



Universidad Autónoma de Querétaro
Facultad de Informática



Reconocimiento facial en imágenes con profundidad por medio de la
extracción de características

Tesis

Que como parte de los requisitos
para obtener el Grado de
Maestro en Ciencias de la Computación

Presenta

Rafael Duarte Pérez

Dirigido por:

Dr. Hugo Jiménez Hernández

Querétaro, Qro. a 26 de enero de 2024



Dirección General de Bibliotecas y Servicios Digitales
de Información



Reconocimiento facial en imágenes con profundidad
por medio de la extracción de características

por

Rafael Duarte Pérez

se distribuye bajo una [Licencia Creative Commons
Atribución-NoComercial-SinDerivadas 4.0 Internacional](#).

Clave RI: IFMAC-278778



Universidad Autónoma de Querétaro

Facultad de Informática

Maestría en Ciencias Computacionales

Reconocimiento facial en imágenes con profundidad por medio de la extracción de características

Tesis

Que como parte de los requisitos para obtener el Grado
Maestro en Ciencias Computacionales

Presenta

I.S. Rafael Duarte Pérez

Dirigido por:

Dr. Hugo Jiménez Hernández

Dr. Hugo Jiménez Hernández
Presidente

Dr. José Alfredo Acuña García
Secretario

M.I.S.D. Erika del Río Magaña
Vocal

Dra. Sandra Luz Canchola Magdaleno
Suplente

Dra. Ana Marcela Herrera Navarro
Suplente

Centro Universitario, Querétaro, Qro.
Enero 2024
México

Quiero expresar mi profundo agradecimiento a aquellas instituciones y personas que han contribuido de manera significativa al desarrollo de esta investigación.

En primer lugar, agradezco al Consejo Nacional de Humanidades, Ciencias y Tecnologías (CONAHCYT) por su apoyo financiero a través de la beca del plan académico MCC durante el periodo de realización de este trabajo. Su respaldo ha sido fundamental para llevar a cabo esta investigación y para mi formación académica.

Asimismo, deseo manifestar mi reconocimiento a la Facultad de Informática por brindarme las herramientas, recursos y el ambiente propicio para la realización de este proyecto. Agradezco especialmente a la Dra. Gabriela Xicoténcatl Ramírez, a la Dra. Ana Marcela Herrera Navarro y al Dr. Hugo Jiménez Hernández por su orientación, valiosos consejos y su constante apoyo a lo largo de este camino académico.

También, mi gratitud se extiende a mis compañeros de estudio, colegas, amigos y familia, cuyo apoyo, comprensión y ánimo fueron fundamentales en cada etapa de esta investigación. Hago también mención de la M.S.C. Gabriela Pacheco Sánchez, quien fue un apoyo importante en el proceso.

Sin la colaboración, guía y respaldo de estas instituciones y personas mencionadas, este trabajo no habría sido posible.

Muchas gracias.

ÍNDICE

I. Introducción	1
I.1 Planteamiento del Problema.....	2
I.2 Preguntas de investigación	4
I.3 Hipótesis	4
I.4 Objetivos	5
I.5 Justificación	5
II. Antecedentes y/o Fundamentación Teórica	7
III. Metodología de Investigación	21
3.1 Análisis de la situación.	23
3.2 Desarrollo de las soluciones (Propuesta).....	25
3.2.1 Banco de imágenes.....	25
3.2.2 Extracción de características	29
3.2.3 Generación del grafo de identidad.....	29
3.3 Implementación.....	29
3.3.1 Banco de imágenes.....	30
3.3.2 Extracción de Características.....	38
3.4 Producción de documentación	39
IV. Resultados	40
V. Conclusiones	46
VI. Referencias	48

ÍNDICE DE TABLAS

Tabla 3.1 Características de la CCPR.....	28
Tabla 3.2 Iteraciones del sistema SCRCP	36

ÍNDICE DE FIGURAS

Figura 2.1.....	10
Figura 2.2.....	11
Figura 2.3.....	13
Figura 2.4.....	14
Figura 2.5.....	15
Figura 2.6.....	16
Figura 2.7.....	18
Figura 2.8.....	21
Figura 3.1.....	22
Figura 4.1.....	40
Figura 4.2.....	41
Figura 4.3.....	42
Figura 4.4.....	43
Figura 4.5.....	44
Figura 4.6.....	45

ABREVIATURAS Y SIGLAS

BMP - Imagen de Mapa de Bits

CCPR – Colección a Color y a Profundidad de Rostros

FAST – Features from Accelerated Segment Test

INDAUTOR – Instituto Nacional del Derecho de Autor

JPG, JPEG - Joint Photographic Experts Group

OBJ - Wavefront OBJ

ORB - Oriented FAST and Rotated BRIEF

ORL - Olivetti Research Laboratory

PLY - Polygon File Format

RGB – Red Green Blue

SCRCP – Sistema Capturador de Rostros a Color y a Profundidad

SIFT - Scale-Invariant Feature Transform

SURF - Speeded Up Robust Features

ToF – Tiempo de Vuelo (Time of Flight)

