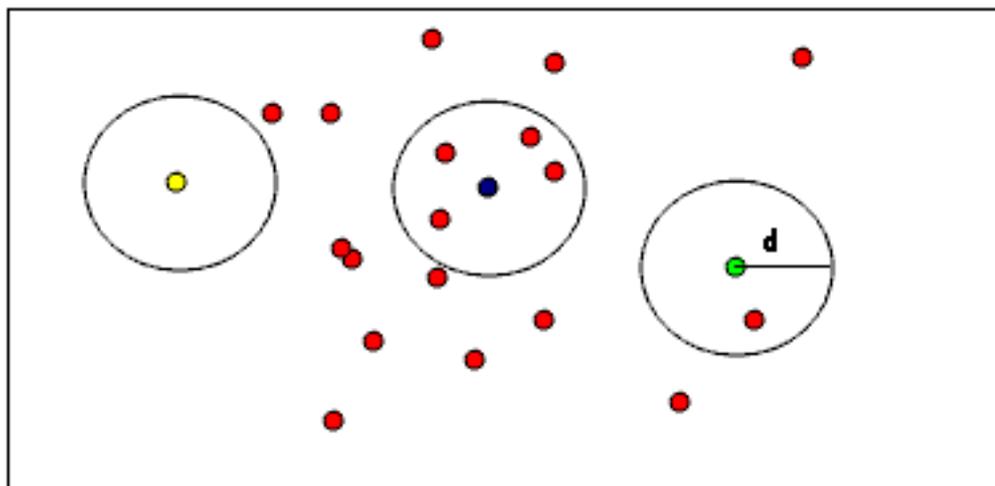


Figura 21: Eliminación basada en componentes conexas aplicando ROR.

(Fuente: «PCL - Point Cloud Documentation» [En línea]. Disponible: <http://www.pointclouds.org/documentation/>)

Una vez ejecutado el programa se obtienen las primeras nubes de puntos. En la figura 22 puede verse la clara aportación de los citados filtros. En la imagen superior se muestra la nube obtenida directamente y en la inferior, la nube filtrada. Además, la nube filtrada ha sido coloreada en función de la profundidad de los puntos que la forman. En otras palabras, los puntos más cercanos se han coloreado de un tono azul y los más lejanos de color rojo. Los puntos del intervalo situados entre el más próximo y el más lejano se colorean progresivamente con un color intermedio entre el rojo y el azul. De esta manera, la visualización de algunos detalles, como el de la dentadura, es más clara.

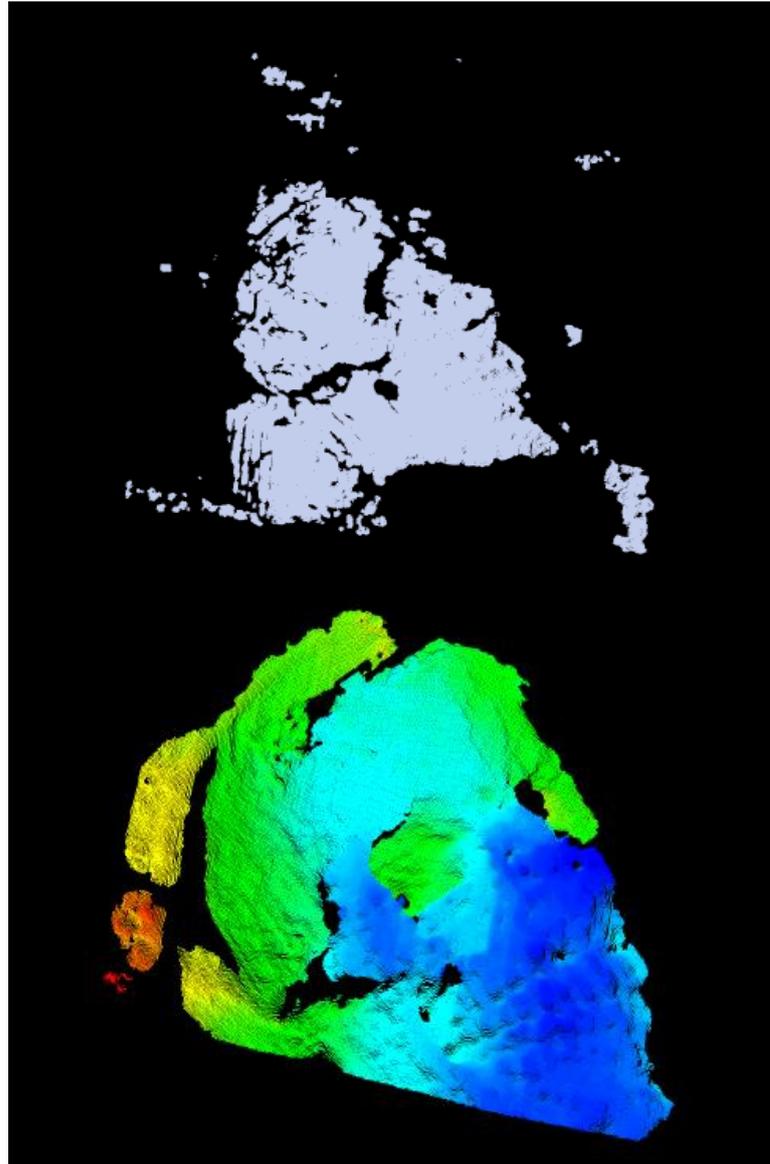


Figura 22: Nubes de puntos inicial (arriba) y filtrada (abajo) obtenidas.

(Fuente: Elaboración propia)

A pesar de que la nube de puntos generada en la figura anterior parezca lo suficientemente buena, si se amplía la imagen, de manera que pueda visualizarse la superficie creada en la nube de una forma más detallada, se puede observar cómo se manifiesta un ruido considerable sobre ella. Este defecto puede apreciarse en la siguiente figura.

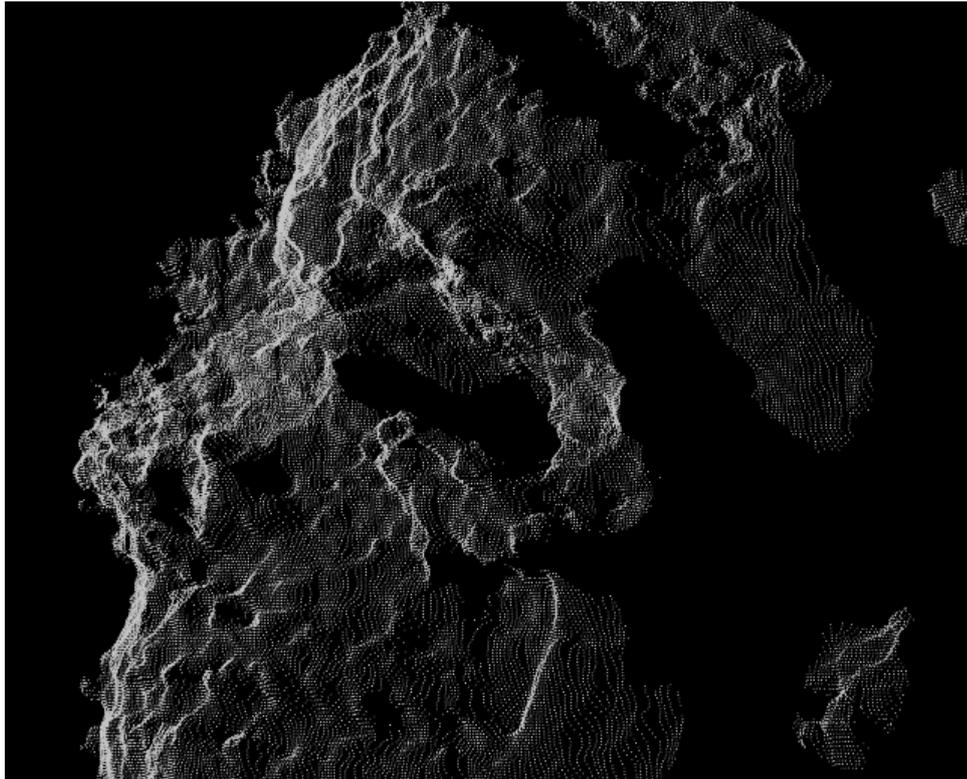


Figura 23: Visualización detallada de la superficie obtenida.

(Fuente: Elaboración propia)

Además, este ruido creado sobre la superficie es aleatorio, es decir, una misma superficie toma diferentes valores para dos tomas de imágenes realizadas bajo las mismas condiciones y desde la misma posición.

Uno de los objetivos, en esta parte del desarrollo, consiste en poder realizar una reconstrucción tridimensional a partir de un registro entre varias nubes de puntos. De manera resumida, un registro consiste en buscar las transformaciones geométricas que hay que aplicar a las nubes de puntos de manera que se puedan alinear dos o más nubes de puntos a partir de índices característicos de ellas. Sin embargo, debido al defecto aleatorio que surge sobre las superficies, hace casi imposible el poder realizar un registro entre varias nubes de puntos, ya que en cada caso se calculan unos puntos característicos diferentes y las alineaciones no son correctas.

Por tanto, en busca de una solución a este problema, se deciden buscar vías alternativas, siendo una de ellas la implementación de un nuevo sistema de visión artificial.

### 4.3 IMPLEMENTACIÓN DE CÁMARAS TOF

Concretamente se trabajará con una cámara ToF de la marca “Basler”, como la que se puede ver en la siguiente figura.



Figura 24: Cámara ToF utilizada en el proyecto.

(Fuente: «Basler ToF Camera» [En línea]. Disponible: <https://www.baslerweb.com/en/>)

Está equipada con ocho LEDs de alta potencia, los cuales forman el set de iluminación, y serán los responsables de la creación de las imágenes. Además, a diferencia del sistema anterior, estas se generarán automáticamente con información tridimensional a través de las nubes de puntos que nos proporcione la cámara.

Antes de nada, a través del programa “Pylon Camera Software Suite” proporcionado por Basler, se ajustarán diferentes parámetros de la cámara para la captura de imágenes. Estos ajustes se almacenarán en un archivo con extensión “.pfs” de manera que puedan ser leídos por los programas de se desarrollen posteriormente con el objetivo de cargar la configuración que genera la mejor nube de puntos posible bajo las condiciones en las que se trabaja.

Para proceder a la obtención de imágenes 3D, será necesario, al igual que en el caso anterior, colocar el cráneo en su soporte y ubicar la cámara en una posición estable. La distancia a la cual se posiciona el dispositivo deberá encontrarse entre los valores de máximo y mínimo nivel de alcance definidos en el archivo contenedor de las configuraciones. En la figura 25, mostrada a continuación, puede observarse cuál será la zona de trabajo.



Figura 25: Disposición de trabajo para la captura de la cámara ToF.

(Fuente: Elaboración propia)

Para poder obtener nubes de puntos de calidad es necesario que la cámara supere un periodo de “warm-up” de 20 minutos. Es decir, que se tendrá que tener la cámara alimentada, al menos, 20 minutos antes de proceder a la captura de imágenes para asegurar un funcionamiento óptimo.

Una vez superado, se orienta el cráneo con el ángulo que se desee frente al objetivo y se inicializa el proceso de captura. Inicialmente se hace una declaración de todas las variables del sistema y el programa lee y carga la configuración que se proporcione a partir del archivo con extensión “.pfs” correspondiente. Una vez queden implementados todos los parámetros, se inicializa una ventana en la cual se puede visualizar en tiempo real lo que la cámara ToF está viendo. En el momento en el que se pulse la letra “q” el programa finalizará la visualización y almacena el último conjunto de puntos generados por la cámara ToF y la convierte a una nube de puntos PCD con la cual se trabaja posteriormente.

Como ya se ha explicado anteriormente, se han ajustado las distancias máximas y mínimas donde se permite la detección de objetos. Sin embargo, el tamaño de la nube que se obtiene es fijo. Debido a esto, todos los puntos que no se encuentren dentro del rango establecido no contendrán ningún valor de profundidad. Por lo tanto, en la fila correspondiente a estos puntos en el archivo

PCD en vez de aparecer el valor de sus coordenadas tridimensionales, se mostrará un valor “NaN” en cada coordenada, haciendo referencia a no haber podido identificar la profundidad en dicho punto.

Esto hace que nuestra nube de puntos contenga multitud de puntos sin información, pero que a la hora de procesar la nube se tienen en cuenta. Por causa de este motivo, un primer paso indispensable consiste en realizar una eliminación de todo aquel índice que no proporcione información útil con el objetivo de reducir el número de puntos de nuestra nube, únicamente a aquellos que son valiosos. Para ello se aplica la función “*removeNaNFromPointCloud*” implementada por las PCL, de forma que elimine aquellos puntos con x, y o z igual a “NaN”. En la siguiente figura puede apreciarse el efecto de aplicar dicho filtro en los archivos PCD.

```
# .PCD v0.7 - Point Cloud Data file format # .PCD v0.7 - Point Cloud Data file format
VERSION 0.7                               VERSION 0.7
FIELDS x y z rgb                          FIELDS x y z rgb
SIZE 4 4 4 4                              SIZE 4 4 4 4
TYPE F F F U                              TYPE F F F U
COUNT 1 1 1 1                            COUNT 1 1 1 1
WIDTH 745762                              WIDTH 364928
HEIGHT 1                                   HEIGHT 1
VIEWPOINT 0 0 0 1 0 0 0                   VIEWPOINT 0 0 0 1 0 0 0
POINTS 745762                             POINTS 364928
DATA ascii                                 DATA ascii
nan nan nan 0                             46.216888 -82.877968 684.815 4289967027
nan nan nan 0                             46.966608 -82.82856 683.8894 4290032820
nan nan nan 0                             47.854294 -82.828384 683.34094 4290032820
nan nan nan 0                             48.883301 -82.877487 683.16901 4290098613
46.216888 -82.877968 684.815 4289967027   50.125778 -83.000633 683.56177 4290164406
nan nan nan 0                             51.156662 -83.049438 683.3858 4290164406
46.966608 -82.82856 683.8894 4290032820   52.259613 -83.122826 683.39667 4290164406
nan nan nan 0                             53.29187 -83.171303 683.21692 4290164406
nan nan nan 0                             44.280792 -82.095917 688.46045 4289901234
nan nan nan 0                             44.343781 -81.803947 685.64209 4289967027
nan nan nan 0                             45.234009 -81.804382 685.09784 4290032820
nan nan nan 0                             46.053802 -81.78022 684.36285 4290032820
nan nan nan 0                             46.802277 -81.731453 683.4375 4290098613
47.854294 -82.828384 683.34094 4290032820 47.619431 -81.706879 682.7002 4290164406
nan nan nan 0                             48.576813 -81.730972 682.33972 4290230199
nan nan nan 0                             49.676102 -81.803749 682.35541 4290361785
nan nan nan 0                             50.633957 -81.827499 681.99158 4290427578
48.883301 -82.877487 683.16901 4290098613 51.591736 -81.851074 681.62598 4290427578
```

Figura 26: Resultado de eliminar puntos NaN.

(Fuente: Elaboración propia)

Se puede comprobar como la reducción de los puntos que conforman la nube es considerable, pero a pesar de la disminución de los índices puede no ser suficiente. Posteriormente se va a realizar un registro entre diversas nubes de puntos similares a esta, por lo tanto, para agilizar el proceso se precisa implementar alguna de las técnicas de “*Downsampling*”. Estas técnicas, dentro del procesamiento digital, están asociadas con procesos de remuestreo de un sistema, en este caso concreto, el de las nubes de puntos. En consiguiente, se

va a aplicar un filtro denominado “VoxelGrid”. Lo que este aporta, es la reducción de muestras de los datos de entrada aplicando técnicas geométricas y con ayuda del acople de una cuadrícula en tres dimensiones sobre la nube de puntos. La rejilla tridimensional creada, está compuesta por una cantidad de *voxels3D*, en otras palabras, por una cantidad de pequeñas cajas 3D. Luego, cada voxel, albergará un número concreto de puntos. Por lo tanto, la forma en la que este filtro reduce la cantidad de muestras de la nube es aproximando todos los puntos presentes dentro de cada voxel con su centroide.

El centroide no es más que el punto correspondiente al centro geométrico de un objeto. Si se divide este en elementos de volumen  $dV$ , la ubicación del centroide para el volumen del objeto puede ser determinada calculando los “momentos” de los elementos con respecto a cada uno de los ejes coordenados, obteniendo como resultado las siguientes ecuaciones:

$$\bar{X}_C = \frac{\int_V x \cdot dV}{\int_V dV}, \quad \bar{Y}_C = \frac{\int_V y \cdot dV}{\int_V dV}, \quad \bar{Z}_C = \frac{\int_V z \cdot dV}{\int_V dV}$$

En la siguiente figura se muestra en que consiste este filtro. De manera gráfica en el voxel izquierdo se tienen todos los puntos de la nube de entrada y en el de la derecha la simplificación de todos los puntos a la del centroide.

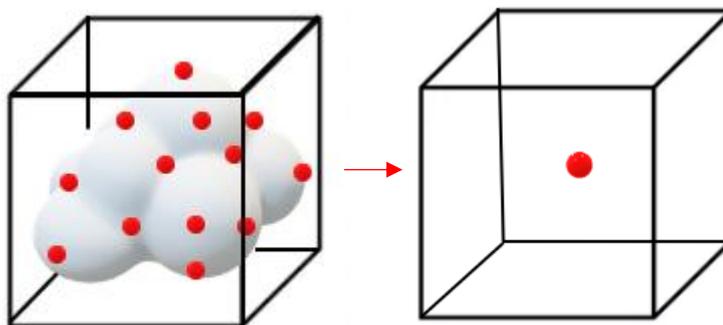
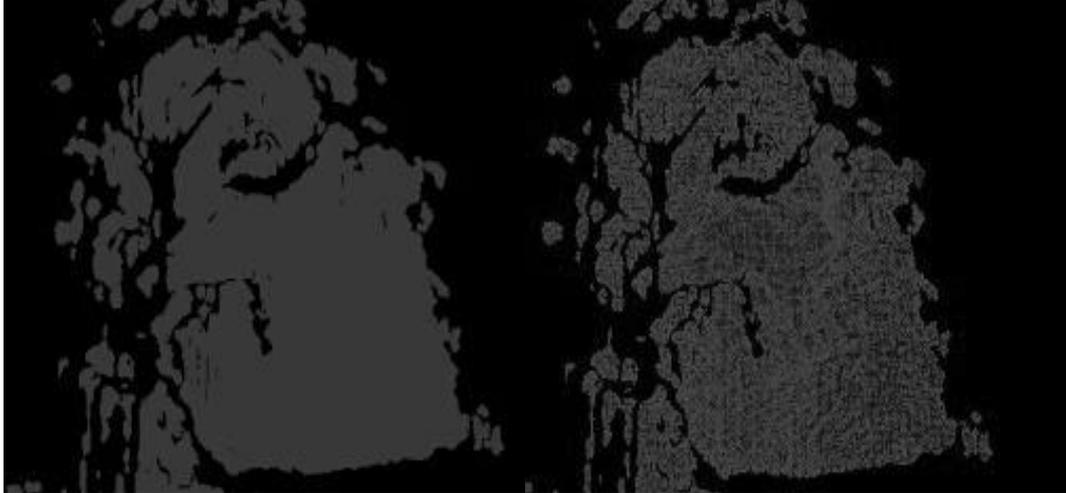


Figura 27: Esquema de aplicación del filtro VoxelGrid.

(Fuente: Elaboración propia).

Se ajusta el tamaño de estos cubos con el fin de reducir el tamaño de la muestra, pero sin perder información relevante sobre ella. En la práctica, un ejemplo del resultado obtenido tras la aplicación de este filtro puede verse en

la figura 28, donde claramente se puede apreciar la diferencia de densidades de puntos entre la nube de entrada, situada a mano izquierda, y la de salida.



*Figura 28: Resultado de aplicar el filtro VoxelGrid sobre las nubes de puntos.*

*(Fuente: Elaboración propia)*

Con este filtrado se obtienen unas nubes de puntos con menor peso, computacionalmente hablando, de manera que para realizar su posterior registro resultan más rápidas de procesar.

Como ya se ha definido previamente, el registro consiste en las diferentes transformaciones geométricas, sean de traslación o rotación, de las nubes de puntos de tal forma que estas se alineen entre sí. Para ello es necesario que se capturen imágenes del cráneo desde diversas perspectivas. En el siguiente conjunto de nubes de la figura 29 se muestran dos de las tomas realizadas. Estas muestras surgen tras un acondicionamiento exhaustivo de los parámetros de la cámara, un posicionamiento estable y un procesado como el citado anteriormente. Las nubes han sido representadas con un tamaño de punto mayor, de manera que las superficies parezcan continuas.