Resumen

En esta investigación, se aborda detalladamente la problemática asociada al reconocimiento facial en imágenes convencionales, destacando la carencia de información sobre la topología facial. Esta limitación surge debido a que, en la mayoría de los casos, el análisis se limita al uso exclusivo del color como factor principal, lo cual resulta insuficiente para lograr una representación completa y precisa de las características faciales. Para superar esta deficiencia, se optó por la utilización de imágenes de profundidad capturadas mediante una cámara ToF (Time of Flight). Este enfoque se orienta a extraer características basadas en las distancias entre los objetos capturados, proporcionando así información tridimensional detallada. El objetivo fue la creación de un grafo que almacene información crucial para un reconocimiento facial efectivo.

La inclusión de datos de profundidad permite extraer características que enriquecen el proceso de reconocimiento facial. Se procede a la construcción de un grafo donde cada nodo representa puntos clave en la topología facial, estableciendo relaciones espaciales a partir de las distancias exactas entre ellos. Este grafo resultante permitió almacenar información de las características obtenidas de los diversos rostros, logrando así, la identificación facial por medio del isomorfismo que se usó para validar el grafo creado con el grafo obtenido en la lectura del rostro.

Palabras clave: Grafo identidad, Harris-Corner, imágenes a profundidad, reconocimiento facial

Abstract

In this research, the issue associated with facial recognition in conventional images is thoroughly addressed, emphasizing the lack of information regarding facial topology. This limitation arises because, in most cases, the analysis is exclusively focused on the use of color as the primary factor, which proves insufficient to achieve a complete and precise representation of facial features. To overcome this deficiency, the utilization of depth images captured through a Time of Flight (ToF) camera was chosen. This approach aims to extract features based on distances between captured objects, providing detailed three-dimensional information. The objective was the creation of a graph that stores crucial information for effective facial recognition.

The inclusion of depth data allows for the extraction of features that enhance the facial recognition process. The construction of a graph follows, where each node represents key points in facial topology, establishing spatial relationships based on exact distances between them. This resulting graph facilitated the storage of information from features obtained from various faces, thereby achieving facial identification through isomorphism, which was used to validate the graph created with the graph obtained in face reading.

Keywords: Depth images, facial recognition, Harris-Corner, identity graph.

I. Introducción

Los sistemas biométricos son aquellos que utilizan características físicas de los humanos para algún tipo de reconocimiento, este trabajo se centra en la biometría facial que involucra el análisis de rostros y sus características para poder realizar un reconocimiento de este.

La identificación de rostros obtenida a través de la visión por computadora ha sido muy estudiada y trabajada a lo largo de los últimos años debido a sus múltiples aplicaciones, específicamente para sus implementaciones en sistemas de seguridad (Madrigal et al., 2007).

La biometría facial tiene aplicaciones que vemos desde las redes sociales y usos recreativos hasta en temas de seguridad como la autenticación de sistemas bancarios. Por lo que existen múltiples modelos y metodologías para lograr la biometría facial, dichos modelos se exponen a continuación en el presente trabajo.

En mayor medida se ha centrado la biometría facial en elementos bidimensionales como son las fotografías e imágenes convencionales, las cuales están conformadas por una matriz 2D de píxeles. Sin embargo, este formato puede verse afectado por factores externos como lo son la iluminación, la escala, la perspectiva y la posible superposición de otra imagen 2D (Vázquez-López, 2014).

En contraste con lo anterior, los sistemas biométricos en imágenes 3D perciben información adicional, como la profundidad, que ayuda a evitar problemas con la perspectiva, iluminación

e incluso expresiones faciales. Sumando el hecho de que falsificar o sobre poner otra imagen resulta muy complicado, el reconocimiento facial 3D presenta mayores virtudes para su uso en la biometría facial. Sin embargo, es importante aclarar que este proceso conlleva mayor exigencia de procesamiento.

I.1 Planteamiento del Problema

La biometría requiere de la información de la geometría facial para identificar las características que hacen distintivo al rostro. En este proceso se han desarrollado técnicas directas e indirectas de medir dicha geometría (Adjabi et al., 2020).

En el desarrollo de sensores ópticos, las cámaras de espectro visible han trasladado el problema de detectar asimetrías en la biometría a reconocer contrastes debido a la intensidad, que posteriormente se asocian a la geometría del rostro. Sin embargo, se ha mostrado que el uso de intensidades puede estar propenso a ser afectado por las condiciones de iluminación y la perspectiva de la luz que incide en el objeto generando falsas geometrías o evitando su detección (Chihaoui et al., 2016).

En tal sentido, un área de oportunidad es el uso de una instrumentación que no se vea afectada por las variaciones de luz y tenga como objetivo detectar la geometría del rostro utilizando valores de profundidad (Hansard et al., 2021).

Esta situación da lugar a utilizar sensores de tiempo de vuelo en combinación de descriptores de forma asociadas a las imágenes de distancia que, al trabajar con la posición en el plano (x, y) de cada píxel y agregar el valor de profundidad (z) , proporcionan un entendimiento geométrico de los rostros que permite tener una mayor cantidad de información para trabajar el modelo de reconocimiento facial en comparación de los sensores que no contemplan la profundidad (Adjabi et al., 2020).

Además, en la fundamentación teórica se expone la existencia de un bajo número de bases de datos para imágenes de rostros con profundidad en comparación al amplio número de bases de datos para imágenes 2D. Además de la falta de actualización de las primeras. Por lo cual en el presente trabajo se expone la intención de crear una base de datos de imágenes de rostros con profundidad.

Según Soltanpour et al., (2017), existe un número escaso de modelos de reconocimiento facial que utilicen la extracción de características de la imagen, lo que representa un área que se puede abordar más, además de hacer notar que los sensores de tiempo de vuelo son más accesibles y baratos en la actualidad en comparación a los años de algunas citas.

Por último, el reconocimiento facial convencional sólo toma en cuenta factores de la imagen en 2D por lo que información como la profundidad o la distancia entre regiones faciales es ignorada, el reconocimiento facial 3D busca incorporar esa información para realizar un modelo robusto de identificación facial (Ouamane et al., 2014).

I.2 Preguntas de investigación

Pregunta general: ¿Es posible elaborar una métrica basada en la detección de características globales en imágenes de profundidad que permitan desarrollar criterios de similitud para la identificación de rostros?

Preguntas específicas:

- ¿Cómo se interpretará la información de la imagen para su preprocesamiento?
- ¿Cómo resulta útil utilizar Harris Corner?
- ¿Qué límites conlleva la adquisición de imágenes con profundidad?
- ¿Qué información nos resulta útil al utilizar imágenes con profundidad en comparación a las imágenes convencionales 2D?
- ¿Se comprobará el modelo desarrollado con otra base de datos? ¿Se obtiene el comportamiento deseado?

I.3 Hipótesis

Si es posible generar un grafo de identificación para el reconocimiento facial de individuos a partir de la extracción de características de imágenes con profundidad capturadas con

una cámara con sensor infrarrojo, entonces se puede usar el grafo de identificación para validar la identidad de los rostros analizados.

I.4 Objetivos

Objetivo general: Desarrollar un modelo de reconocimiento facial para obtener una descripción geométrica del rostro en imágenes de distancia mediante la adaptación de descriptores globales.

Objetivos Específicos:

- Construir una base de datos de rostros con profundidad con alumnos de la Facultad de Informática de la Universidad Autónoma de Querétaro para su uso en este y futuros proyectos.
- Interpretar las imágenes capturadas dentro del modelo para la adquisición de características utilizando el descriptor Harris-Corner.
- Generar un grafo de identificación del rostro mediante las características extraídas para validar la identidad del rostro y lograr el reconocimiento facial.

I.5 Justificación

El análisis biométrico entendido como la búsqueda de información distintiva del rostro es un problema abierto porque los rostros pueden tener una gama amplia de características en

cuestión de forma, tamaño, color de piel, expresiones, imperfecciones, marcas y poses. También porque los mismos rostros pueden cambiar debido al tiempo, tener un grado de oclusión al usar lentes o maquillaje. Las condiciones de captura también influyen en el problema, así como la representación y aprendizaje de los rostros.

Una posible solución que se ha utilizado es el uso del espectro visible y el contraste de color para identificar zonas topológicas para identificar rostros, sin embargo, según Ouamane et al., (2014), el reconocimiento facial trabajado solamente con imágenes 2D no brinda la precisión deseada debido a la sensibilidad de iluminación, pose y variación de expresiones mientras que el uso de imágenes 3D puede evitar en gran medida esas limitantes.

Cada vez surgen más métodos que toman en cuenta el uso de imágenes 3D para el reconocimiento facial y como señalan Soltanpour et al., (2017), hacen falta modelos de reconocimiento facial que tomen en cuenta la extracción de características de imágenes con profundidad.

Las bases de datos que contienen imágenes de rostros con profundidad están desactualizadas además de ser pocas en comparación a las bases de datos de imágenes en 2D. En este proyecto proponemos la necesidad de tener una base de datos propia para generar el modelo de reconocimiento facial, y promover la base de datos construida para su uso en otros proyectos de investigación.