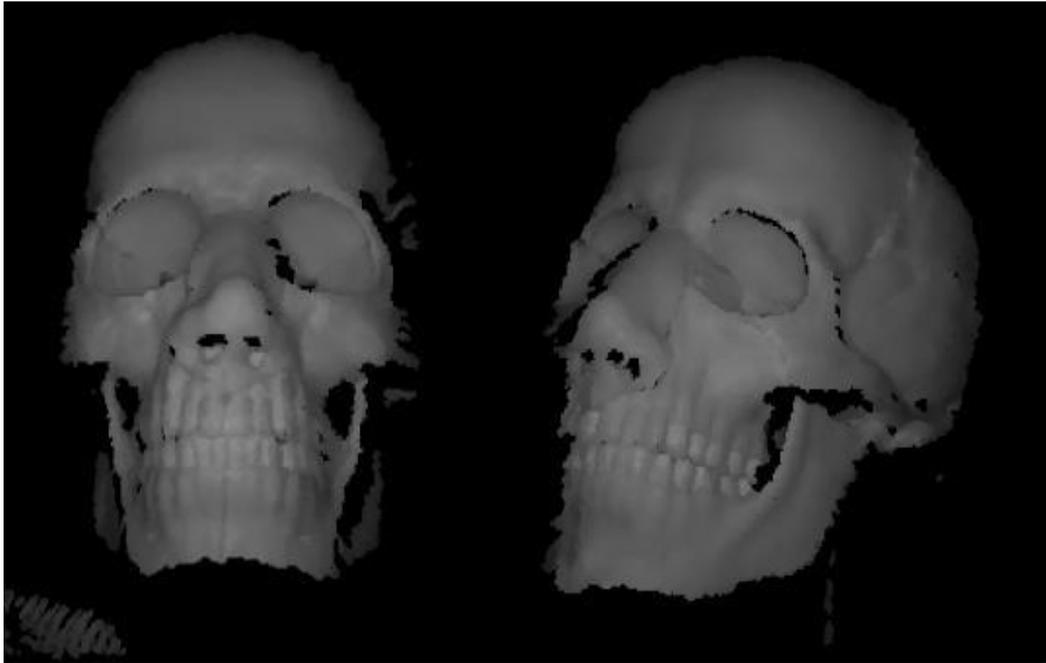


Figura 29: Representación de las nubes obtenidas.

(Fuente: Elaboración propia)

A simple vista se comprueba que, el resultado obtenido a través de la cámara ToF es mucho más limpio, con mayor calidad de detalle y con mayor densidad de puntos que la que se obtenía a través de las cámaras de visión estéreo. Por lo tanto, se puede decir que la mejora ha sido considerable

Sin embargo, las superficies siguen presentando ruido, menor que en el caso de la visión estereoscópica, pero suficiente como para que los métodos de búsqueda de puntos característicos no den con una solución correcta a la hora de intentar casar las distintas nubes entre sí. Este defecto puede verse en la siguiente figura, donde se muestra una de las superficies desde una proximidad mayor.

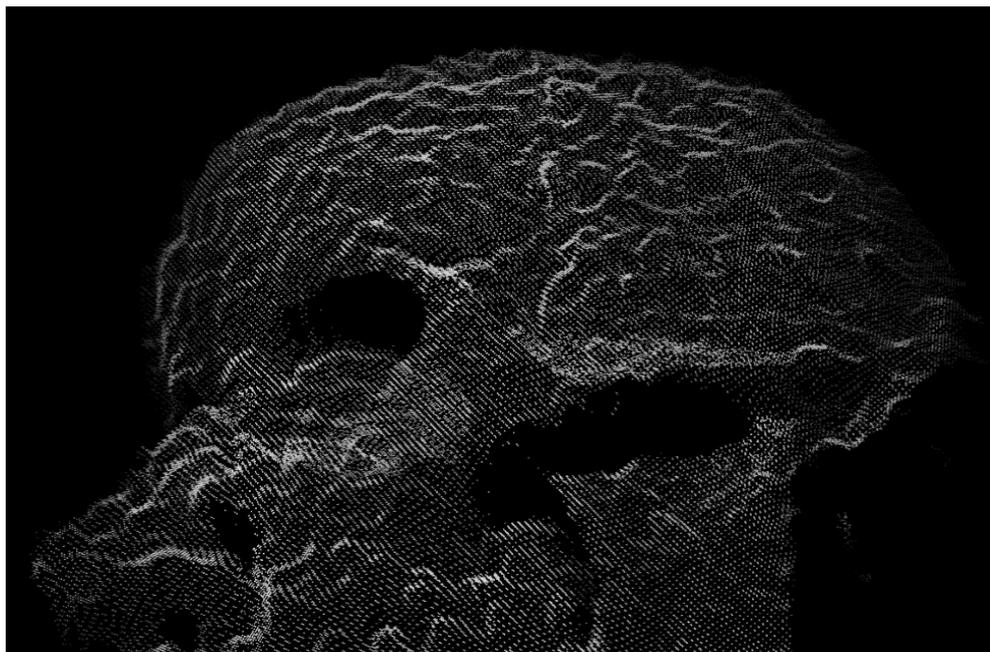


Figura 30: Visualización detallada de la superficie obtenida.

(Fuente: Elaboración propia)

Debido a este problema, se procede a buscar una alternativa. Como se ha explicado, la captura de las distintas nubes de puntos es necesaria para poder realizar un registro y posteriormente una reconstrucción de lo que sería el cráneo completo de forma que se consiga desarrollar el modelo externo de la cabeza.

Hasta ahora la adquisición de imágenes se realiza sin seguir ninguna pauta concreta. Básicamente, se basa en la toma de ciertas imágenes alrededor de la cabeza del paciente, el cráneo en este caso, de manera manual con el fin de encontrar en cada una de las nubes obtenidas puntos significativos con los cuales posteriormente realizar un registro completo. El hecho de que la toma de las nubes fuese manual, tiene como resultado muestras en las que parámetros como la distancia de la cabeza a la cámara o la posición y el ángulo de esta sean aleatorios. Debido a ello, las transformaciones geométricas entre una nube y otra, a pesar de realizar tomas intentando mantener siempre una posición determinada, son diferentes casi en la totalidad de los casos. Es por ello por lo que ahora, se va a intentar eliminar la aleatoriedad del proceso para poder alcanzar los objetivos deseados.

Con el fin de mejorar el método previamente descrito, ahora se va a proceder a realizar una toma de imágenes contralada, eliminando así, las posiciones fortuitas desde las que se hacía anteriormente. Para ello se va a



hacer uso de un brazo robótico, concretamente del UR3 de *UNIVERSAL ROBOTS*, el cual, además es colaborativo. Esta última funcionalidad le permite ser instalado sin tener que utilizar un vallado de seguridad gracias a la “sensibilidad” que tiene ante el contacto con otros objetos. Debido a esta virtud se le permite trabajar con personas, gracias a los sensores de esfuerzo que lleva implementados en cada uno de los ejes. Además, es justamente este “cobot” el que está pensado para que asista al médico en las operaciones de cirugía endoscópica transesfenoidal. Por lo tanto, el realizar la adquisición de imágenes con este brazo robótico al paciente, una vez este situado en el quirófano en el momento previo a la operación, no supondrá ningún estorbo ya que dicho robot se encuentra ya inmerso en el espacio de trabajo.

Uno de los inconvenientes a la hora de realizar las tomas manuales, consiste en el ruido generado por la alta sensibilidad a los movimientos de las cámaras ToF. Por tanto, uno de los grandes beneficios de automatizar el proceso usando el “cobot”, es la pérdida completa del ruido por movimientos durante las tomas. Por otro lado, con este nuevo método se conocen con precisión los puntos desde los cuales se han realizado las capturas. Debido a esto, las transformaciones geométricas a realizar en el procesado quedan totalmente definidas.

Para llevar a cabo este método se ha realizado una prueba en la cual se toman 3 capturas de la cabeza del paciente. Una justo perpendicular a la cara y dos laterales. En la siguiente figura puede verse el entorno, de manera conceptual, sobre el que se ha trabajado.

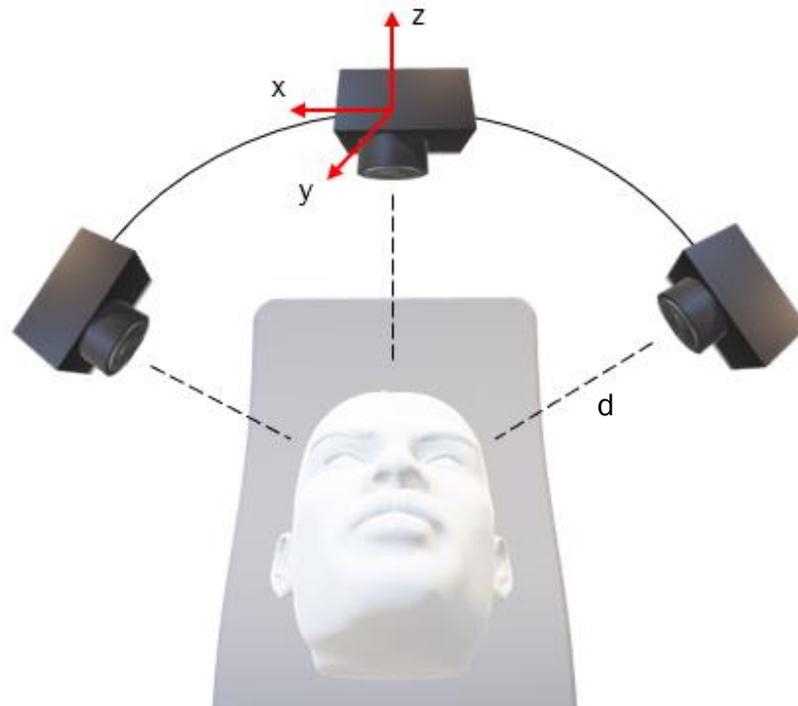


Figura 31: Esquema de posicionamiento de la cámara para las capturas.

(Fuente: Elaboración propia)

Las tomas han sido realizadas a una distancia  $d$ , comprendida entre los 65 y los 120 centímetros, debido a que las muestras obtenidas a una distancia por debajo de los 40 centímetros pierden cierta resolución. El ajuste del rango admisible se ha impuesto basándose en los límites de la zona de trabajo del robot y, además, teniendo en cuenta una separación razonable a la que se puede encontrar el paciente sobre la mesa de operaciones.

Al igual que en el método anterior, el objetivo sigue siendo el mismo, realizar un registro con las nubes de puntos obtenidas en las diferentes tomas. Ya se ha comentado como a través de este nuevo procedimiento, se pierde la necesidad de extraer cualquier tipo de punto significativo de las muestras para poder hacer un “*matching*” entre ellas. Al saber exactamente cuál ha sido la posición desde la cual se han tomado las imágenes, nos bastará con hacer una transformación de las nubes de puntos para ponerlas en correspondencia. Para ello se hará uso de la librería de alto nivel, que ofrece c++, llamada *Eigen*, con la cual se pueden realizar operaciones lineales, matriciales, vectoriales y transformaciones geométricas entre otras cosas. Por lo tanto, con la ayuda de dicha librería, se procederá a crear y configurar una matriz de transformación, entre cada par de nubes, con la cual poder realizar el registro de las muestras en busca de una nube final que las contenga a todas.

Debido al sistema físico con el que se ha trabajado y teniendo en cuenta el sistema de coordenadas con el que trabaja la cámara, el cual se representa de color rojo sobre una de las cámaras de la figura 31, únicamente se deberá tener en cuenta una traslación sobre el eje X y una rotación sobre el eje Y. Para la obtención de las posiciones se ha colocado el robot en una posición en la cual solo mueve uno de los ejes al realizar los movimientos, concretamente el tercer eje. Para determinar cuáles son los valores correspondientes a las transformaciones se ha tenido en cuenta la configuración geométrica formada por el robot. En la siguiente figura puede apreciarse a mano izquierda la postura que adopta el robot durante las tomas y a mano derecha el esquema geométrico utilizado.

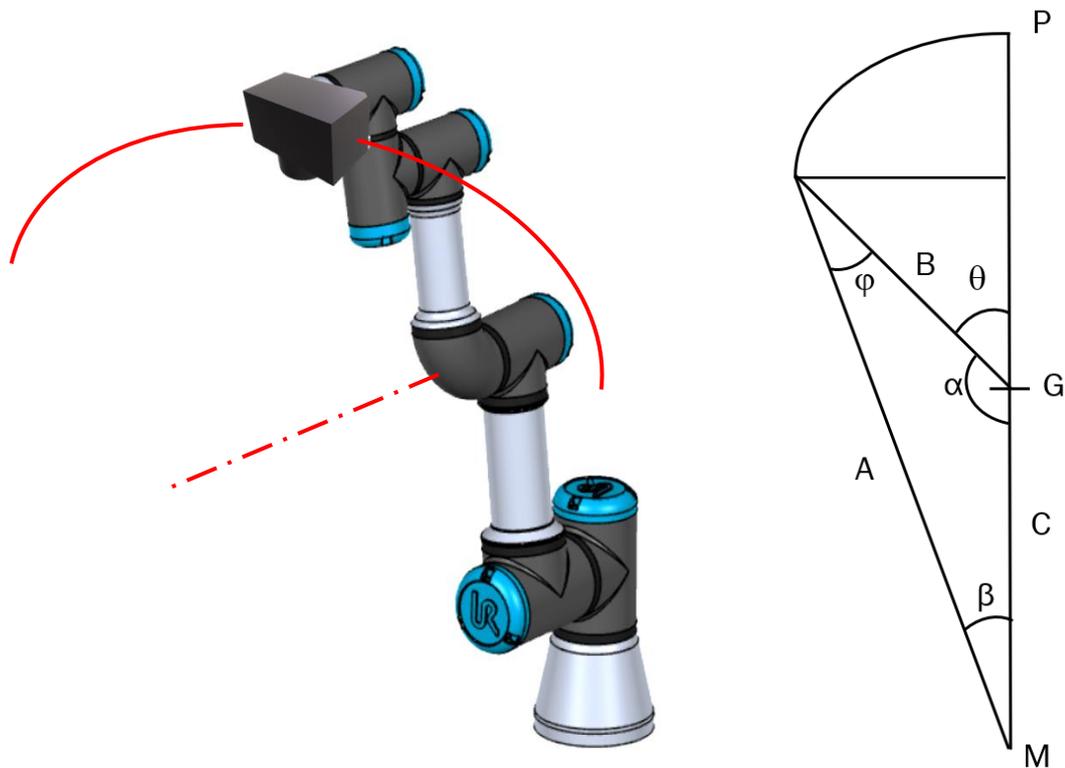


Figura 32: Robot colaborativo UR3 (izquierda) y modelo geométrico (derecha).

(Fuente: Elaboración propia)

El punto M aproxima el punto de giro respecto de la cabeza que se coloque sobre la mesa. El punto G hace referencia al punto respecto al que gira el robot por su tercer eje. Finalmente, el punto P hace referencia a la pinza del brazo robótico, en la cual, en este caso, se incorpora la cámara ToF. Por lo tanto, las distancias MG y GP son conocidas, es decir, se conoce el tamaño de

los segmentos B y C. Por otro lado, también se conoce el valor del ángulo  $\theta$ , por lo tanto, puede obtenerse  $\alpha$  como:

$$\alpha = 180 - \theta$$

Ahora bien, el dato de interés en este esquema es  $\varphi$ . Este proporciona el valor angular necesario que hay que girar la cámara para que esta apunte al objetivo situado en el punto M, en este caso, la cabeza. Para obtener su valor se emplearán el teorema del seno y del coseno.

$$\text{Teorema del seno: } \frac{A}{\sin \alpha} = \frac{B}{\sin \beta} = \frac{C}{\sin \varphi}$$

$$\text{Teorema del coseno: } A^2 = B^2 + C^2 - 2BC \cos \beta$$

Tomando  $\beta$  como:

$$\beta = 180 - \alpha - \varphi = 180 - (180 - \theta) - \varphi = \theta - \varphi$$

Por lo tanto, con la ayuda de estos dos teoremas, se pueden calcular los diferentes valores angulares de  $\varphi$  en función de lo que gire el brazo robótico, es decir, en función del ángulo  $\theta$ .

Conocida la configuración del modelo geométrico de la figura 32, se procede a la obtención de las nubes de puntos y a su posterior procesamiento. La primera muestra que se introduzca será la que se utilice como referencia, y serán las siguientes nubes las que sufran transformaciones geométricas. El resultado del registro obtenido a través de este método puede apreciarse en la siguiente figura.



Figura 33: Resultado de llevar a cabo el registro de las nubes de puntos.

(Fuente: Elaboración propia)

El registro a través de este nuevo método es satisfactorio. Con estos resultados se aprecia nítidamente la silueta de la persona capturada y se pueden diferenciar con suficiente claridad sus diferentes aspectos faciales.

Por otro lado, cabe destacar un avance llevado a cabo en las pruebas realizadas con el cráneo, simulando con este, la cabeza del paciente. Este desarrollo consiste en implementar la eliminación de la superficie posterior, en la cual estaría apoyada la cabeza del paciente, a la hora de proceder a la adquisición de las diferentes tomas. El objetivo principal es la considerable reducción en los puntos que conforman las nubes, haciéndolas más rápidas de procesar y almacenando en ellas únicamente valores útiles.

Para eliminar dicha superficie se hace uso de un filtro denominado *RANSAC*, el cual es la abreviación de “*RAN*dOm *SA*mple *C*onsensus” y podría traducirse como “consenso de la muestra tomada al azar”. Este, implementa un algoritmo iterativo con una técnica de aprendizaje no supervisado para estimar los parámetros de un modelo matemático mediante un muestreo aleatorio de los datos de entrada. El algoritmo *RANSAC* asume que los datos que analiza están constituidos tanto de *inliers* y *outliers*. Los *inliers* pueden definirse como el conjunto de puntos que pueden ser definidos por un modelo con unos determinados parámetros, mientras que los *outliers*, o valores atípicos, no encajan con ese modelo en ninguna circunstancia. También, es



necesario disponer de un procedimiento que permita estimar de forma óptima los parámetros del modelo matemático elegido a partir de los datos.

Los datos de entrada a este filtro son la muestra a observar, un modelo parametrizado capaz de explicar o adaptarse a las observaciones y ciertos umbrales de confianza que definen las desviaciones máximas permitidas respecto al modelo. Por otro lado, el algoritmo se compone esencialmente de dos pasos que se repiten iterativamente:

- Se selecciona de manera aleatoria un subconjunto de los datos originales. Estos valores elegidos son hipotéticos, y esta hipótesis ha de comprobarse. Para ello, se calcula el modelo junto con sus parámetros correspondientes al grupo de puntos seleccionado.
- El algoritmo comprueba que elementos de todo el conjunto de datos son consistentes con el modelo, es decir, cuantos de los índices encajan con él. En este punto del algoritmo se atienden las posibles desviaciones, las cuales, normalmente están atribuidas a posibles efectos de ruido. Finalmente, se puntúa el grado de consenso, por ejemplo, utilizando el número de *inliers* dentro del umbral.

El conjunto de *inliers* obtenidos por el ajuste del modelo se denominan “conjunto de consenso”. El algoritmo *RANSAC* repite iterativamente los dos pasos anteriores hasta que el consenso obtenido en una determinada iteración tenga suficientes *inliers* o hasta llegar a un número máximo de iteraciones. En este último caso, el algoritmo devuelve el conjunto de consenso con mayor puntuación encontrado hasta el momento. En las siguientes figuras se muestran cuáles serían los pasos seguidos por el algoritmo, teniendo como modelo matemático una recta, una desviación estándar de valor  $\delta$  y como nube de entrada, la mostrada en la figura 34. El proceso se realizará mostrando dos casos distintos de aplicación del algoritmo.

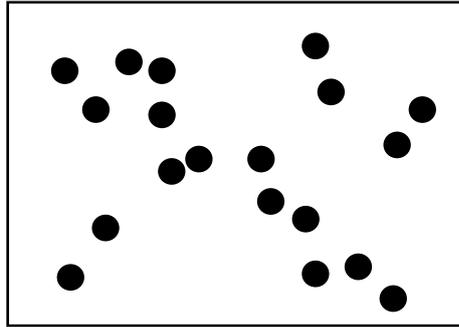


Figura 34: Nube de puntos de partida para aplicar RANSAC.

(Fuente: Elaboración propia)

1. Selección aleatoria de un subconjunto de la nube

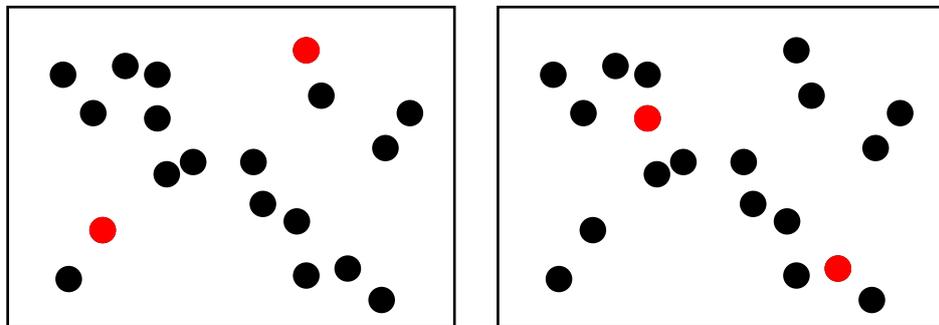


Figura 35: Selección aleatoria.