Se podría lograr una identificación de rostros más confiable si se obtiene una mayor cantidad de información que ayude a describir la estructura geométrica del rostro y por esta razón el hardware basado en la medición de distancia juega un papel indispensable en el proyecto.

II. Antecedentes y/o Fundamentación Teórica

Los sistemas biométricos utilizan características físicas del cuerpo humano con la finalidad de identificar a diferentes personas (Ruíz et al., 2009). Este proyecto se centra en el análisis biométrico facial utilizando imágenes que censan la profundidad con respecto a la cámara.

La identificación de rostros que se obtiene a través de la visión por computadora se ha estudiado y trabajado a lo largo de los años dando pie a múltiples aplicaciones en el entorno de los sistemas de seguridad (Madrigal-González et al., 2007).

La biometría facial ha sido bien recibida como proceso de autenticación, gracias a que no requiere ninguna herramienta adicional, como sí lo hacen otros procesos de autenticación como las tarjetas, además, no requiere memorizar algún patrón o contraseña, los cuales corren el riesgo de ser olvidados o extraviados por el usuario (Choi & Lee, 2009).

La biometría facial se realiza a través de distintas técnicas de reconocimiento facial donde los patrones asociados a la forma y fisiología del rostro son analizados para identificar un rostro en un video o imagen (Tamilselvi & Karthikeyan, 2018).

En la literatura se observa que las imágenes y videos utilizados para el reconocimiento facial se enfocan en analizar las intensidades de color sobre el plano imagen en el espectro visible, lo que presenta limitantes importantes como lo son los cambios en la intensidad de luz, cambios en las expresiones faciales o la presencia del uso de cosméticos. El uso de imágenes de profundidad tiene el objetivo de proveer información sobre la topología del rostro para lograr un reconocimiento facial menos sensible a las limitaciones previamente señaladas (Adjabi et al., 2020).

Tanto para imágenes de espectro visible como para imágenes de profundidad, el reconocimiento facial usa filtros generados por computadora que transforman los rostros en valores numéricos para ser comparados y determinar su similitud. Según Mahouachi y Akhloufi (2023) el proceso de reconocimiento facial consiste en los siguientes pasos:

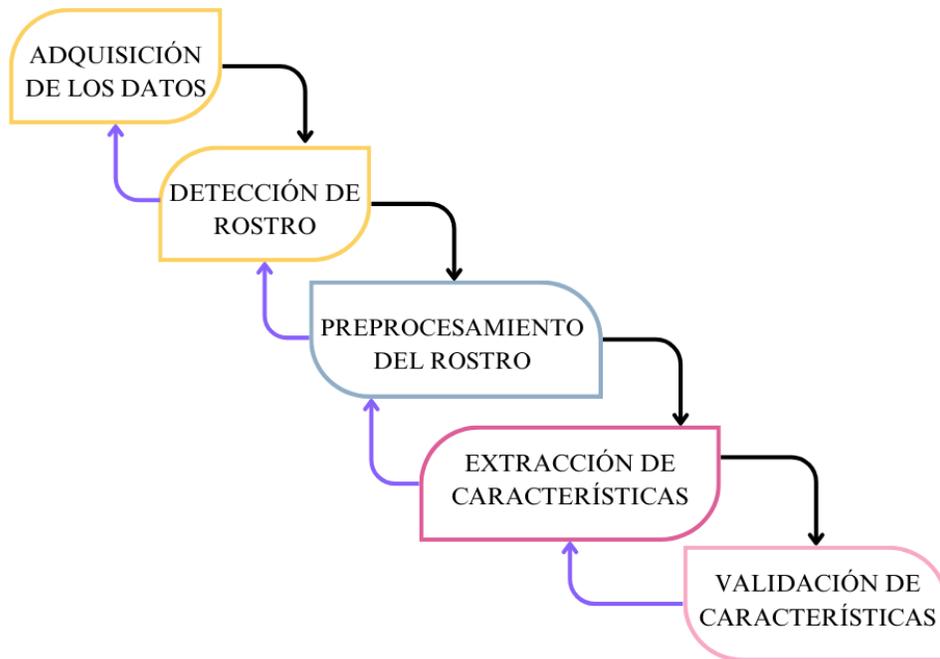
1. Detección de rostro, fase que se enfoca en el escaneo completo de la imagen con la finalidad de identificar si es que existe un área que corresponde a un rostro. En caso de identificar un rostro, pasamos a la siguiente fase.

2. Preprocesamiento del rostro, fase que trabaja sobre el área de interés en la que se realizan operaciones para reducción de ruido, mejora del contraste o resaltar detalles.
3. Extracción de características, consiste en identificar los cambios de valores en la imagen para mostrar las áreas donde cambian las distancias en la imagen. Esto ayuda a definir las regiones faciales asociadas a los ojos, nariz, boca, mejillas, etc., así como identificar la relación geométrica entre dichas regiones.
4. Validación de características, fase que consiste en la comparación de las características encontradas en el rostro de la imagen y las características de los rostros ya conocidos o registrados.

En la Figura 2.1 se expone un diagrama que contiene el proceso de reconocimiento facial para el presente trabajo.

Figura 2.1

Etapas para el reconocimiento facial

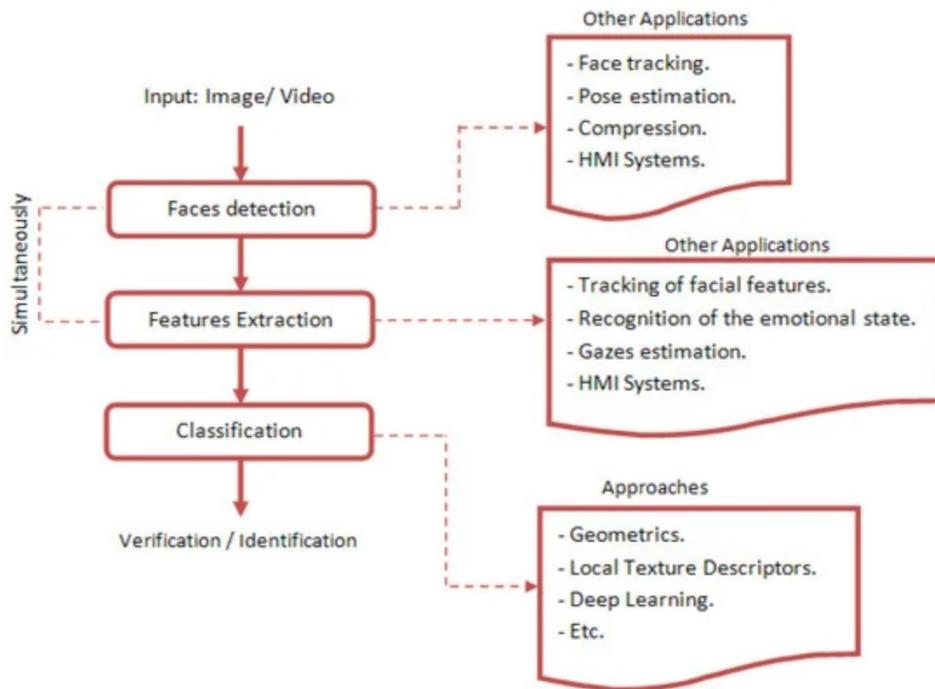


Nota: Elaboración propia a partir de la revisión de la literatura.

A lo largo de la revisión de la literatura se encontraron fases similares, como las que señalan Adjabi et al., 2020, en la Figura 2.2.

Figura 2.2

Fases del Reconocimiento



Nota: Tomado de Past, Present, and Future of Face Recognition: A Review. (Adjabi et al., 2020)

Los modelos antes mencionados utilizan un modelo sobre la topología del rostro, esto permite adaptar un conjunto de características preexistentes para la detección del rostro. En este sentido, este proyecto aborda el problema desde una perspectiva más general, donde es necesario conocer todas las zonas relevantes de un rostro para lograr construir una métrica que ayude a su detección e identificación. En este trabajo va encaminado a la detección automática de características biométricas, utilizando descriptores de imágenes,

los cuales ayudaran a la identificación eficiente de rostros. Por esta razón a continuación mostramos un conjunto de DataSets de Rostros que se pueden utilizar para su evaluación y luego un conjunto de descriptores en espectro visible, que pueden adaptarse a imágenes de profundidad para proporcionar información sobre la topología del rostro.

Durante la fase de detección de rostro es crítico que se tenga una base de información de rostros que sirva para probar, desarrollar y comprobar el reconocimiento en esta y las demás fases (Kortli et al., 2020). Para trabajar con imágenes 2D existe una gran variedad de bases de datos como las siguientes:

- Olivetti Research Laboratory (ORL): Base de datos que contiene 400 imágenes de 40 personas diferentes, las imágenes fueron tomadas en tiempos diferentes, con una variación de iluminación, diferentes expresiones y detalles faciales propensos a oclusión. Las imágenes tienen un tamaño de 92 x 112 píxeles y están en escala de grises con un valor de 0-255 por píxel (Zhu et al., 2019). La Figura 2.3 es una muestra de las imágenes contenidas en ORL (He, 2021).

Figura 2.3

Muestra de las imágenes contenidas en ORL



Nota: Obtenido de Research on Face Image Digital Processing and Recognition Based on Data Dimensionality Reduction Algorithm (He, 2021).