(Fuente: Elaboración propia)

2. Calcula el modelo de la muestra

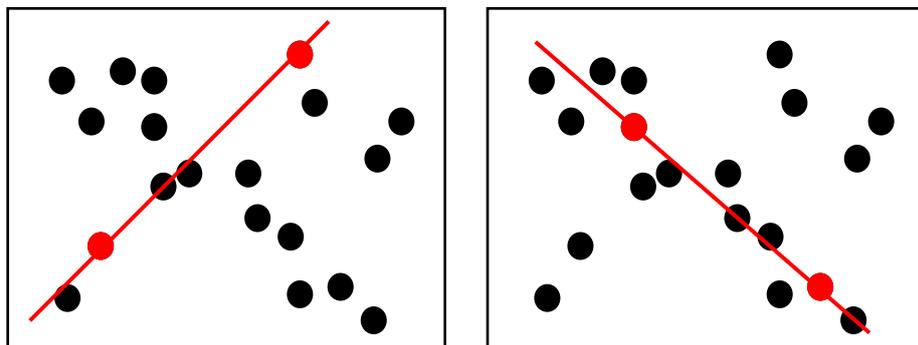


Figura 36: Calculo del modelo lineal.

(Fuente: Elaboración propia)

### 3. Tener en cuenta el valor de desviación estándar

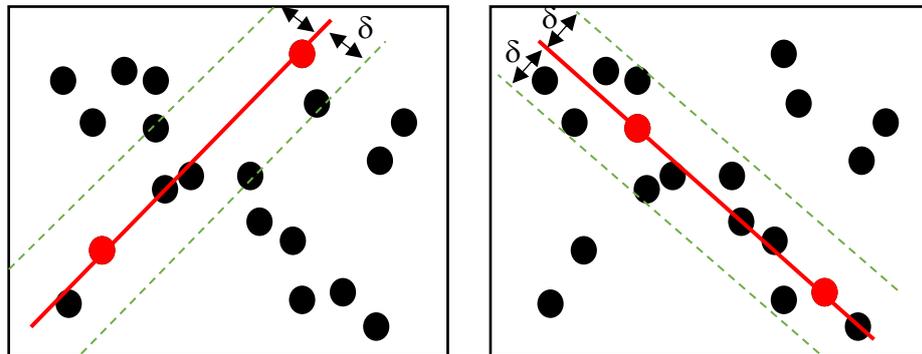


Figura 37: Valoración de la desviación estándar.

(Fuente: Elaboración propia)

### 4. Obtener puntuación del conjunto de consenso

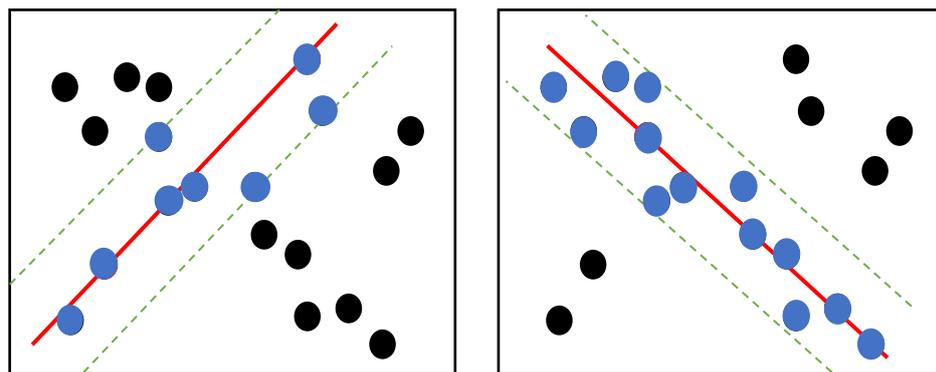


Figura 38: Obtención de puntuación.

(Fuente: Elaboración propia)

### 5. Valoración de los resultados

Suponiendo que el algoritmo RANSAC, en este caso, únicamente realiza estas dos iteraciones. Los inliers y outliers que se obtendrían como resultado son los mostrados en la figura 39, de izquierda a derecha respectivamente. Estos corresponden a la iteración desarrollada a mano derecha en las figuras anteriores, debido a que su puntuación es de 13, frente a los 8 obtenidos en la mano izquierda.

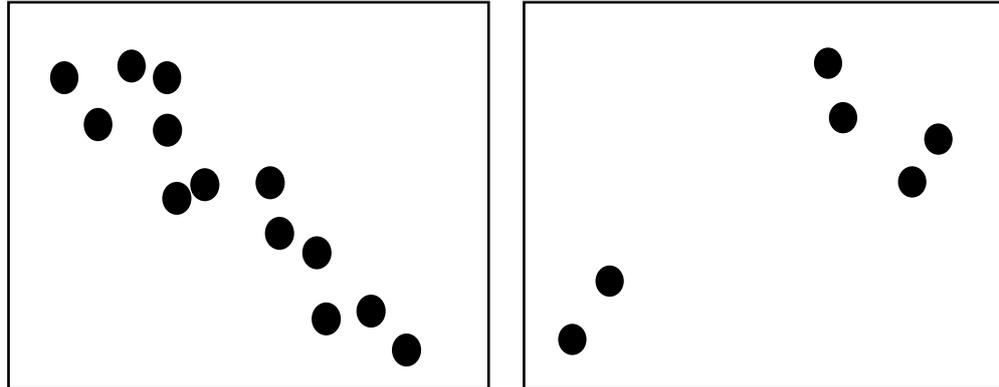


Figura 39: Resultado obtenido por RANSAC, mostrando inliers (izquierda) y outliers(derecha).

(Fuente: Elaboración propia)

Una gran ventaja de este método es su capacidad para realizar una estimación sólida de los parámetros del modelo, es decir, puede estimar los parámetros con un alto grado de precisión incluso cuando hay un número significativo de valores atípicos presentes en el conjunto de datos.

Concretamente, como ya se ha citado previamente, con la ayuda de este método se va a proceder a la eliminación de la superficie plana situada en la parte posterior de la cabeza del paciente. Esta superficie equivaldría a la cabecera de la mesa quirúrgica. Por lo tanto, el modelo matemático buscado ahora es un plano y la desviación estándar se determina para que todos los puntos del plano que contengan ruido se ajusten al modelo también. En este caso, los inliers serán los puntos que se hayan eliminado, con el objetivo de obtener únicamente el conjunto de puntos de la nube no pertenecientes al plano y, por consiguiente, pertenecientes a la cabeza del paciente. Los resultados de la aplicación del algoritmo RANSAC se ve en las siguientes figuras.

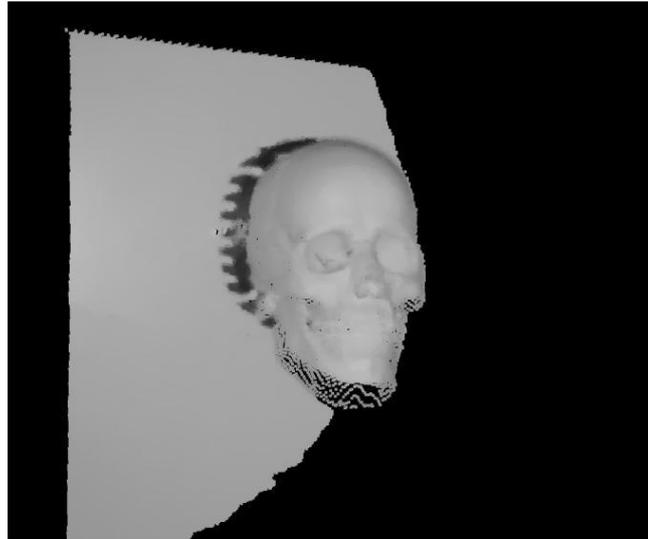


Figura 40: Nube de puntos de entrada que contiene un plano a eliminar.

(Fuente: Elaboración propia)

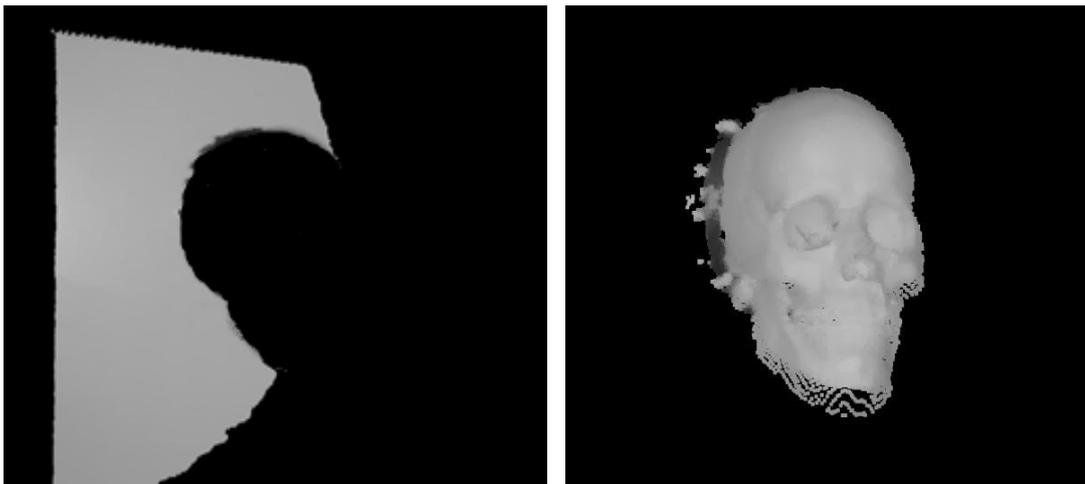


Figura 41: Resultado obtenido por RANSAC, mostrando inliers (izquierda) y outliers (derecha).

(Fuente: Elaboración propia)



Universidad de Valladolid

# Integración de un sistema de visión 3D en un entorno robotizado de cirugía Ander Loidi Yarza



ESCUELA DE INGENIERÍAS  
INDUSTRIALES

## 5 PROCESAMIENTO DE IRMS

### 5.1 INTRODUCCIÓN AL CAPÍTULO

En el análisis de las imágenes de resonancia magnética (IRM), se va a realizar un procedimiento, el cual permita obtener diferentes modelos tridimensionales de estructuras anatómicas humanas.

A modo de aclaración va a detallarse cuales son los planos anatómicos en un ser humano. Los planos principales son tres: el horizontal o axial, el plano frontal o coronal y el sagital o medio. Pueden apreciarse los cortes de los planos citados en la siguiente figura.

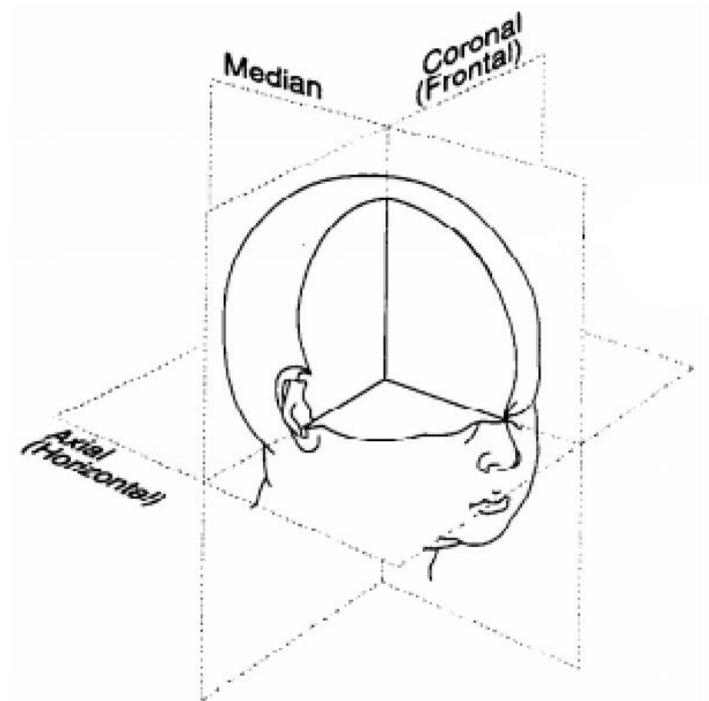


Figura 42: Representación de los tres planos anatómicos principales.

(Fuente: Elaboración propia a partir de Timor-Tritsch L.F. y Monteagudo A. (1996). "Transvaginal fetal neurosonography: standardization of the planes and sections by anatomic landmarks". Ultrasound in Obstetrics and Gynecology, 8, 42-47.)

## 5.2 GENERACIÓN DE LOS TRES PLANOS ANATÓMICOS

Los datos de entrada de los que se dispone consisten en un conjunto de imágenes, concretamente doscientas, del plano axial de la cabeza humana. Algunas de ellas, pueden verse en la siguiente figura.

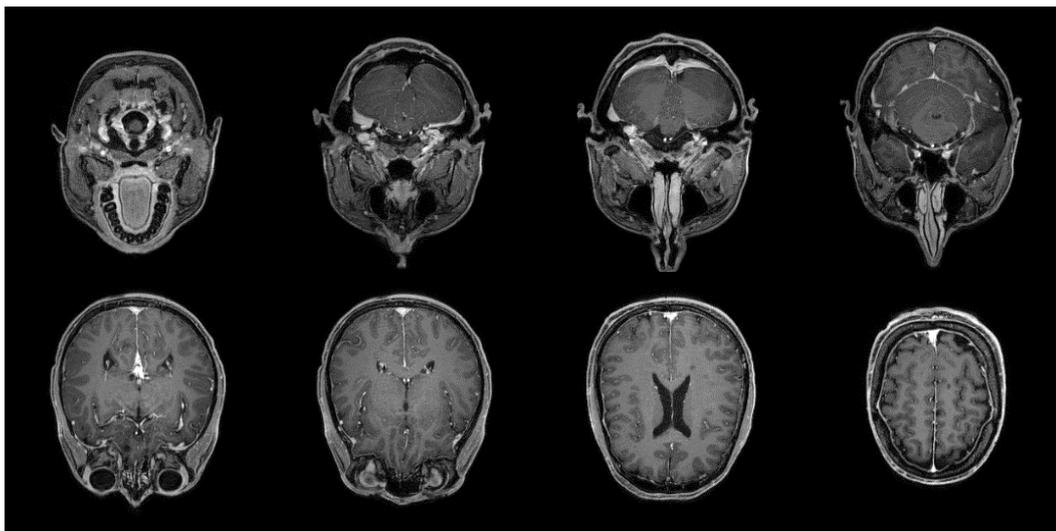


Figura 43: IRM correspondientes al plano axial.

(Fuente: Proporcionadas por el Hospital Clínico Universitario de Valladolid para su uso en el proyecto)

Estos cortes aportan una cantidad enorme de información a los médicos a la hora de diagnosticar posibles daños, síntomas o enfermedades. Sin embargo, a través del corte en un únicamente plano no siempre es fácil distinguir con claridad lo que se está viendo. Debido a ello, se pretenden generar las imágenes correspondientes a los otros dos planos, el sagital y el coronal. Para ello, se van a examinar de manera iterativa la totalidad de las imágenes pertenecientes al plano axial con el objetivo de obtener la información necesaria de ellas para la formación del resto de los planos.

### 5.2.1 Plano sagital

En este caso, la información que interesa de las imágenes del plano axial son los píxeles contenidos en cada una de las columnas de cada imagen. El primer corte del plano sagital estará formado por la primera columna de cada

uno de los cortes del plano axial. Por lo tanto, se deberá extraer el valor de los citados píxeles por cada columna y superponer el conjunto de datos ordenadamente uno encima del otro para la generación del plano sagital. De manera gráfica puede apreciarse en la figura 44 como sería la generación de uno de los cortes sagitales, donde las líneas verticales de color negro en cada uno de los cortes hacen referencia a las columnas de la imagen.

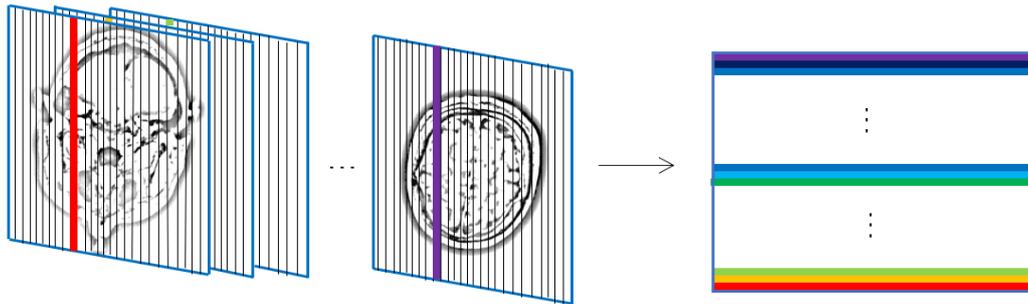


Figura 44: Esquema de la generación de las IRM del plano sagital a partir del axial.

(Fuente: Elaboración propia)

Como ya se ha dicho, puede observarse como la formación de la imagen del plano sagital, representada a mano derecha, surge de la superposición de todas las columnas de las imágenes del plano axial.

El pseudocódigo para la generación de todas las imágenes sería el siguiente:

```
for(número de columnas "ncol")
{
  for(número de imágenes del plano axial "nf")
  {
    For(número de filas "nfil")
    {
      Sagital(nfil,nf) = axial(nfil,ncol)
    }
  }
  rotación(Sagital con 90 grados)
  vectorSagital.pushback(SagitalRotada)
}
```

De esta manera se barre y almacena debidamente el valor de cada píxel de todas las imágenes del plano sagital. Posteriormente se implementa una rotación de la imagen para disponerla verticalmente para su futura visualización. El resultado obtenido se ve en el conjunto de imágenes de la siguiente figura.

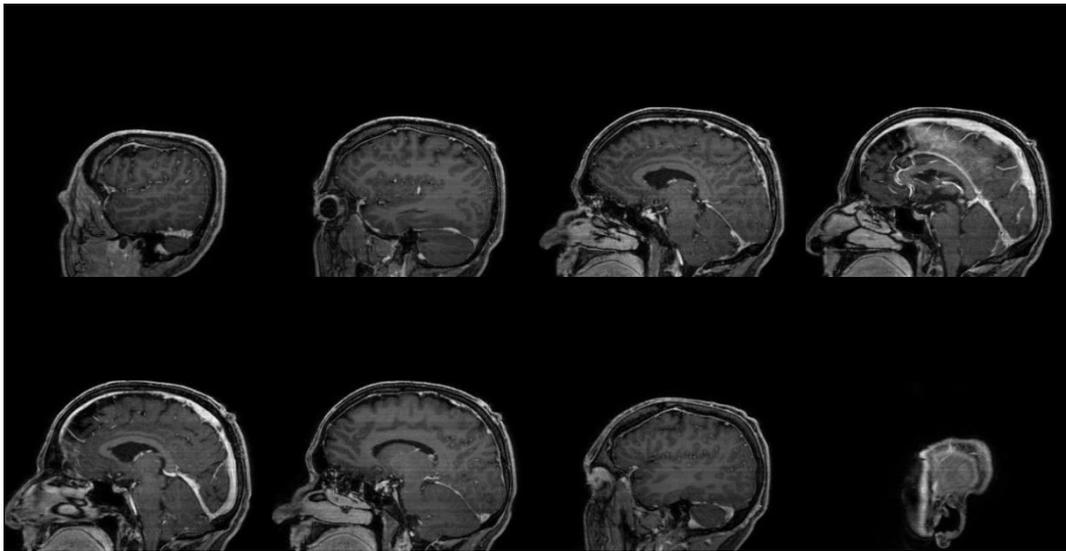


Figura 45: IRM correspondientes al plano sagital.

(Fuente: Elaboración propia)

### 5.2.2 Plano coronal

Para la construcción de este plano, la información de relevancia de las imágenes del plano axial son sus filas, concretamente los píxeles de cada una de ellas. En otras palabras, el primer corte del plano coronal estará formado por la primera fila de cada uno de los cortes del plano axial. Debido a ello, será necesario identificar cada una de las filas para poder acceder a los píxeles que la forman y construir las imágenes.