- Institute of Automation, Chinese Academy of Sciences, Near Infrared – Visible Light (CASIA NIR-VIS): Base de datos que contiene imágenes de los rostros de 725 personas diferentes con dos tipos de formato diferentes; JPEG con un número de 1-22 imágenes por persona y BMP con un número de 5-50 imágenes por persona. Para ambos casos se utilizó una resolución de 640x480 píxeles (Stan et al., 2013). En la Figura 2.4 se muestran 6 fotos de la base de datos CASIA.

Figura 2.4

Imágenes de CASIA en formatos JPEG y BMP



Nota: Tomados de la base de datos Casia Nir-Vis 2.0 (Stan et al., 2013).

- RGB Depth and Thermal (RGB-D-T): Base de datos con 45,900 imágenes de 51 personas diferentes, orientada a hombres entre 20 - 40 años, el sensor utilizado fue un Microsoft Kinect y capturó 3 tipos de imágenes; RGB con dimensiones de 640x480 píxeles, imágenes de profundidad con dimensiones de 640x480 píxeles e imágenes térmicas con una resolución de 384x288 píxeles. En la Figura 2.5 se muestran ejemplos de las imágenes para cada formato Simón et al., (2016).

Figura 2.5

Rostros en formatos RGB, profundidad y térmico

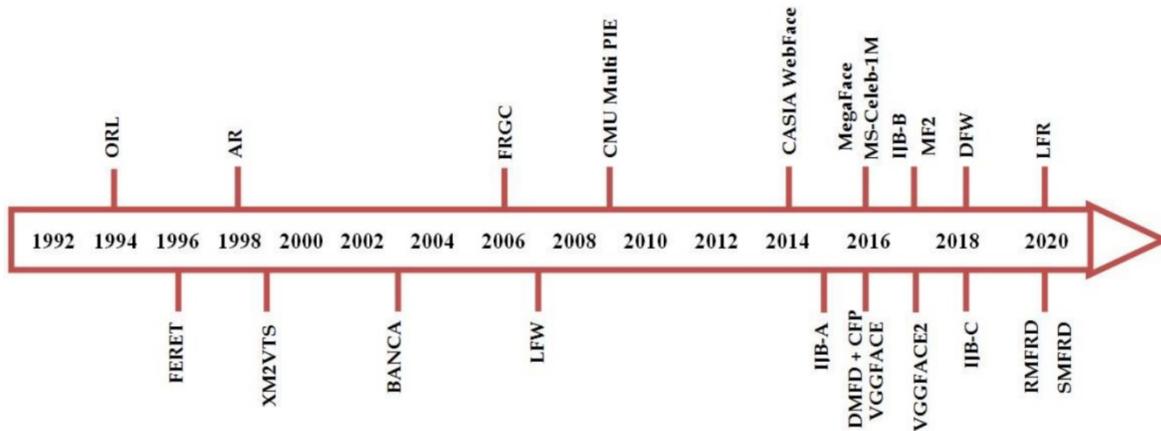


Nota: Tomado de la base de datos Improved RGB-D-T based face recognition (Simón et al., 2016).

Existen más bases de datos sobre rostros como FERET propuesta por Phillips, Wechsler, Huang y Rauss (1998), AR Dataset desarrollada por Martinez and Benavente (1998), IARPA Janus Benchmark-A recolectada por Klare et al., (2015) entre otras. Adjabi et al., (2020) muestra en la Figura 2.6 una recolección de bases de datos para el reconocimiento de rostros.

Figura 2.6

Línea de tiempo del desarrollo de bases de datos de rostros



Nota: Tomado de Past, Present, and Future of Face Recognition: A Review. (Adjabi et al., 2020).

Las distintas bases de datos sobre rostros varían en cuanto al sensor que utilizan y por lo tanto cambian sus valores en las dimensiones, formatos y poses que capturan. Chihaoui et al., (2016) muestran las variaciones que diferencian a cada base de datos a continuación:

- Número de imágenes contenidas en cada base de datos.
- Número de imágenes por individuo, adquiridas con diferentes condiciones como la orientación, pose, iluminación y presencia de oclusión.
- La dimensión de las imágenes.
- Pose y orientación de los rostros.

- Cambios en la iluminación.
- Género de las personas que participan en la adquisición de rostros.
- Fondo uniforme de la imagen.
- Tiempo entre capturas de rostro.

En cuanto al reconocimiento de rostros en imágenes 3D, Adjabi et al., (2020) menciona que la información adquirida por las cámaras con sensor infrarrojo tiene ventajas comparadas con la información capturada con las cámaras comunes, una ventaja clara es la captura de imágenes bajo condiciones de poca iluminación y tienen una mayor robustez al detectar cambios en la expresión facial.