En este caso, en vez de generar un código específico para la creación de los distintos cortes se va a aprovechar el explicado anteriormente. Para poder reutilizar el código, únicamente se tendrá que rotar las imágenes del plano axial 90 grados. De esta manera, las filas pasan a ser columnas y viceversa, y ya que el código de generación del plano sagital obtenía la información de cada una de las columnas, podrá ser implementado de nuevo. El pseudocódigo sería el siguiente:

```
for(número de imágenes del plano axial)
{
    rotación(Axial con 90 grados)
}

for(número de columnas "ncol")
{
    for(número de imágenes del plano axial "nf")
    {
        For(número de filas "nfil")
        {
            Coronal(nfil,nf) = AxialRotada(nfil,ncol)
        }
    }
    rotación(Coronal con 90 grados)
    vectorCoronal.pushback(CoronalRotada)
}
```

De manera gráfica puede apreciarse en la siguiente figura como sería la generación de uno de los cortes coronales, donde las líneas verticales en cada uno de los cortes vuelven a hacer referencia a las columnas de la imagen.

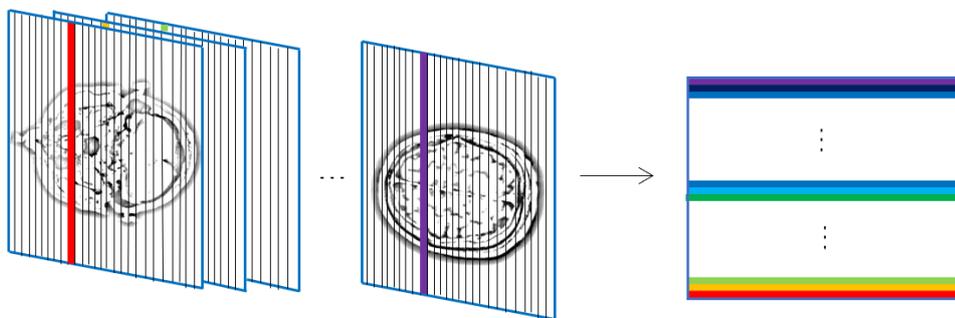


Figura 46: Esquema de la generación de las IRM del plano coronal a partir del axial.

(Fuente: Elaboración propia)

Por lo tanto, mediante la ejecución del código explicado anteriormente, se vuelve a barrer y almacenar debidamente el valor de cada píxel de todas las imágenes del plano coronal. También se aplica una rotación a cada imagen creada de manera que el resultado pueda ser visualizado de la forma más

cómoda posible. En el siguiente conjunto de imágenes de la figura 47 se aprecia cual ha sido dicho resultado.

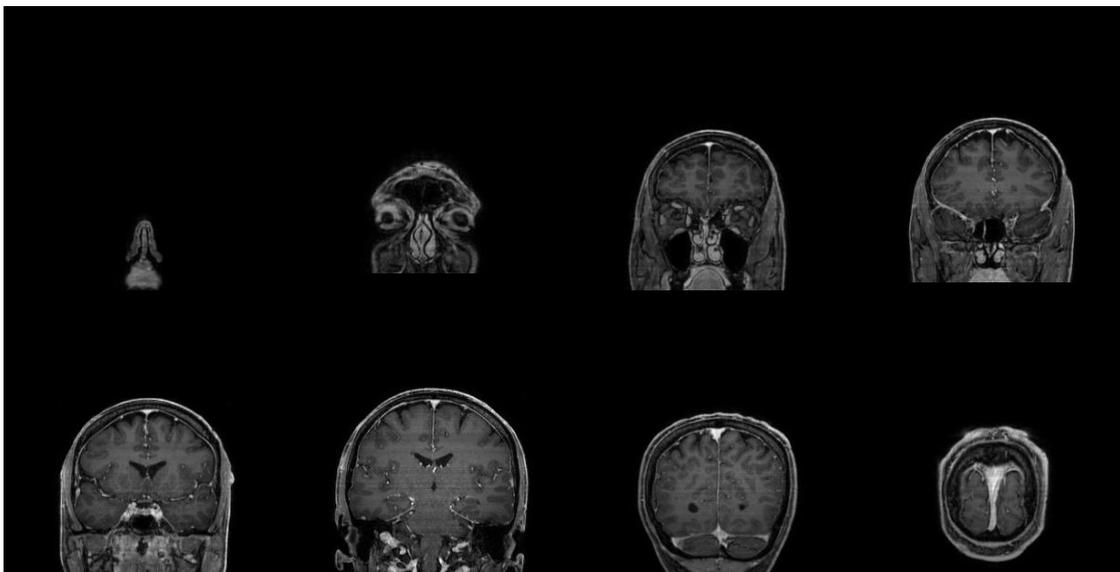


Figura 47: IRM correspondientes al plano coronal.

(Fuente: Elaboración propia)

### 5.3 OBTENCIÓN TRIDIMENSIONAL DE LAS NARINAS

En este caso se parte del conjunto de IRM de la cabeza proporcionadas por el plano axial. Lo que se pretende es buscar las narinas en ellas. Para ello se ha determinado una región de interés (ROI) en la cual puede encontrarse la nariz. Una ROI no es más que una fracción de la imagen, en la cual se pretende focalizar la atención en busca de posible información relevante. Para poder especificar las coordenadas de la ROI dentro de la imagen se han tenido en cuenta las medidas y proporciones de una cara humana promediando los datos obtenidos por Leong S.C. y White P.S. en su trabajo “*A comparison of aesthetic proportions between the Oriental and Caucasian nose*”. Se obtiene que la nariz humana tiene una anchura promedio de 3,75 centímetros y una largura promedio de 3,425 centímetros.

Para asegurar el encuadre de la nariz en su totalidad dentro de la ROI, se tomará un valor de 4 centímetros para la anchura. Sin embargo, para la largura no bastara con ajustar un poco el valor obtenido ya que las fosas nasales no se apreciarían en su totalidad. Por lo tanto, se determina una largura de 7.5 centímetros para la ROI.

Ahora solo queda situar el centro de esta región de interés dentro de nuestra imagen. Al encontrarse la cabeza centrada en las IRM, es fácil centrar la región de interés verticalmente justo en la mitad de la imagen, es decir, sobre alguno punto del eje representado en amarillo en la figura 48. Para centrarla horizontalmente, se comenzará buscando aquel punto que contenga el mayor valor de la coordenada "Y". Puede que el punto que se encuentre no esté situado sobre el eje amarillo, pero esto no supone ningún problema. En la figura 48 se representa una posición aleatoria de dicho índice representado por un punto de color verde. Posteriormente se le resta, al valor "Y" de dicha posición, la mitad del valor de la altura definida para la ROI, y se traza el eje horizontal verde. De esta manera, queda definido el centro, a partir del cual se generará la ROI, como la intersección del eje verde y el eje amarillo. Orientativamente, puede apreciarse la localización del centro de la región en la figura 48 representado por un punto rojo.

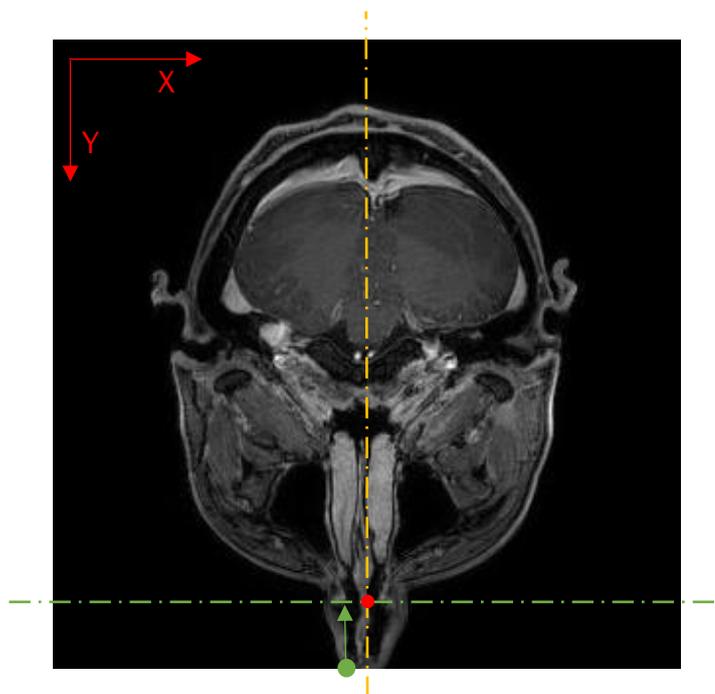


Figura 48: Localización del centro de la región de interés (ROI) sobre imagen axial.

(Fuente: Elaboración propia)

Una vez queda definido en que consiste y como se calcula la región de interés, se procede a la obtención de las narinas.

Para ello se barrerán todas las imágenes del plano axial, y para cada una de ellas se realizarán los siguientes pasos:



- Lectura de la imagen y aplicación de un filtro de suavizado, con el objetivo de reducir el ruido y conseguir que las intensidades de pequeños elementos se mezclen con el fondo de manera que se detecten únicamente las estructuras de mayor tamaño. En este caso se aplica el filtro Gaussiano, el cual realiza una convolución para cada uno de los puntos de la matriz de entrada con un kernel gaussiano y luego suma todos ellos para producir una imagen de salida. El kernel o máscara que se implementa se caracteriza por dar un mayor peso al píxel central y sus vecinos cercanos, y menor peso a los lejanos. La máscara M utilizada es la siguiente:

$$M = \frac{1}{9} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

Numéricamente, un gaussiano en dos dimensiones puede definirse según la siguiente ecuación:

$$G_0(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

Donde  $\mu$  es la media y  $\sigma$  representa la desviación estándar, para cada una de las variables x e y.

- Una vez aplicado el filtro se procede a la binarización de la imagen, de tal forma que, posteriormente se pueda trabajar de una manera más eficaz sobre ella. Para la binarización se emplea un umbral adaptativo. A diferencia de la binarización simple, en la cual el valor umbral es global y se aplica por igual para todos los píxeles de la imagen, este método adaptativo calcula el valor umbral para las regiones más pequeñas, de manera que, existirán diferentes valores umbrales para las distintas regiones. El modelo por el cual se calcula el valor del umbral para cada región es el “ADAPTATIVE\_THRESH\_GAUSSIAN\_C”.

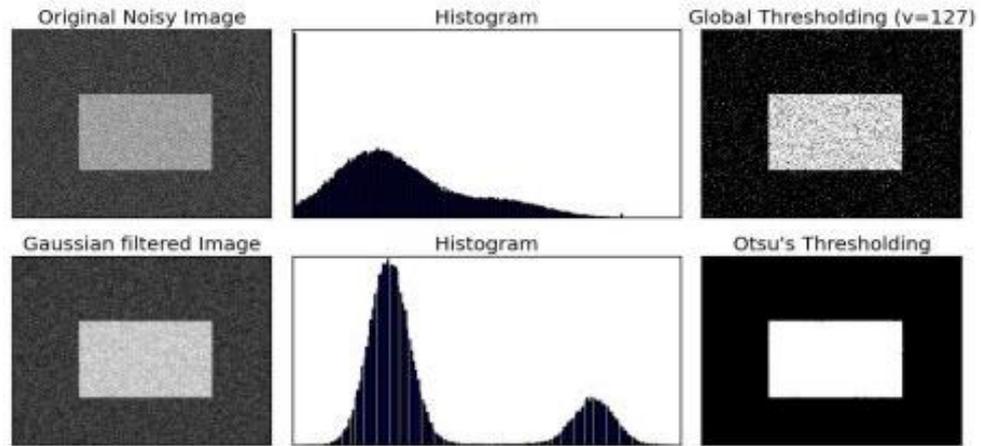


Figura 49: Diferencia entre aplicar un umbral global (arriba) y uno adaptativo (abajo).

(Fuente: «OpenCV» [En línea]. Disponible: <http://opencv.org/>)

Al igual que en el filtro gaussiano, el valor del umbral es la suma ponderada de los valores del vecindario donde la máscara sigue una distribución gaussiana. La letra “C” hace referencia a una constante que se resta de la media ponderada calculada.

Se puede ver en la siguiente figura el resultado de esta binarización.

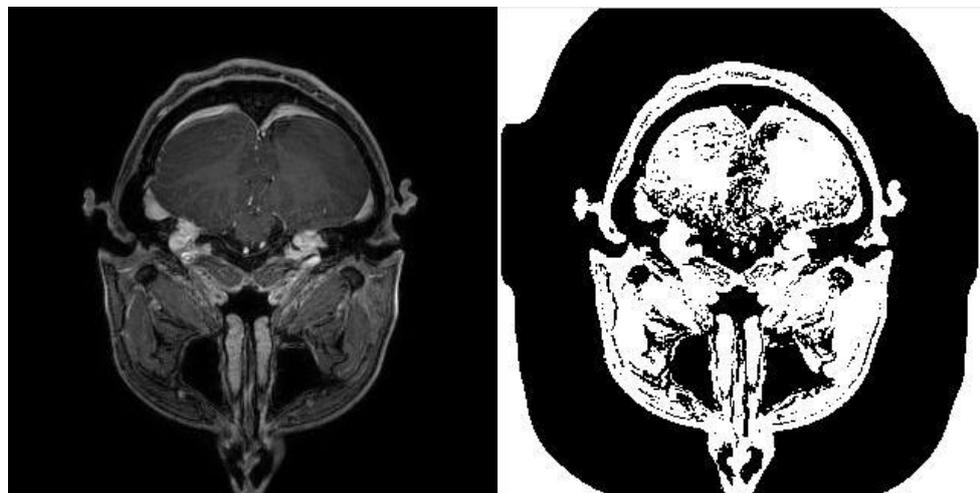


Figura 50: Resultado de binarizar con umbral adaptativo.

(Fuente: Elaboración propia)

- Ahora se procede a recortar la imagen y quedarse únicamente con la región de interés previamente calculada. En la siguiente imagen puede verse a mano izquierda, la ROI marcada por un rectángulo de bordes rojos sobre la imagen binarizada y posteriormente, a mano derecha, el recorte citado con el objetivo de analizar a partir de este punto únicamente esta zona deseada.

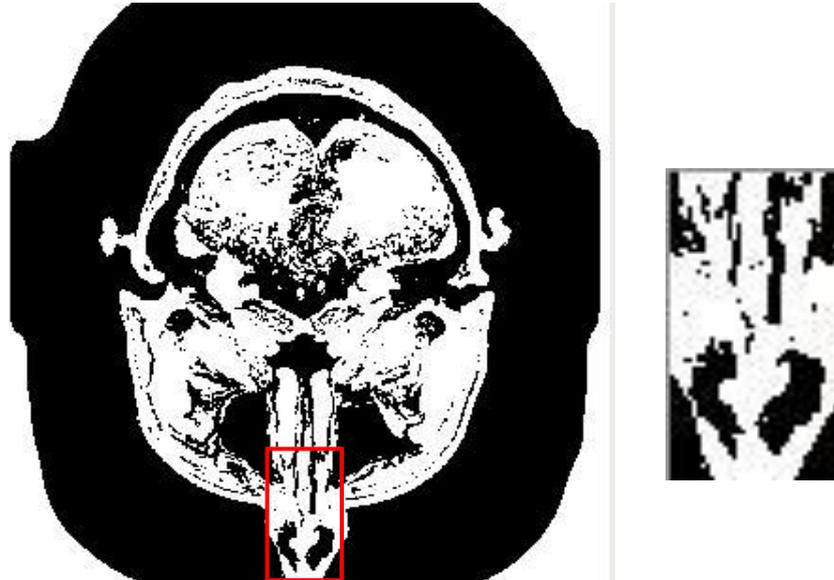


Figura 51: Caracterización de la región de interés (ROI).

(Fuente: Elaboración propia)

- Una vez se obtiene la ROI de manera independiente, se le aplica una transformación morfológica. Estas transformaciones consisten en operaciones sencillas basadas en la forma de las imágenes. El proceso a través del cual se genera la imagen de salida es aplicando sobre la imagen de entrada un elemento estructurador, el cual se asemeja al kernel citado anteriormente, de manera que este decida la naturaleza de la operación.

En este caso, concretamente se implementa un cierre, que no es más que una dilatación seguida de una erosión. Numéricamente, el cierre de una imagen  $A$  con un elemento  $B$ , consiste en lo siguiente:

$$A \bullet B = (A \oplus B) \ominus B$$

La erosión es una de las operaciones fundamentales en el procesamiento morfológico de las imágenes y en las que se basan el resto de las transformaciones. En la erosión binaria la idea principal es

sondear una imagen con un elemento estructurador, de conformación sencilla y predefinida, de manera que se saquen conclusiones sobre como este encaja o no con las distintas formas de la imagen. El elemento de sondeo es a su vez una imagen binaria, es decir, un subconjunto del espacio, y en la erosión, este determina la forma del vecindario de pixeles sobre el que se toma el mínimo. Numéricamente, la erosión de una imagen A con un elemento B, consiste en lo siguiente:

$$A \ominus B = C$$

Donde el valor de cada píxel de la imagen resultante C se calcula como se muestra a continuación:

$$C(x, y) = \min[A(x + x_{ee}, y + y_{ee})]$$

Siendo  $x_e$  y  $y_e$  un valor específico determinado en función del elemento estructural (ee) implementado. En la siguiente figura se muestra de manera gráfica un ejemplo de erosión sobre la imagen A con un ee B, teniendo como resultado el conjunto de pixeles contenidos en C.

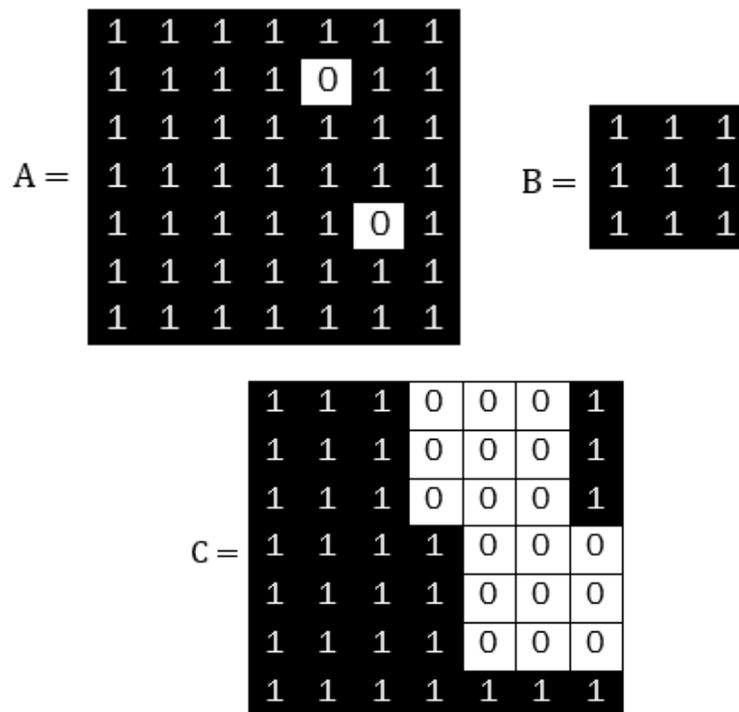


Figura 52: Ejemplo esquemático de aplicar una erosión sobre A con un ee B.

(Fuente: Elaboración propia)

Por otro lado, la dilatación es justo lo contrario a una erosión. En esta operación morfológica, con que uno de los píxeles del elemento estructural este a '1' en el sondeo, el valor del píxel de salida se pone a '1' también. Numéricamente, la dilatación de una imagen A con un elemento estructurador B consiste en lo siguiente:

$$A \oplus B = C$$

Donde el valor de cada píxel de la imagen resultante C se calcula como se muestra a continuación:

$$C(x, y) = \max[A(x + x_{ee}, y + y_{ee})]$$

Siendo  $x_e$  y  $y_e$  un valor específico determinado en función del elemento estructural (ee) implementado. En la siguiente figura se muestra de manera gráfica un ejemplo de dilatación sobre la imagen A con un ee B, teniendo como resultado el conjunto de píxeles contenidos en C.

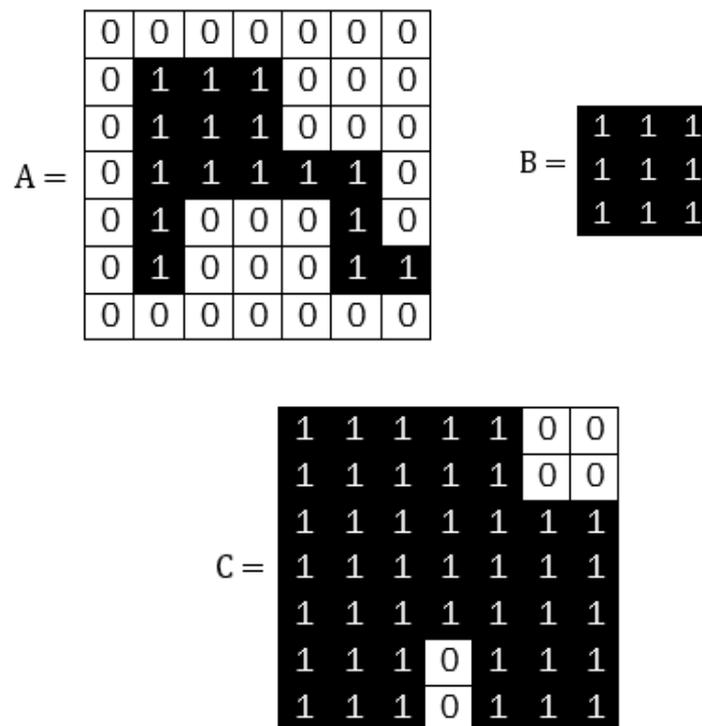


Figura 53: Ejemplo esquemático de aplicar una dilatación sobre A con un ee B.