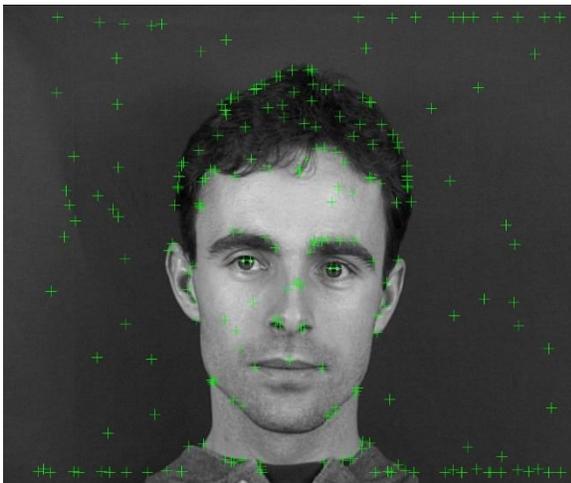
A lo largo de la revisión de la literatura se encontró un bajo número de bases de datos con imágenes de rostros con profundidad, además se encontró que dichas bases de datos están desactualizadas debido a que la más reciente tiene fecha del 2010, dicha base de datos es Texas 3D Face Recognition Database (Texas 3DFRD) que cuenta con 1140 imágenes de 118 personas, es de libre acceso y fue desarrollada por Gupta, Castleman, Markey, and Bovik (2010).

En cuanto a la extracción de características se encontraron los siguientes en la revisión de la literatura:

- Harris Corner Detector – El detector desarrollado por Harris y Stephen (1998) extrae esquinas e infiere características de la imagen. El detector toma las diferenciales del marcador de esquinas en una cuenta con referencia a la dirección directa, ha mostrado ser apto para distinguir entre bordes y esquinas. De acuerdo con el estudio de Hamouz (2014), el detector de Harris es eficiente para sacar las siguientes características de un rostro: ojos, fosas nasales, las corneas internas y la boca, como se muestra en la Figura 2.7.

Figura 2.7

Respuesta Típica del detector Harris Corner en un rostro



Nota: Tomado de Feature-based affine-invariant detection and localization of faces (Hamouz 2014).

- Scale-Invariant Feature Transform (SIFT): detector y descriptor de puntos característicos. SIFT detecta los puntos en las imágenes mediante la función DoG (Diferencia de gaussianas). SIFT se desempeña principalmente en tareas de reconocimiento de objetos. De acuerdo con el trabajo de Krizaj et al., (2010) este descriptor funciona muy bien para analizar los puntos de interés de la imagen para la detección de rostros. Sin embargo, presenta un problema al momento de generar los puntos de interés, ya que resulta variante a la iluminación lo que hace que sea un método poco confiable en entornos no controlados. En el trabajo de Krizaj et al., (2010) Utilizan el SIFT con puntos aprendidos por etapas de aprendizaje para evitar el problema de la iluminación.

Se usó por Zhu, Yi, Lei y Li (2014) para generar un nuevo modelo morfológico 3D que encontrará correspondencias entre puntos 3D y puntos 2D

- Speeded Up Robust Features (SURF): Desarrollado por Bay et al., (2008), parte del modelo SIFT. SURF es un detector y descriptor de puntos característicos. La mejora de este modelo es la velocidad de generación de los descriptores, ya que reduce la dimensión de los vectores y la complejidad del cálculo.

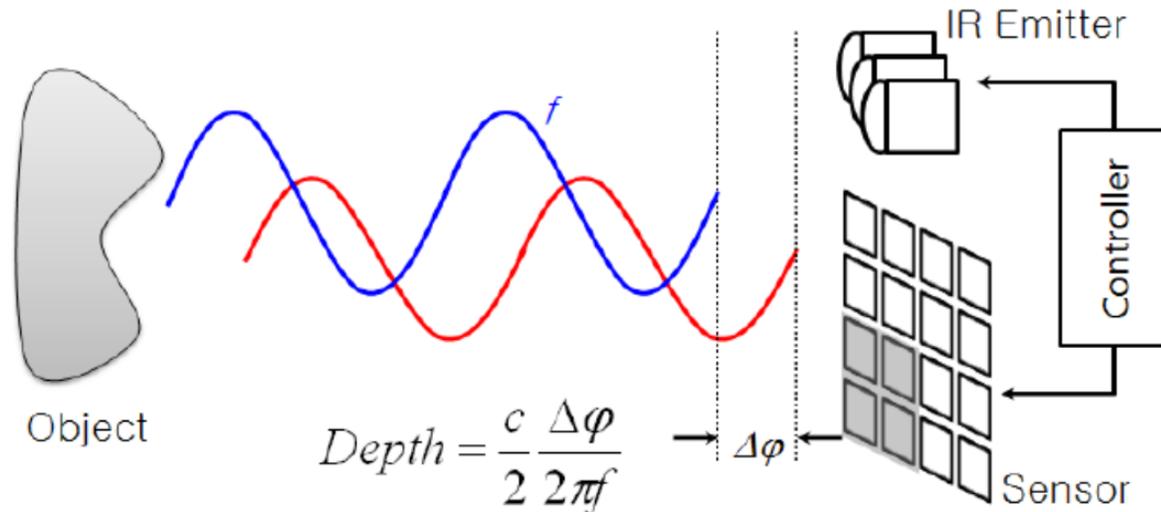
- Oriented FAST and Rotated BRIEF (ORB): Desarrollado por Rublee et al., (2011) en los laboratorios de OpenCV. Naciendo de la necesidad de un algoritmo libre, ORB toma características de SURF y SIFT. Logra ser tan preciso para detectar las características como SIFT mientras que es más rápido.