(Fuente: Elaboración propia)

Por lo tanto, como se ha definido previamente, el cierre consiste en la implementación de las citadas transformaciones geométricas, es decir, de una dilatación seguida de una erosión. En este caso, el cierre que se aplica es bastante suave, en otras palabras, el elemento estructurador empleado para la morfología es de pequeño tamaño. Como de la propia palabra puede intuirse, esta operación morfológica se encargará de cubrir pequeños huecos que puedan surgir en la imagen. En la siguiente figura se muestran los efectos más significativos para poder apreciar su implicación en el proceso.

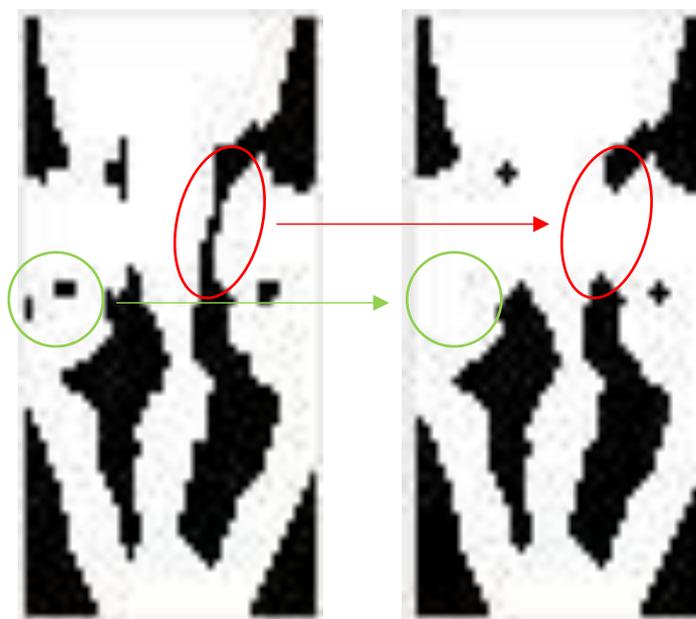


Figura 54: Resultado de aplicar la operación morfológica de cierre sobre la ROI.

(Fuente: Elaboración propia)

- Seguidamente, se procede a la localización de contornos en nuestra ROI. La búsqueda consiste en localizar todos aquellos puntos continuos, a lo largo del límite, con el mismo color o intensidad. De esta manera, se puede diferenciar y detectar los distintos objetos presentes en cada imagen. El resultado de la búsqueda de contorno puede verse en la siguiente figura.

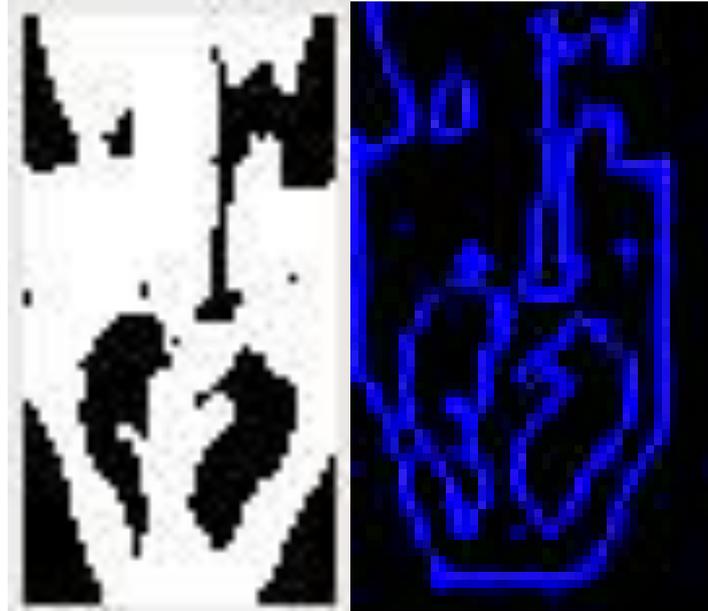


Figura 55: Detección de contornos sobre la ROI.

(Fuente: Elaboración propia)

Para poder identificar y almacenar únicamente con las narinas en estas secciones se procede a discriminar los objetos en función del valor de su superficie. Es decir, se aplicará un filtro donde se establecen unos límites para los valores de las áreas de manera que solo se almacenen aquellos objetos que cumplan los requisitos. La diferencia entre las áreas en todas las imágenes es bastante grande, donde se pueden encontrar elementos con valores muy pequeños (representado por el color verde en la figura 56) o muy grandes (rojo), los cuales pueden considerarse despreciables, y valores intermedios (azul), correspondientes a las fosas nasales.

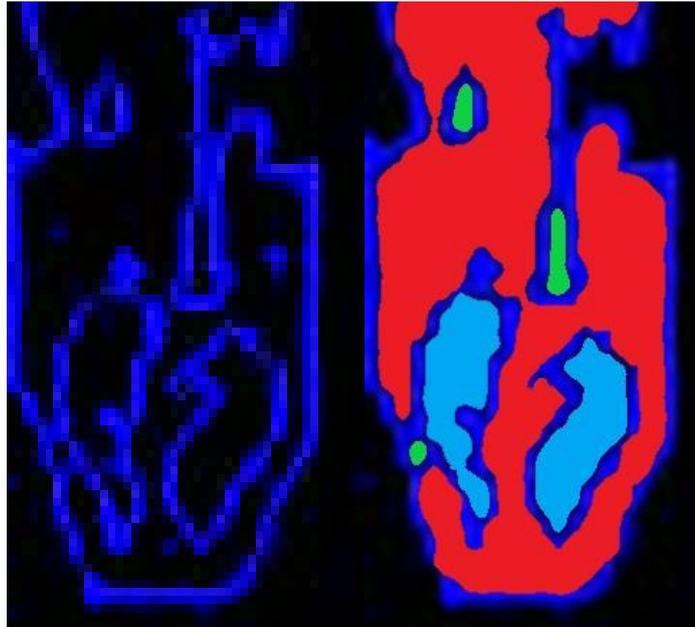


Figura 56: Representación por colores de las diferentes superficies en función de su tamaño.

(Fuente: Elaboración propia)

Por lo tanto, es bastante sencilla la definición de los límites. Una vez estos quedan definidos, tanto el inferior como el superior, se obtiene el resultado que se puede ver en la figura 57 al aplicar la citada segmentación.

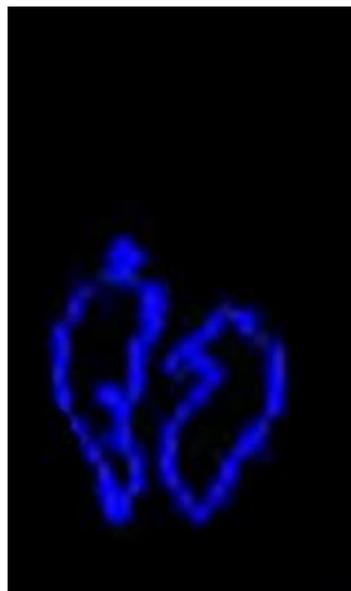


Figura 57: Segmentación de las narinas por superficie.

(Fuente: Elaboración propia)



Surge un gran interés en poder diferenciar la fosa nasal izquierda de la derecha con vista al futuro, ya que, de esta manera, podrían calcularse trayectorias personalizadas para cada una de las narinas. Debido a esto, se procede a la identificación individualizada de cada una.

Para ello se realiza una cuantificación de los momentos de los contornos. Estos se calculan utilizando la fórmula de Green, la cual relaciona una línea integral alrededor de una curva cerrada simple (C) y una doble integral sobre la región plana (D) limitada por dicha curva. Matemáticamente el teorema queda definido por la siguiente ecuación:

$$\oint_C (L dx + M dy) = \iint_D \left( \frac{\delta M}{\delta x} - \frac{\delta L}{\delta y} \right) dx dy$$

Donde L y M son funciones de (x, y) definidas en una región abierta que contiene D y tienen derivadas parciales continuas allí.

Este teorema es de gran utilidad en la geometría plana, ya que se puede utilizar para determinar el centro de las figuras planas únicamente mediante la integración sobre el perímetro.

Los momentos centrales se computan como:

$$mu_{ji} = \sum_{x,y} (matriz(x, y) \cdot (x - \bar{x})^j \cdot (y - \bar{y})^i)$$

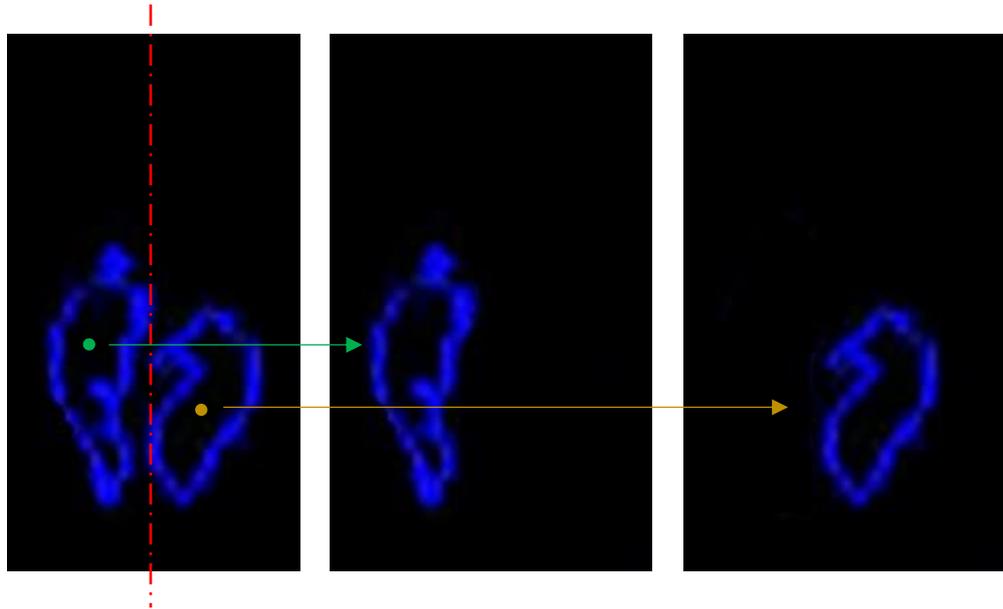
Donde  $\bar{x}$  e  $\bar{y}$  son el centro de masa:

$$\bar{x} = \frac{m_{10}}{m_{00}}, \bar{y} = \frac{m_{01}}{m_{00}}$$

Por lo tanto, conocidas las coordenadas de los centros de las superficies creadas por los diferentes contornos puede diferenciarse las fosas nasales en la de la izquierda y la derecha en función de la posición relativa de los calculados centros. Para ello se divide la ROI en dos partes iguales verticalmente y se imponen las condiciones siguientes:

- Si las coordenadas del centro del elemento se encuentran en la parte izquierda de la ROI, pertenecen a la fosa nasal derecha.
- Si las coordenadas del centro del elemento se encuentran en la parte derecha de la ROI, pertenecen a la fosa nasal izquierda.

En el siguiente conjunto de imágenes de la figura 58, se puede observar el proceso explicado y la manera en la que se obtienen las fosas de manera individualizada. Además, en la figura 11 se muestran 2 ejemplos de la obtención de las narinas.



*Figura 58: Proceso de clasificación de las fosas nasales individualmente.*

*(Fuente: Elaboración propia)*

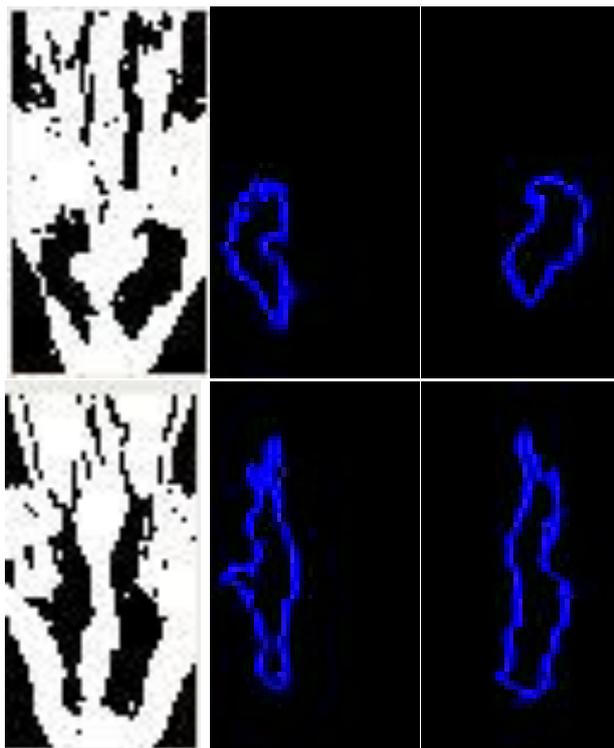


Figura 59: Ejemplos de obtención de las narinas individualmente.

(Fuente: Elaboración propia)

Ahora una vez obtenidas y almacenadas todas las fosas nasales, para cada uno de los cortes axiales donde se localizan, se procede a generar una nube de puntos de forma que se puedan visualizar de manera tridimensional.

Para ello, por cada una de las secciones, se itera a través de las matrices que contienen los índices que determinan la fosa nasal correspondiente. El bucle que se implementa recorre toda la matriz, y para cada uno de los puntos  $(i, j)$  comprueba si el valor del color en dicho punto es distinto de cero. Únicamente los índices donde se encuentre el contorno de la fosa cumplirán esta propiedad, ya que el resto de los puntos serán negros, en otras palabras, tienen asociado el valor cero por su color.

Los puntos que cumplan la condición se introducen en el archivo correspondiente a la nube de puntos. Asignándole a las coordenadas “X” e “Y” los valores de “i” y “j” respectivamente. El valor de la “Z” se asigna en función del corte axial en el que se encuentre.

De esta manera se construye el modelo tridimensional de las fosas nasales, el cual puede verse en la siguiente figura.

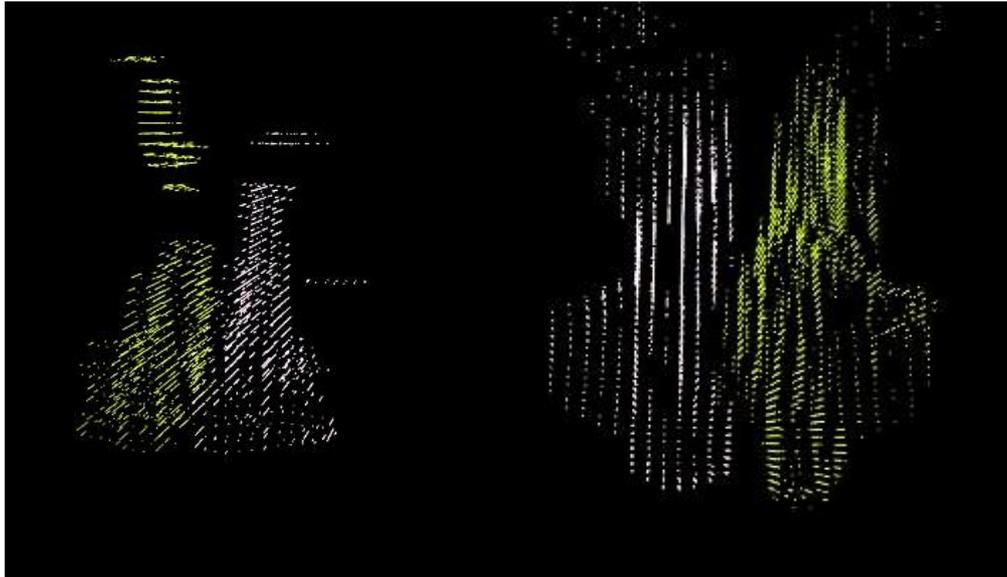


Figura 60: Vista frontal e inferior del modelo 3D de las fosas nasales.

(Fuente: Elaboración propia)

De manera que la nube de puntos obtenida sea más visual y fácil de contextualizar, se ha seguido un procedimiento parecido al explicado para determinar la parte externa de la nariz. Se ha ampliado la ROI y en cada uno de los cortes se almacena únicamente el contorno externo. El proceso seguido para la obtención de la nariz tiene como resultado lo siguiente:

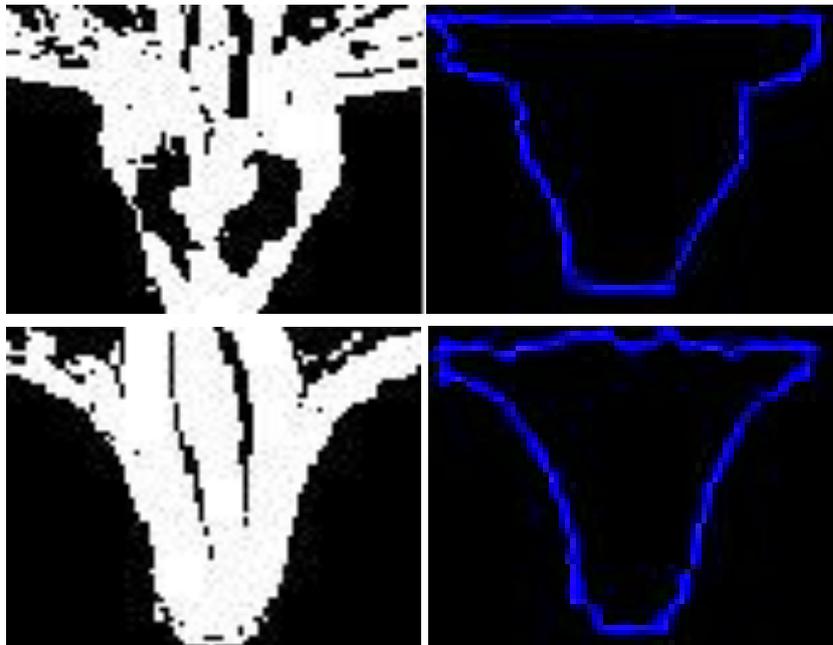


Figura 61: Ejemplos de obtención del contorno externo de la nariz.

(Fuente: Elaboración propia)

Realizando la misma iteración que la explicada anteriormente se genera un modelo completo, implementando ambas fosas nasales y la nariz, como el que se ve en las siguientes figuras.

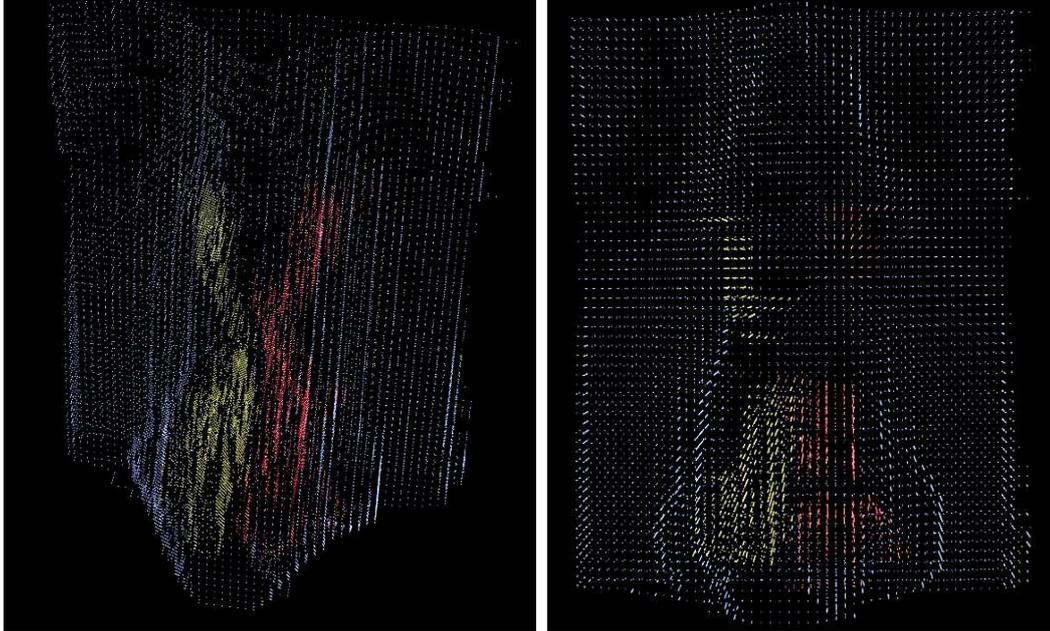


Figura 62: Vista frontal del modelo de las fosas nasales junto con el de la nariz.

(Fuente: Elaboración propia).

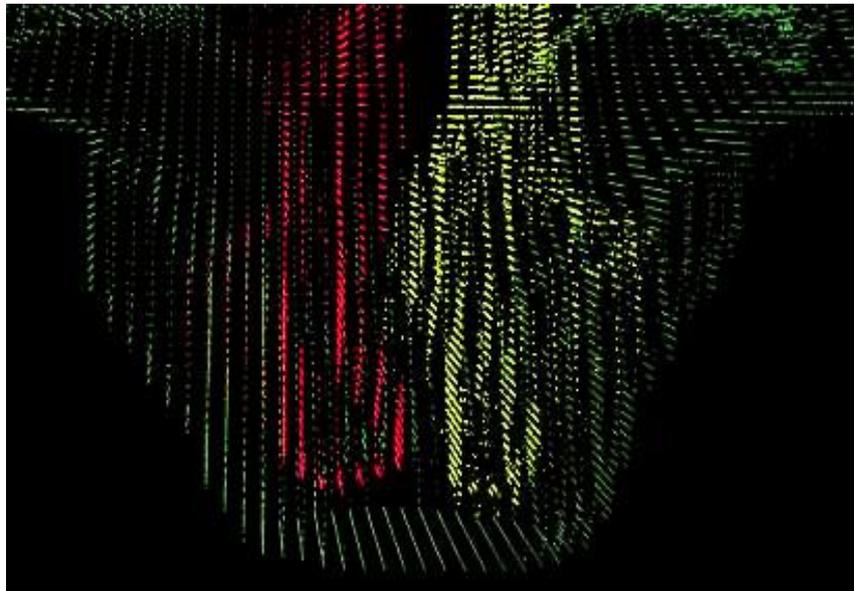


Figura 63: Vista inferior del modelo de las fosas nasales junto con el de la nariz.

(Fuente: Elaboración propia)



## 5.4 OBTENCIÓN TRIDIMENSIONAL DE LA CABEZA

En esta parte del análisis de las imágenes de resonancia magnética se busca determinar y obtener un modelo tridimensional de la cabeza completa del paciente. Para ello se puede hacer uso de cualquiera los tres conjuntos de imágenes de los planos anatómicos. En este caso, se va a implementar a través del plano axial. Se obtendrán dos modelos 3D de la cabeza. Por un lado, uno que contenga únicamente la información referida al contorno y, por otro lado, uno que contenga, además de la información del contorno, todos los detalles de las estructuras anatómicas internas de la cabeza.

Para ello se van a analizar todas las imágenes del plano axial, y para cada una de ellas se van a realizar una serie de pasos. Inicialmente para poder diferenciar entre la nube que contiene únicamente el contorno y la que aporta toda la información interna los pasos realizados son distintos. Esta diferencia, se explica a continuación.

### 5.4.1 Nube contorno

Para la obtención del contorno se realiza un procesado a cada una de las IRM del plano axial explicado en los siguientes pasos:

- Lectura de la imagen y su correspondiente binarización. En este caso no es necesario realizar una binarización adaptativa debido al método que se va a implementar. Por lo tanto, se aplica una binarización global, en la cual, si el valor del píxel es mayor que el del umbral, se le asigna un determinado índice definido. Se implementa a través del comando *threshold* ofrecido por OpenCV. Este proporciona diferentes tipos de binarización, entre ellos la binarización invertida, la cual se emplea en este proceso. La diferencia de estos tipos se aprecia en el siguiente conjunto de diagramas de la figura 64. El primero hace referencia a un valor de entrada, representado por una línea negra, y un determinado umbral, representado por una línea roja. El resto de los diagramas corresponden a algunos de los citados métodos. El primer par lo forman la binarización simple y la invertida, por otro lado, el segundo par consta de la truncada y la binarización a cero.

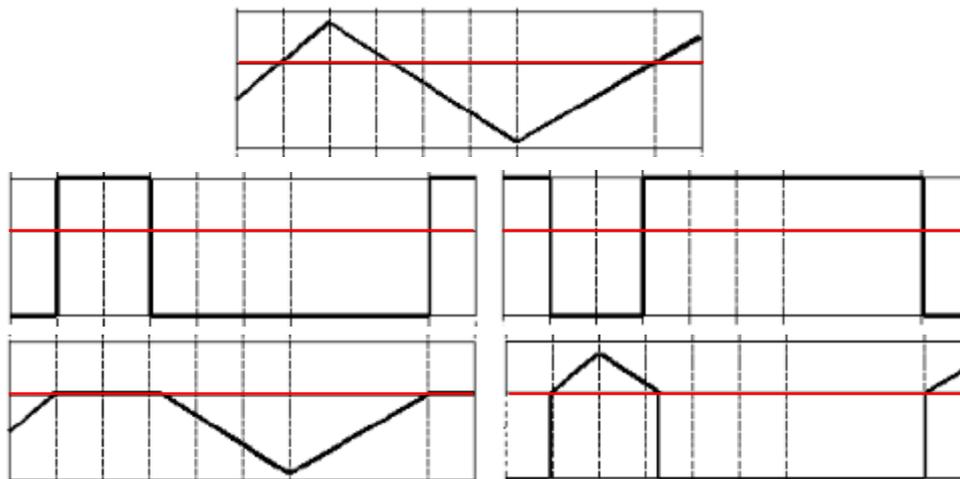


Figura 64: Diferentes binarizaciones.

(Fuente: «OpenCV» [En línea]. Disponible: <http://opencv.org/>)

Como previamente se ha citado, el método empleado en este caso es el de la binarización invertida, obteniendo el resultado mostrado en la siguiente figura.

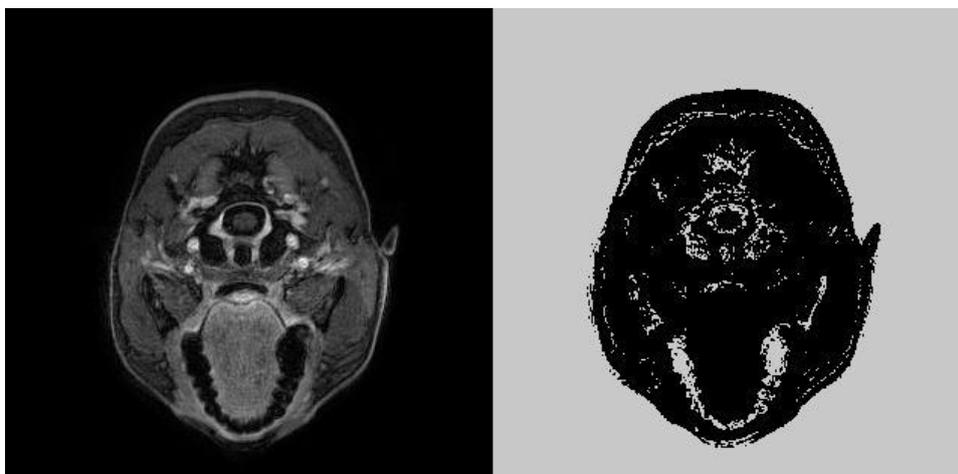


Figura 65: Imagen original y su correspondiente binarización inversa.

(Fuente: Elaboración propia)

- Una vez se tiene la imagen binarizada, se procede a rellenar esta. Para ello se utiliza la función *floodfill*, la cual llena un componente conectado a partir de un punto de partida; o semilla, con un color especificado. La conectividad la determina la cercanía del color y brillo de los píxeles vecinos. A través de este método se selecciona el conjunto de píxeles continuos del mismo color que el de la posición inicial, concretamente,

la (0, 0), y se sustituyen por el color correspondiente al valor 255, es decir, el blanco. Se puede establecer un rango de tolerancia si el color a extraer no es uniforme, pero la previa binarización que se ha realizado sobre la imagen evita la implantación de dicha característica. El resultado que se logra se muestra en la figura siguiente.

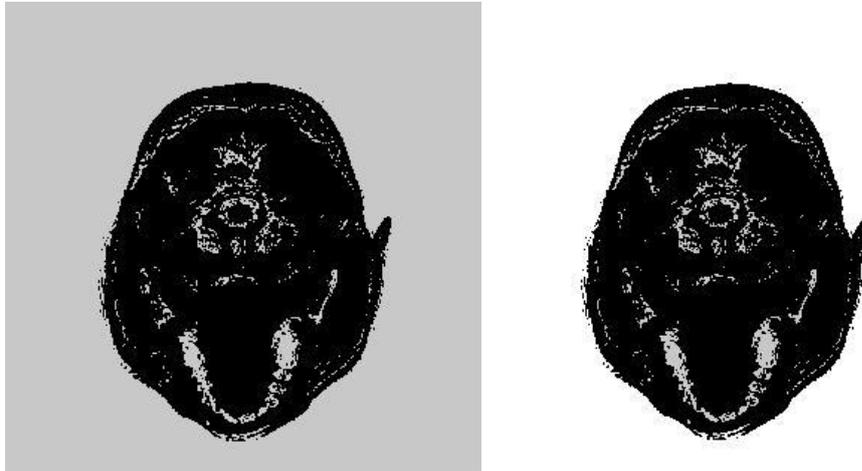


Figura 66: Resultado de implementar el relleno con floodfill.

(Fuente: Elaboración propia)

Una vez se obtiene la imagen rellena, se realiza una inversión de la completa para seguir su procesado. Por lo tanto, finalmente se obtiene la imagen mostrada en la siguiente figura.

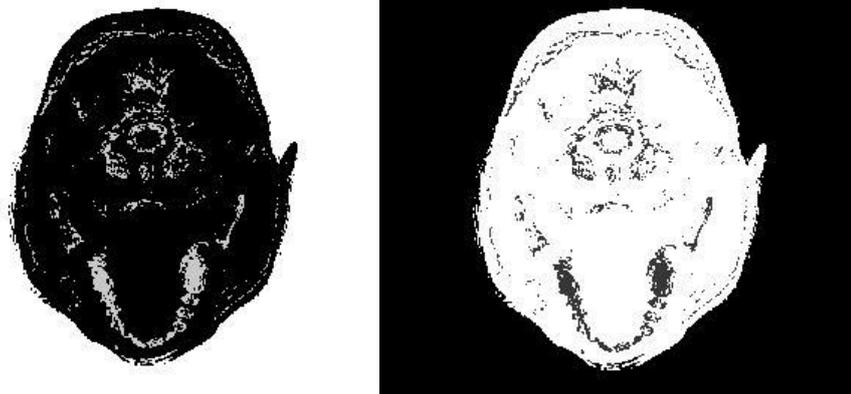


Figura 67: Imagen rellena y su correspondiente inversa.

(Fuente: Elaboración propia)

- Finalmente se procede a combinar la imagen binarizada obtenida en el primero de los pasos y la imagen rellenada e invertida del paso número dos. El proceso de unificación se muestra en el conjunto de imágenes de la figura 68. Se puede decir que se realiza una “suma” de las dos primeras imágenes, para dar lugar a las imágenes situadas más a la derecha.

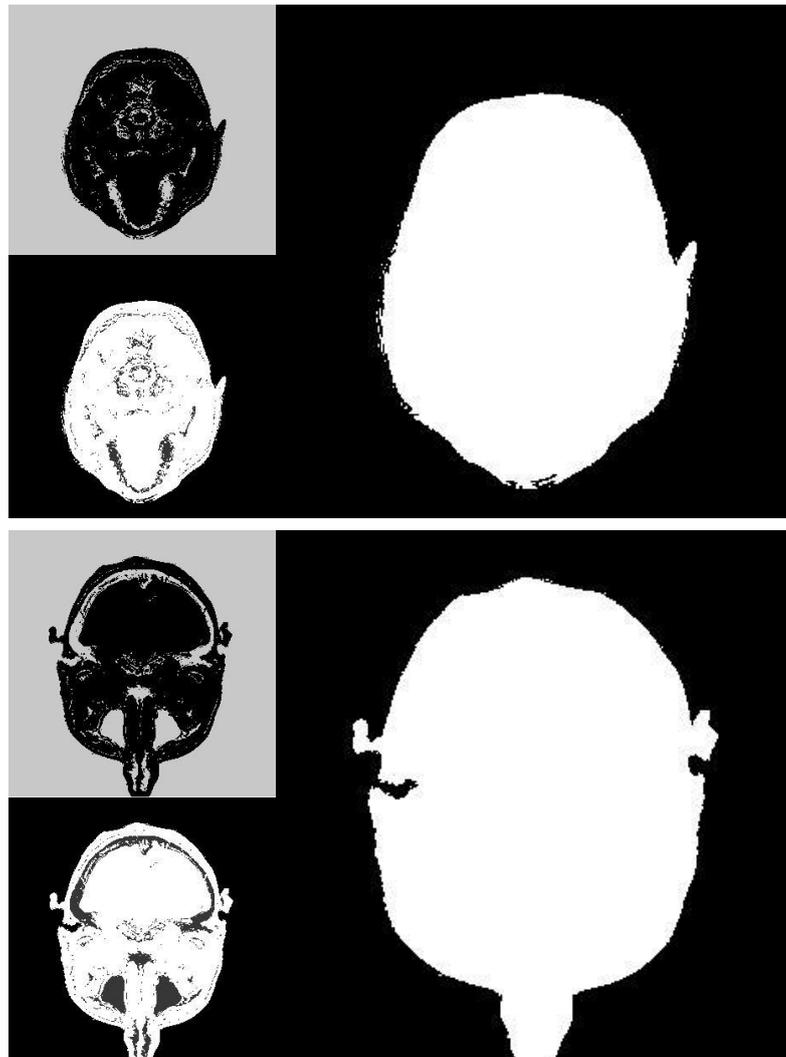


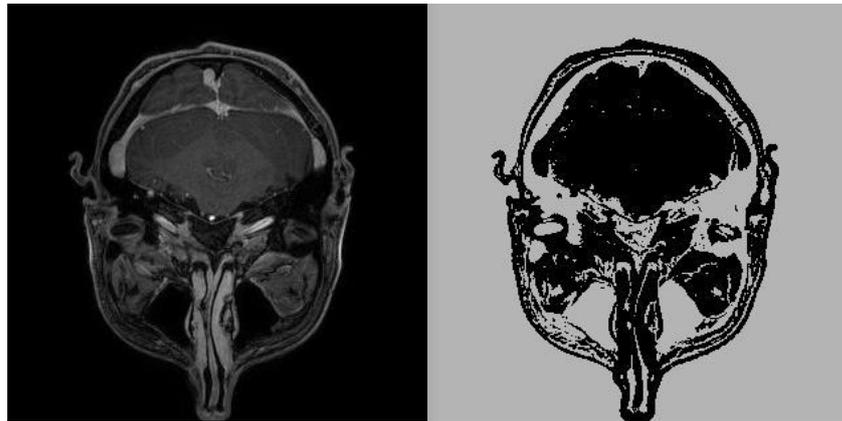
Figura 68: Ejemplos del resultado obtenido con las imágenes combinadas.

(Fuente: Elaboración propia)

#### 5.4.2 Nube completa

En este caso, para obtener toda la información perteneciente al interior de la cabeza, así como del contorno de esta, únicamente se realiza una

binarización global a cada una de las imágenes. Se hace uso del mismo método de binarización que se emplea en el primero de los pasos para la detección del contorno de las imágenes. La diferencia entre ambos reside en el ajuste del umbral. A modo de ejemplo, puede apreciarse en la siguiente figura una demostración de esta binarización.



*Figura 69: Imagen original y su correspondiente binarización.*

*(Fuente: Elaboración propia)*

Ahora sí, el proceso que se sigue a continuación es el mismo para la generación de ambas nubes, por lo tanto, se explica de manera conjunta.

Para cada una de las imágenes se buscan los contornos con el método explicado en el apartado correspondiente a la obtención tridimensional de las narinas. Por lo tanto, se pretende localizar todos aquellos puntos continuos, a lo largo de un límite, con el mismo color o intensidad. El objetivo es identificar cualquier estructura anatómica para el caso de la nube completa, y en el caso del contorno detectar únicamente los puntos correspondientes al exterior de la cabeza. En el siguiente conjunto de imágenes de la figura 70 se ve el resultado de buscar los diferentes contornos de los elementos que conforman cada imagen. Para su representación se ha asignado un color aleatorio entre el verde, el rojo y el azul a cada contorno encontrado con el fin de poder distinguir un elemento de otro con mayor claridad a la hora de la visualización.

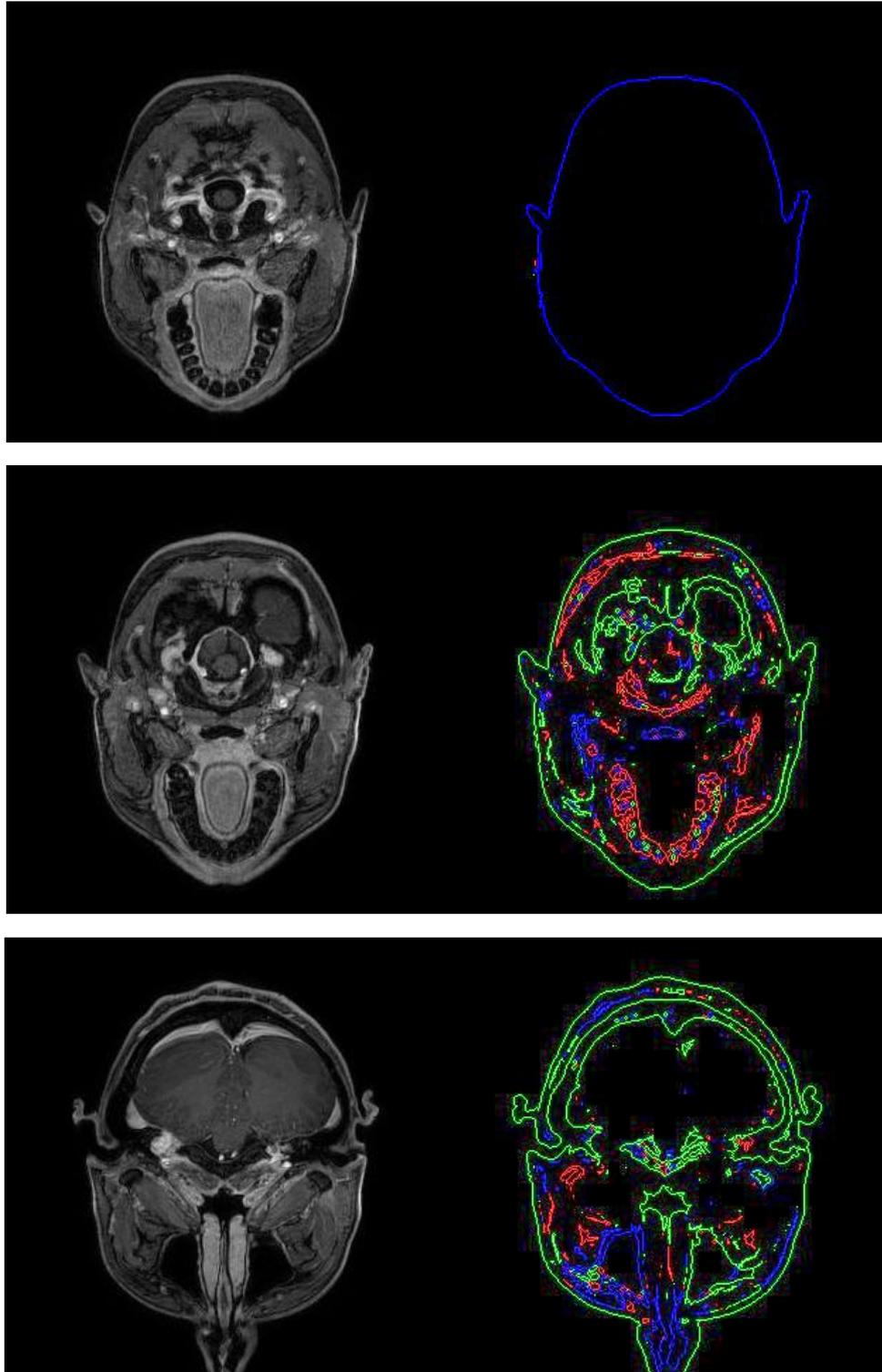
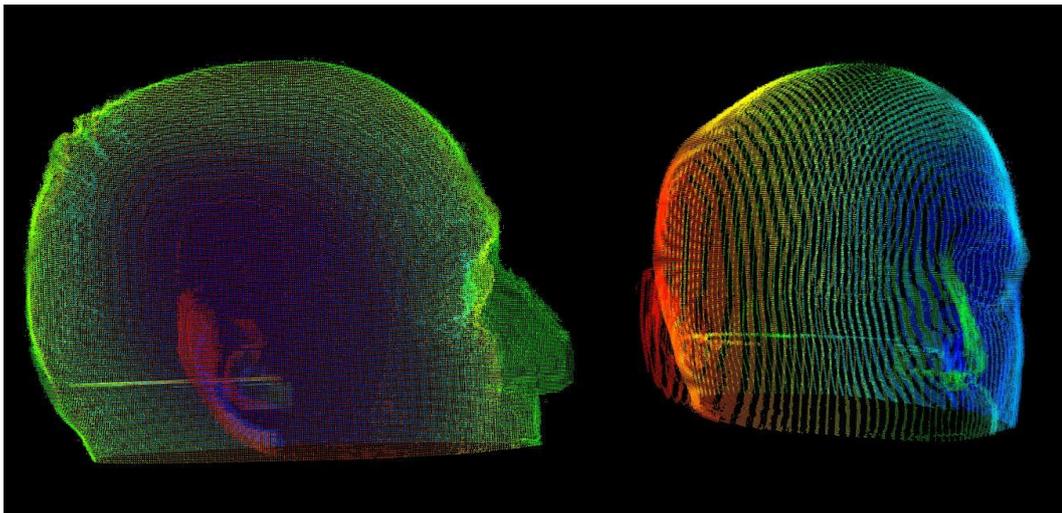


Figura 70: Ejemplos de la obtención de los contornos en cada corte de las IRM.

(Fuente: Elaboración propia)

En la primera imagen se muestran los contornos encontrados para el caso de la nube encargada del modelo externo de la cabeza y, en las dos imágenes siguientes, los contornos referentes a la nube con toda la información.

Una vez se han procesado todas las imágenes se da paso a la generación de las nubes de puntos. Para ello, se vuelve a utilizar el método empleado en la generación tridimensional de las fosas nasales. Se analizan todas las imágenes creadas y para cada una de ellas se itera a través de su correspondiente matriz en busca de aquellos puntos  $(i, j)$  donde el valor del color es distinto de cero. Son estos los puntos que se introducen en la nube, excluyendo el resto. Por lo tanto, el resultado que se obtiene en este caso es el mostrado a continuación en las figuras 71, correspondiente al modelo externo, e 72, correspondiente al modelo completo. En el caso de estas nubes, se han representado coloreadas en función de la profundidad de los puntos para una interpretación más clara.



*Figura 71: Modelo tridimensional correspondiente al exterior de la cabeza.*

*(Fuente: Elaboración propia)*

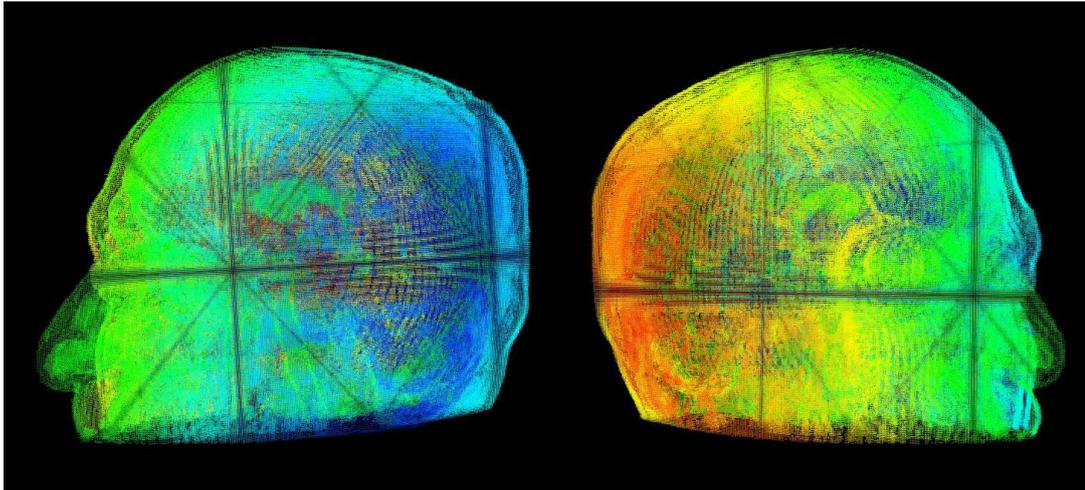


Figura 72: Modelo tridimensional completo de la cabeza.

(Fuente: Elaboración propia)

## 5.5 OBTENCIÓN TRIDIMENSIONAL DEL CEREBRO

Un proceso similar al citado en los dos subapartados anteriores puede aplicarse para la obtención de distintas estructuras anatómicas que se pueden encontrar en la cabeza. Concretamente en este caso, se ha desarrollado un procedimiento para la obtención del cerebro humano. En esta ocasión se parte de los cortes del plano anatómico sagital para la obtención tridimensional. Los pasos realizados en cada una de las imágenes son los siguientes:

- Una vez se ha leído la imagen en la iteración correspondiente, se aplica un filtro de suavizado. Se hace uso del filtro Gaussiano explicado en el subapartado 6.3, teniendo de nuevo el objetivo de reducir el ruido y conseguir que las intensidades de pequeños elementos se mezclen con el fondo de manera que se detecten únicamente las estructuras de mayor tamaño.
- Posteriormente se procede a la binarización de la imagen por medio de un umbral adaptativo. Como ya se ha explicado anteriormente, este método de binarización calcula un valor umbral diferente para cada una de las regiones en las que se divide la imagen, obteniendo así, resultados más detallados y precisos. En la siguiente figura se puede ver el resultado de esta binarización.

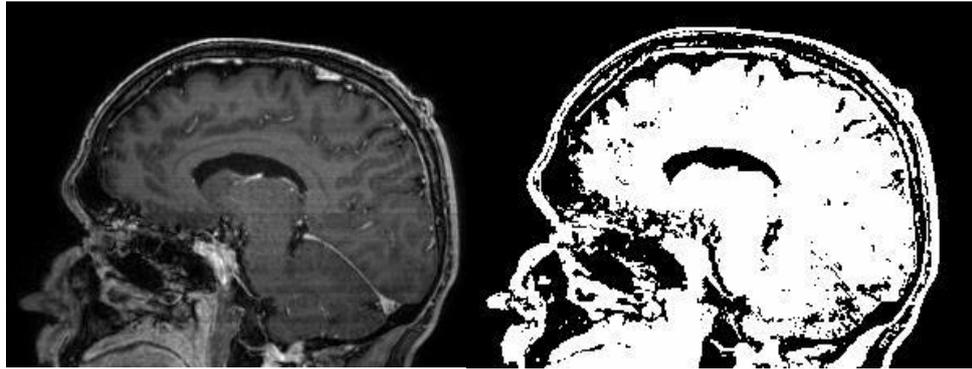


Figura 73: Imagen original y su correspondiente binarización global.

(Fuente: Elaboración propia)

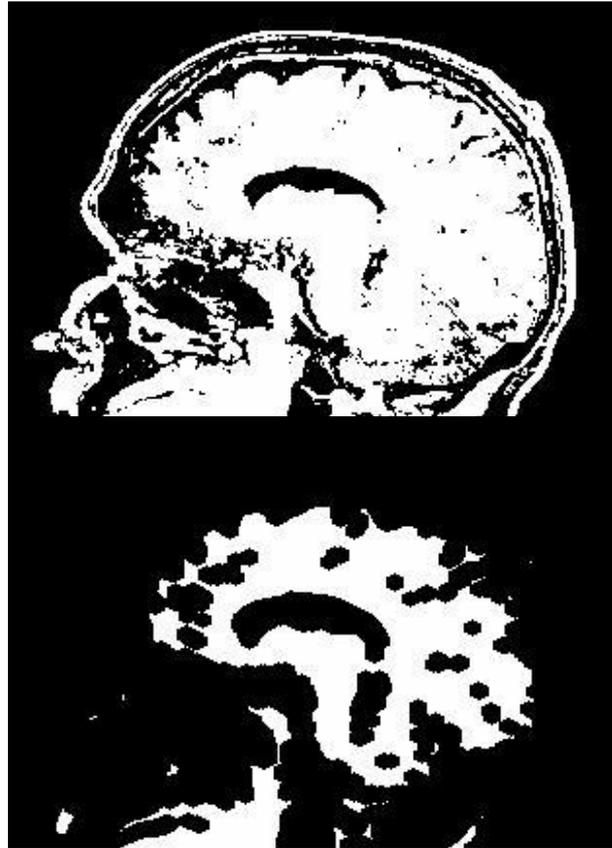
El trabajar con estas imágenes binarizadas, las cuales esta construidas por matrices formadas de ceros y unos, facilitan su posterior análisis y procesamiento.

- Ahora, con la imagen binaria, se procede a realizar una serie de operaciones morfológicas sobre ella. Primero se implementa una erosión y posteriormente un cierre y finalmente se refuerza con una dilatación. Con estas transformaciones se busca eliminar los elementos que no interesan y dar más relevancia a los pertenecientes a la estructura cerebral.

El elemento estructural utilizado en este proceso es elíptico, y a modo de ejemplo, para un tamaño del kernel de 5x5 se obtendría la siguiente matriz haciendo referencia al citado elemento estructural.

$$ee = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

En la siguiente figura se puede apreciar la erosión realizada sobre la imagen previamente binarizada, obteniendo como resultado la imagen situada en la parte inferior de la figura.



*Figura 74: Resultado de aplicar la erosión sobre la imagen.*

*(Fuente: Elaboración propia)*

Posteriormente, se implementa un cierre. Esta transformación morfológica consiste básicamente en una dilatación seguida de una erosión, pero en este caso, se añade una dilatación a mayores para mejorar los efectos. Por lo tanto, el resultado que se obtiene del conjunto de estas dos transformaciones es el que se muestra en la siguiente figura.



Figura 75: Resultado de aplicar el cierre y la dilatación de refuerzo sobre la imagen.

(Fuente: Elaboración propia)

- Ahora se procede a buscar los contornos de la imagen transformada morfológicamente. Como ya se sabe, esta implementación consiste en localizar todos aquellos puntos continuos, a lo largo del límite, con el mismo color o intensidad. De esta manera, es posible diferenciar las diferentes estructuras presentes en cada corte. A modo de ejemplo, se muestra en la figura siguiente los contornos encontrados para el corte sagital obtenido en el paso anterior.



Figura 76: Representación de todos los contornos encontrados en la imagen.

(Fuente: Elaboración propia)

- Una vez llegado a este punto, interesa almacenar únicamente los contornos referidos al cerebro. Debido a esto, debe implementarse algún criterio para su identificación y selección.

En este caso, al igual que en la localización de las fosas nasales, se han calculado las coordenadas de los centros de masas de cada uno de los elementos encontrados en la búsqueda de contornos. La discriminación se vuelve a realizar en función de la ubicación de dichos puntos y el área correspondiente al elemento. El criterio de posición se cumple si las coordenadas del centro de masas se encuentran dentro de una circunferencia establecida.

Para determinar esta circunferencia, inicialmente se ha elegido ubicar su centro en una posición aproximada a la mitad del cerebro. Posteriormente se ha asignado un radio, respecto al citado punto, que define la superficie de aceptación para la posición de los centros de masas. En la siguiente figura puede apreciarse el centro de la circunferencia, representada en rojo, y la región de aceptación de las coordenadas de los centros de masas, delimitada por la curva de color verde.

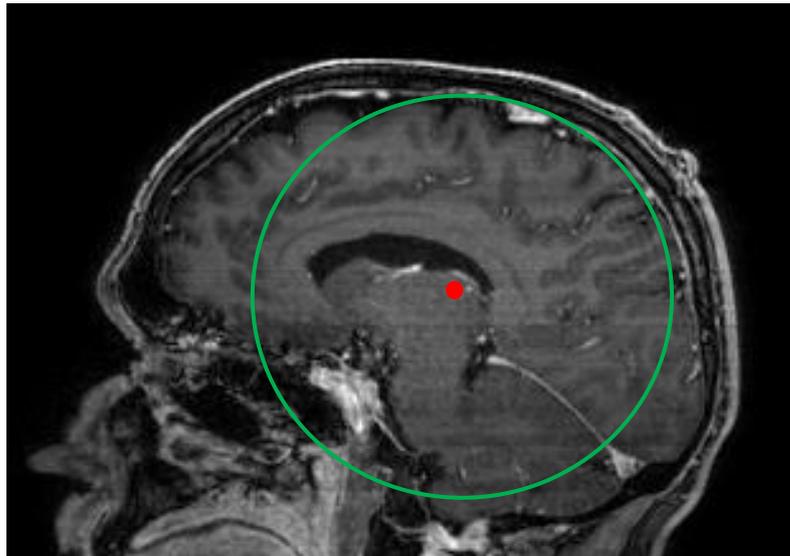


Figura 77: Definición de la región de aceptación junto con su centro.

(Fuente: Elaboración propia)

Por lo tanto, con los contornos y sus respectivos centros de masas encontrados, a continuación, se muestra el ejemplo donde se visualiza cuáles son los elementos que cumplen la condición de pertenecer a la región de aceptación, representados por el color verde, y cuáles no, representados por el color rojo. Además de este requisito, también se excluyen aquellas superficies cuyo valor de área sea muy pequeño. En

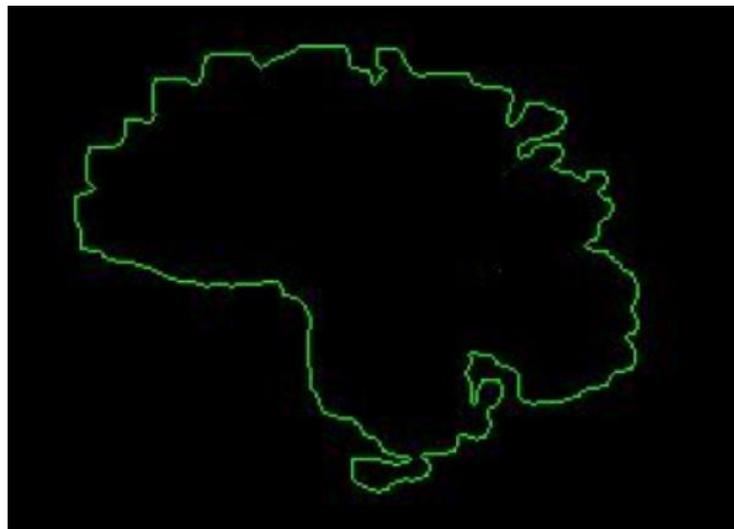
la figura siguiente se muestra cual es el resultado de aplicar las citadas condiciones.



*Figura 78:* Representación de los elementos en función del cumplimiento de las condiciones.

(Fuente: Elaboración propia)

Inmediatamente después de identificar cuáles son los contornos de los elementos válidos, se obtienen de manera individual como se muestra en la siguiente figura.



*Figura 79:* Representación del contorno valido que cumple las condiciones.

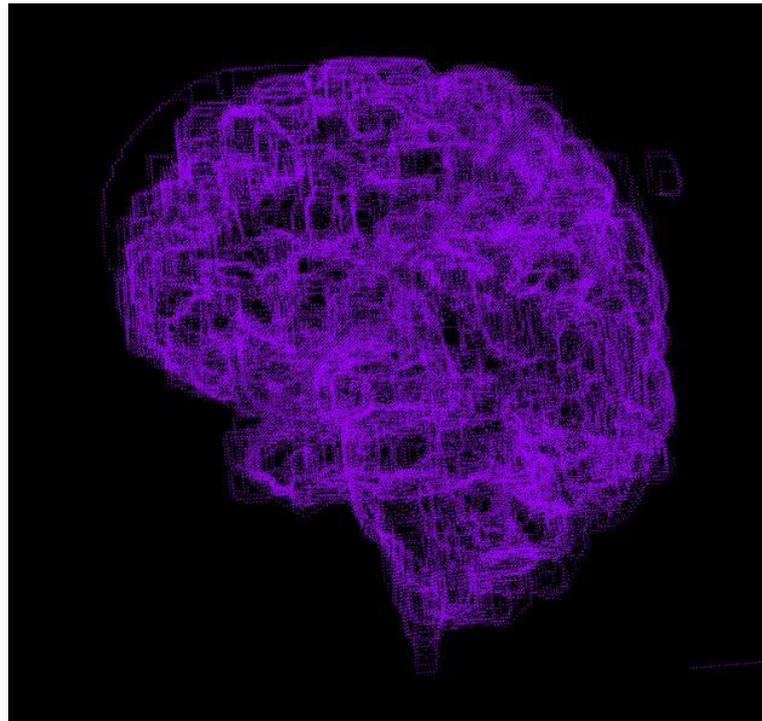
(Fuente: Elaboración propia).

Inmediatamente obtenidos y almacenados todos los cortes cerebrales, para cada una de las imágenes del plano anatómico sagital, se procede a generar una nube de puntos de forma que se pueda visualizar de manera tridimensional.

Para ello, por cada una de las imágenes, se itera a través de las matrices que contienen los índices que determinan el corte cerebral correspondiente. El bucle que se implementa recorre toda la matriz, y para cada uno de los puntos (i, j) comprueba si el valor del color en dicho punto es distinto de cero. Únicamente los índices donde se encuentre el contorno del corte cerebral cumplirán esta propiedad, ya que el resto de los puntos tienen asociado el valor '0'.

Los puntos que cumplan la condición se introducen en el archivo correspondiente a la nube de puntos. Asignándole a las coordenadas "X" e "Y" los valores de "i" y "j" respectivamente. El valor de la "Z" se asigna en función del corte sagital en el que se encuentre el bucle.

De esta manera se construye el modelo tridimensional cerebral, el cual puede verse en la siguiente figura.

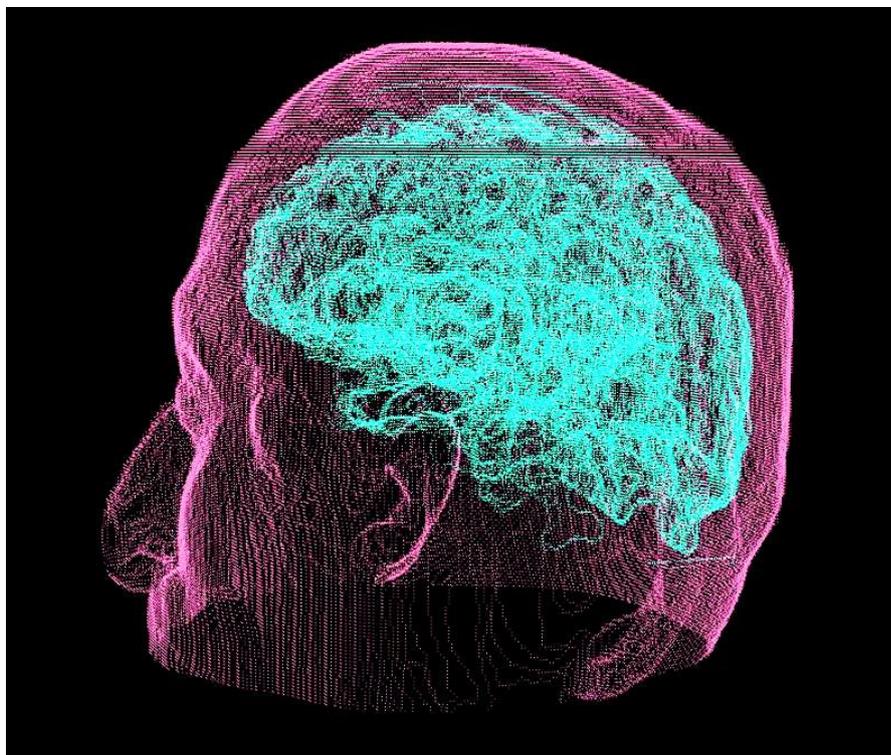


*Figura 80: Representación del modelo tridimensional cerebral creado a partir de las IRM.*

*(Fuente: Elaboración propia)*

Para visualizarlo en contexto con la cabeza del paciente, se hace uso del modelo tridimensional externo generado en el subapartado anterior de este

capítulo. De esta manera, el resultado que se obtiene al visualizar simultáneamente ambas nubes se muestra en la siguiente figura.



*Figura 81: Representación del modelo cerebral creado junto con el contorno exterior de la cabeza.*

*(Fuente: Elaboración propia)*

## 5.6 SISTEMA DE NAVEGACIÓN QUIRÚRGICO

Otro aspecto importante en esta parte del desarrollo es la creación de un sistema de navegación quirúrgico 3D en tiempo real. Este sistema permite a los médicos acceder dinámicamente a lo largo de la operación a los distintos modelos o secciones tridimensionales creados de la cabeza del paciente. De esta manera, se pueden adquirir y transmitir datos adicionales que mejoren características críticas en los abordajes transesfenoidales como pueden ser la posición y orientación que debe mantener el robot asistencial, así como la trayectoria que esté obligado a seguir.

Debido a esto, se han implantado una serie de menús que permiten elegir cual es la vista que se desea visualizar. Para ello, inicialmente se da a elegir entre los tres planos anatómicos construidos, es decir, el plano axial, el sagital

y el coronal. Posteriormente se selecciona el tipo de visualización que se desea tener, pudiendo seleccionar entre la vista del modelo completo, del modelo hasta una determinada sección o de una sección específica perteneciente al modelo. Finalmente, independientemente de la vista elegida, se selecciona si se quiere visualizar el modelo tridimensional con toda la información interna o el modelo que contiene únicamente la información referida al contorno de la cabeza.

En el conjunto de imágenes de la siguiente figura se muestran los menús recientemente citados.

```
***** PLANO *****
* Seleccione el plano anatomico que quiere visualizar: *
* 1 -> Plano axial *
* 2 -> Plano sagital *
* 3 -> Plano coronal *
* 4 -> Salir *
*****
SELECCION PLANO: █

***** VISTA *****
* Seleccione la vista de la seccion que quiere visualizar: *
* 1 -> Vista del volumen completa *
* 2 -> Vista del volumen hasta una seccion *
* 3 -> Vista de una seccion individualmente *
* 4 -> Menu anterior *
* 5 -> Salir *
*****
SELECCION VISTA: █

***** FORMATO *****
* Seleccione el formato en el que quiere visualizarla: *
* 1 -> Contornos *
* 2 -> Completa *
* 3 -> Menu anterior *
* 4 -> Salir *
*****
SELECCION FORMATO: █
```

Figura 82: Diferentes menús creados para la navegación quirúrgica.

(Fuente: Elaboración propia)

En el caso de seleccionar la segunda o la tercera opción en el menú “VISTA”, se muestra por pantalla un menú auxiliar con el cual se permite definir la sección de partida de la visualización. En función del plano anatómico que se haya seleccionado previamente, se muestra el rango de cortes que contiene dicho plano. En la siguiente figura se puede ver el citado menú auxiliar.

```
***** SECCION *****
* Seccion -> Entre 1 y 200 *
* -1 -> Menu anterior *
* -2 -> Salir *
*****
SELECCION SECCION: █
```

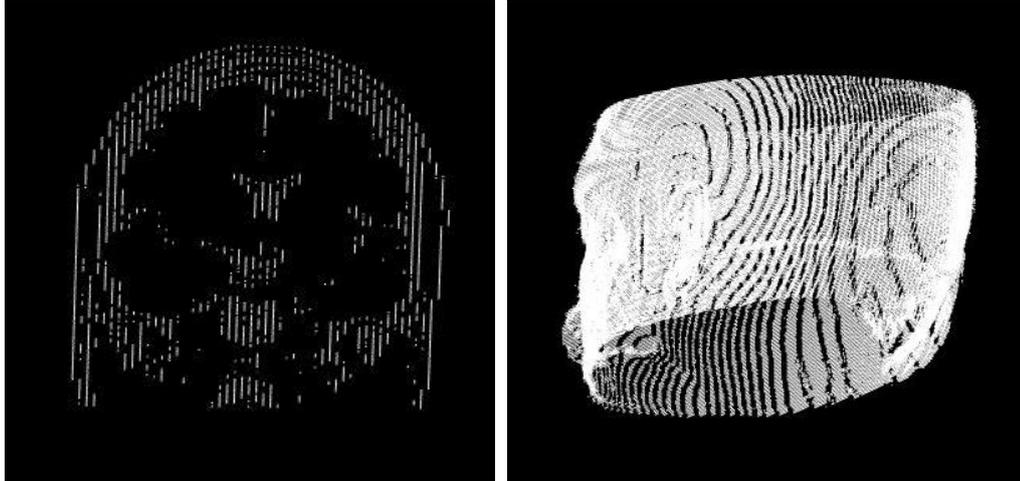
Figura 83: Menú auxiliar para la selección de sección.

(Fuente: Elaboración propia)

Finalmente, para poder navegar en tiempo real a lo largo de los diferentes modelos tridimensionales, se han definido diferentes funcionalidades. Estas vienen resumidas en los siguientes puntos:

- Incremento de sección: Se puede incrementar de uno en uno o de diez en diez el número de secciones.
- Decremento de sección: Se puede decrementar de uno en uno o de diez en diez el número de secciones.
- Modificación del punto de vista
- Volver al menú de selección

Con las dos primeras se podrá ir añadiendo o quitando “capas” a nuestra nube de puntos, de manera que se pueda acceder a visualizar exactamente la zona que el cirujano determine. Con la tercera funcionalidad, se permite modificar la perspectiva desde la cual se observa la nube de puntos. Esta operación es de gran utilidad, ya que permite visualizar cada estructura anatómica desde el ángulo que más interese o más información proporcione. Con la última de las funcionalidades se vuelve a acceder a los menús de selección. En el conjunto de imágenes de la siguiente figura, se pueden ver dos ejemplos de visualización de los modelos seleccionados a través de los mencionados menús de navegación. A mano izquierda se visualiza una sección individual del plano coronal, y, por otro lado, a mano derecha se muestra la visualización del contorno de la cabeza hasta una determinada sección del plano axial.



*Figura 84: Sección individual del plano coronal (izquierda) y superficie externa axial hasta una sección determinada (derecha).*

*(Fuente: Elaboración propia)*



## 6 CONCLUSIONES Y LÍNEAS FUTURAS

---

### 6.1 INTRODUCCIÓN AL CAPÍTULO

A continuación, se procede a realizar una reflexión acerca de los resultados obtenidos, comprobando si se han cumplido los objetivos establecidos. Además, se citarán las conclusiones extraídas debido al desarrollo llevado a cabo junto con algunas de las líneas futuras aplicables al proyecto.

### 6.2 CONCLUSIÓN DEL PROYECTO

Durante el desarrollo de este Trabajo Fin de Grado se han realizado dos desarrollos de visión artificial destinados a proporcionar información útil para ayudar y asistir al cirujano en los abordajes endoscópicos endonasales transesfenoidales. Por un lado, se ha trabajado con la generación de nubes en tres dimensiones, empleando para su captura una cámara de tiempo de vuelo, con su posterior procesado y, por otro lado, se ha desarrollado un análisis de las imágenes de resonancia magnética para la obtención de modelos tridimensionales de diferentes elementos estructurales de la cabeza. Para ello, se ha hecho uso de unas de las herramientas libres más potentes existentes en la actualidad, de las librerías de OpenCV y PCL. A través de ellas, se han podido realizar toda serie de operaciones con el objetivo de obtener los mejores resultados posibles.

Los principales objetivos son dos. El primero de ellos consiste en poder realizar el emparejamiento, o registro, de las diferentes nubes de puntos obtenidas a través de la cámara ToF. El segundo trata de localizar estructuras anatómicas, concretamente en este caso, las fosas nasales del paciente a través de las imágenes de resonancia magnética. Ya que para lograr estos objetivos se han implementado dos desarrollos paralelos, se analizarán los resultados de ambos por separado.



### 6.2.1 Resultados cámara ToF y registro de nubes de puntos

En el capítulo 4 se explicaba el método por el cual se obtenían las nubes y como, posteriormente, se realizaba un procesado sobre ellas para obtener el modelo 3D completo. Cabe destacar, que con el método inicial no se alcanzó el objetivo establecido, debido al ruido presente en las nubes capturadas. Sin embargo, con la alternativa de incorporar al proceso el robot UR3, los resultados mejoran drásticamente. En los resultados obtenidos en el citado capítulo se puede observar el registro de las tres nubes de puntos obtenidas a través de las capturas en una sola, donde puede visualizarse con nitidez la cabeza, el cuello y la parte superior del torso del paciente. Es importante decir que las superficies siguen manteniendo un ruido considerable, pero ahora es mucho más fácil emparejar dicho modelo tridimensional obtenido con las cámaras ToF con otro generado por el análisis de las IRM.

Por lo tanto, se puede decir que los objetivos referidos a la creación de un modelo 3D a partir de diferentes capturas de nubes de puntos han sido completados satisfactoriamente.

### 6.2.2 Resultados análisis de IRM

El desarrollo, referido al procesado de las imágenes obtenidas a través de la resonancia magnética sobre la cabeza del paciente, ha sido explicado en el capítulo 5. El principal objetivo de este proyecto consiste en localizar la cavidad de las fosas nasales, con el fin de proporcionar al robot asistencial una posible localización de posicionamiento y una trayectoria a seguir. Con el desarrollo implementado, cabe destacar que los resultados han sido obtenidos en su plenitud.

La obtención de las fosas nasales ha sido clara y detallada. Además, se ha podido diferenciar entre la fosa nasal izquierda y derecha, proporcionando esto una gran utilidad a la hora de la operación, debido a que pueden desarrollarse trayectorias completamente individualizadas. Por otro lado, se ha podido visualizar el modelo de las fosas nasales junto con el de la nariz del paciente, ya que se ha detectado el contorno de ella, de modo que la visualización sea más clara.

Otro de los objetivos de este desarrollo era la generación de un sistema de navegación novedoso que pueda ayudar al cirujano en los abordajes. Debido a esto, se ha desarrollado un sistema que genera una visualización en tiempo real de todos los planos generados en este proyecto, es decir, el axial, el sagital y el coronal. Es posible implementar cualquier vista de los distintos modelos



tridimensionales completos, parciales o por secciones individuales. Además, se puede considerar que se trata de una navegación en tiempo real debido al bajo tiempo de ejecución que tiene. Por lo tanto, se puede concluir que el citado objetivo ha sido completado exitosamente a través del sistema generado y que, además, mejora la mayoría de los actuales sistemas de navegación. Esta mejora reside en la integración de la tercera dimensión, ya que en la actualidad, gran parte de dichos sistemas únicamente trabajan con imágenes en dos dimensiones.

Por otro lado, se han creado modelos tridimensionales de diferentes estructuras de la cabeza del paciente, además del de las fosas nasales. Esto complementa los objetivos inicialmente marcados, ya que proporciona una mayor contextualización de las diferentes nubes 3D obtenidas y aporta mayor información para el abordaje.

Por un lado, se ha generado un modelo tridimensional de la cabeza del paciente. Este modelo contiene dos versiones, las cuales se diferencian en la cantidad de datos que tienen por sección. Hay un modelo que contiene únicamente la información tridimensional del contorno de la cabeza, y otro, que además de esa información, contiene todos los datos internos. El resultado obtenido, en este caso es muy preciso y puede ser utilizado para realizar un *matching*, en otras palabras, casar dicho modelo con el obtenido por otros medios. Concretamente en este caso, puede realizarse el *matching* con el modelo generado a través de la cámara de tiempo de vuelo.

Por otro lado, además del modelo de la cabeza, también se ha obtenido el cerebro en una nube de puntos. Para ello se ha almacenado la corteza cerebral en cada una de las secciones, dando como resultado a un modelo que almacena el cerebro de manera individual y otro que lo almacena junto con el modelo del contorno de la cabeza. Este resultado es de menor precisión, ya que se pierde información característica en los contornos del cerebro en cada una de las secciones debido a las diferentes transformaciones morfológicas que estas sufren. A pesar de ello, de manera aproximada, es de gran utilidad ya que permite conocer su posición, tamaño o incluso si padece de alguna malformación.

Por lo tanto, se puede concluir que el cumplimiento de los objetivos en ambos desarrollos, y, por lo tanto, el del trabajo, han sido completos. Cabe destacar, que en la realización de dicho trabajo se han generado ideas sobre posibles mejoras y líneas futuras que se comentaran en el siguiente capítulo.



## 6.3 LÍNEAS FUTURAS

### 6.3.1 Modulo con acelerómetro/giroscopio

En la actualidad la cabeza del paciente se fija firmemente en una posición concreta antes de comenzar el proceso quirúrgico, con el objetivo de inmovilizarla durante todo el abordaje. Es cierto que esta queda en una posición estática en la cual no se le permite realizar ningún movimiento, pero cabe la posibilidad de que no se fije bien o que por motivos quirúrgicos la cabeza sufra ciertos movimientos. Ante dicha situación, actualmente el robot sería incapaz de reaccionar debido a que no posee información de la posición y orientación de la cabeza en tiempo real. Por lo tanto, una línea futura de gran importancia es generar algún modelo que proporcione datos en tiempo real.

Para ello, una posible solución que no supone modificación alguna en todo el sistema creado es la incorporación de sensores que detecten cambios en la posición y orientación. Dichos sensores corresponden al acelerómetro y al giroscopio, que vienen integrados en multitud de tarjetas electrónicas. A modo de ejemplo podría implementarse con un módulo MPU 9250. De esta manera, cualquier movimiento realizado por la cabeza del paciente durante las operaciones quedaría registrado y podría transferirse toda la información necesaria al robot para que este modifique la trayectoria en tiempo real con el objetivo de no dañar al paciente.

### 6.3.2 Colorear modelos tridimensionales

Otra de las posibles mejoras es la de colorear las nubes de puntos en función de nivel de gris. Como se explicaba en el quinto punto del capítulo 2, cada píxel de las IRM tiene asociado un nivel de gris en función de la estructura anatómica que represente. Esta diferenciación surge debido a la composición química que tiene cada estructura. Por lo tanto, si en la lectura de la imagen se discrimina en función de su intensidad de color, podrían almacenarse y crearse de una manera directa y rápida modelos tridimensionales con diferentes estructuras anatómicas. A modo de ejemplo, esta posible mejora podría ayudar a la obtención de todas las estructuras portadoras de sangre de la cabeza y posteriormente visualizarlas a través de una nube de puntos. De este modo, se mejora drásticamente la visualización y la claridad individualizada de



las diferentes estructuras anatómicas que se pueden encontrar y, por lo tanto, se obtienen análisis más precisos y con mayor detalle.



Universidad de Valladolid

# Integración de un sistema de visión 3D en un entorno robotizado de cirugía Ander Loidi Yarza



ESCUELA DE INGENIERÍAS  
INDUSTRIALES



## 7 BIBLIOGRAFÍA

---

- [1] Agamennoni G., Fontana S., Siegwart R. y Sorrenti D. (2016). "Point clouds registration with probabilistic data association". *IEEE/RSJ International Conference on Intelligent Robots and Systems. IROS 2016*, 4092-4098.
- [2] Ahmed M., Mohamad M., Marshall J. y Greenspan M. (2015). "Registration of noisy point clouds using virtual interest points". *12th Conference on Computer and Robot Vision. CRV 2015*, 31-38.
- [3] Alberich J., Gómez D. y Ferrer A. (2013). "*Percepción visual*". Barcelona: Universitat Oberta de Catalunya.
- [4] Alegre E. (Ed.), Pajares G. (Ed.) y Escalera A (Ed.). (2016). "*Conceptos y métodos en Visión por computador*". España: Grupo de Visión del Comité Español de Automática (CEA).
- [5] Armengot M., Gallego J.M., Gómez M.J., Barcia J.A., Basterra J. y Barcia C. (2011). "Transphenoidal endoscopic approaches for pituitary adenomas: a critical review of our experience". *Acta Otorrinolaringológica Española*, 62(1): 25-30.
- [6] Ayache N. (1995). "Medical computer vision, virtual reality and robotics". *Image and Vision Computing*, 13(4): 295-313.
- [7] Ballard D. y Brown C.M. (1982). "*Computer Vision*". New Jersey: Prentice Hall.
- [8] «Basler ToF Camera» [En línea]. Disponible: <https://www.baslerweb.com/en/>. [Accedido: 07-Jun-2019].
- [9] Besl P.J. y McKay D. (1992). "A method for registration of 3-D shapes". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2), 239-256.
- [10] Buxton R., Edelman R., Rosen B., Wismer G. y Brady T. (1987). "Contrast in Rapid MR Imaging: T1 and T2 Weighted Imaging". *Journal of Computer Assisted Tomography*, 11(1): 7-16.
- [11] Danqui C., Jian X., Ping Z., Xiang H. y Ming X. (2014). "Pituitary adenoma with gangliocytoma: Reporto of two cases". *Oncology Letters*, 8(2), 781-784.



- [12] «Da Vinci Surgery | Da Vinci Surgical System | Robotic Technology» [En línea]. Disponible: <https://www.davincisurgery.com/>. [Accedido: 28-May-2019].
- [13] Dubey S.P., Munjal V.R. (2014) “Endoscopic Endonasal Transsphenoidal Hypophysectomy: Two Hand Versus Four Hand Technique”. *Indian J Otolaryngol Head Neck Surg*, 66(3): 287–90.
- [14] Fuente E. y Trespaderme F.M. (2012). “*Visión artificial industrial: procesamiento de imágenes para inspección automática y robótica*”. Valladolid: Universidad de Valladolid.
- [15] Guerrero J.M., Pajares G. y Guijarro M. (2011). “Técnicas de procesamiento de imágenes estereoscópicas”. Madrid: Universidad Complutense de Madrid.
- [16] González J. (1999). “*Visión por computador*”. España: Ediciones Paraninfo.
- [17] Hansard M., Lee S., Choi O. y Horaud R. (2012). “*Time of Flight Cameras: Principles, Methods, and Applications*”. Grenoble: Springer.
- [18] Kassam A., Carrau R.L., Snyderman C.H., Prevedello D.M., Mintz A., Gardner P. y Masegur H. (2007). “Endoscopic Expanded Endonasal Approaches to de Caudo-Ventral Skull Base”. *Acta Otorrinolaringológica Española: Cirugía endoscópica avanzada de base de cráneo y espacios paranasales*, 58(1): 14-31.
- [19] «Ubuntu 18.04 LTS (Bionic Beaver)» [En línea]. Disponible: <http://releases.ubuntu.com/bionic/>. [Accedido: 20-Feb-2019].
- [20] Larabel M. (2019) “Ubuntu 18.04 LTS With Latest GNOME Update”. *Ubuntu*.
- [21] Leong S.C.L. y White P.S. (2004). “A comparison of aesthetic proportions between the Oriental and Caucasian nose”. *Clin Otolaryngol Allied Sci*, 29(6): 672-676.
- [22] «Motorika rehabilitation robotics» [En línea]. Disponible: <http://motorika.com/>. [Accedido: 18-Jun-2019].
- [23] Nimsky C., Ganslandt O., Hastreiter P. y Fahlbusch R. (2001). “Intraoperative compensation for brain shift”. *Surgical Neurology*, 56(1): 357-364.
- [24] Niveiro M., Aranda F.I. y Peiró G. (2003). “Patología de los adenomas hipofisarios”. *Revista Española de Patología*, 36(4): 357-372.



- [25] «OpenCV» [En línea]. Disponible: <http://opencv.org/>. [Accedido: 05-Jun-2019]
- [26] «PCL - Point Cloud Documentation» [En línea]. Disponible: <http://www.pointclouds.org/documentation/>. [Accedido: 10-Jun-2019].
- [27] «Philips Imaging System» [En línea]. Disponible: <https://www.philips.es/healthcare/>. [Accedido: 20-Jun-2019].
- [28] Pons L.M., Cruz A., Álvarez I., Daudinot Y. y Salomón J. (2012). “Caracterización del uso de gadolinio en pacientes con indicación de resonancia magnética”. *MEDISAN 2012*, 16(4): 498-503.
- [29] «Proyecto CRANEEAL» [En línea]. Disponible: <https://www.uma.es/medical-robotics/info/107696/proyectedpi2016-80391-C3-1-R/>. [Accedido: 04-Jun-2019].
- [30] Pun T., Gerig G. y Ratib O. (1994). “Image Analysis and Computer Vision in Medicine”. *Computerized Medical Imaging and Graphics*, 18(2), 85-96.
- [31] Rojas D. (2017). “Manejo de los tumores de hipófisis”. *Revista Médica Clínica las Condes*, 28(3): 409-419.
- [32] Schnabel R., Wahl R. y Klein R. (2007). “Efficient RANSAC for Point-Cloud Shape Detection” *Computer Graphics Forum*, 26(2): 214–226.
- [33] Sonka M., Hlavac V. y Boyle R. (2008). “*Image Processing, Analysis, and Machine Vision* (3 ed.)”. Leibniz: Thomson.
- [34] Thevenot J., Bordallo M. y Hadid A. (2018). “A survey on Computer Vision for Assistive Medical Diagnosis from Faces”. *IEEE Journal of Biomedical and Health Informatics*, 22(5): 1497-1511.
- [35] Timor-Tritsch L.F. y Monteaguado A. (1996). “Transvaginal fetal neurosonography: standardization of the planes and sections by anatomic landmarks”. *Ultrasound in Obstetrics and Gynaecology*, 8, 42-47.
- [36] «Tratica computer vision research» [En línea]. Disponible: <https://www.tractica.com/>. [Accedido: 14-May-2019].
- [37] «Wolfson Brain Imaging Centre» [En línea]. Disponible: <https://www.wbic.cam.ac.uk/>. [Accedido: 20-Jun-2019].
- [38] Zacharopoulos G.V., Manios A., Kau C.H., Velegrakis G., Tzanakakis G.N. y Bree E. (2016). “Anthropometric analysis of the face”. *Journal of craniofacial Surgery*, 27: 71-75.



Universidad de Valladolid

# Integración de un sistema de visión 3D en un entorno robotizado de cirugía Ander Loidi Yarza



ESCUELA DE INGENIERÍAS  
INDUSTRIALES