ORB utiliza el método FAST para la extracción de características e incorpora el descriptor Binary Robust Independent Elementary Feature (BRIEF), el cual tomará todas las características y las convertirá en vectores característicos binarios. Este descriptor es útil para representar una característica a través de los valores de sus vecinos en un formato de vector binario, sin embargo, no es invariante a la rotación de la imagen por lo que ORB utiliza rBRIEF (rotation-aware BRIEF). rBRIEF siendo una variante de BRIEF que busca solventar el problema de la rotación.

En esta propuesta se utilizarán cámaras de profundidad, por lo cual se expone a manera de referencia su funcionamiento para entender qué información es muestreada al momento de utilizar este tipo de cámaras. La cámara de tiempo de vuelo (ToF) la cual produce una imagen con profundidad donde se registra, por cada píxel, la distancia del objeto capturado en contraste a la posición de la cámara. Estas cámaras funcionan al medir el retraso de la reflexión de la luz infrarroja. En la Figura 8 se muestra el principio de las cámaras de tiempo de vuelo. (Hansard et al., 2021)

Figura 2.8

Principio del funcionamiento de las cámaras de tiempo de vuelo



Nota: Tomado de Time of flight cameras: Principles, methods, and applications (Hansard et al., 2021).

III. Metodología de Investigación

La metodología aplicada en esta investigación es basada en diseño, esto debido a que sus raíces provienen de las ciencias del diseño como, por ejemplo: ingeniería, arquitectura, generación de productos y desarrollo de software, donde esta última en especial resulta atractiva para el proyecto por la propia naturaleza de este. La metodología basada en diseño tiene la característica de desarrollar productos por medio del proceso iterativo, ayudando a entender mejor el cómo y el porqué del funcionamiento del producto. (Bowler & Large, 2008).

Las fases de la metodología basada en el diseño comprenden ciclos para la definición del problema, el diseño, desarrollo, implementación y la evaluación. Estas fases se repiten en sí mismas y son parte de un proceso que puede reiterarse con la finalidad de obtener aprendizaje y cambios que acerquen el diseño del proyecto a lo estipulado en los resultados esperados. (Strobel et al., 2008).

En la Figura 3.1 se muestran las etapas propuestas para este proyecto, así como los trabajos planteados para cada etapa.

Figura 3.1

Actividades propuestas para las fases generales de la investigación basada en diseño en el presente proyecto



Nota: Elaboración propia.

3.1 Análisis de la situación.

Se realizó una investigación documental sobre el reconocimiento facial utilizando el procesamiento de imágenes, los formatos de las imágenes utilizadas así como la dimensionalidad de las mismas (como matrices de píxeles o como imágenes que describen la profundidad de la escena), los problemas que se presentan en el reconocimiento facial como lo son la perspectiva y la iluminación, los materiales usados para capturar las imágenes, el preprocesamiento que se realiza a una imagen así como los extractores de características utilizados para encontrar puntos de interés en las imágenes, entre otros temas de interés del proyecto. Esto permitió una contextualización del panorama actual del reconocimiento facial.

Especialmente se investigó sobre bancos de imágenes usados en otros proyectos ya que en otras investigaciones se usaron bancos propios, de paga o abiertos. Por lo anterior, y ya que un objetivo de este proyecto fue crear un banco de imágenes para desarrollar el grafo de identidad, se comenzó a diseñar un banco de imágenes tomado con una cámara ToF que permitiera obtener la imagen a color de la escena, así como la profundidad de esta. Los principales puntos de atención fueron los siguientes: formato de la captura (manejada principalmente como bmp o jpg para la imagen en 2 dimensiones), el número de capturas por individuo, las perspectivas del rostro capturadas en el banco de imágenes, así como las diferentes expresiones del rostro.

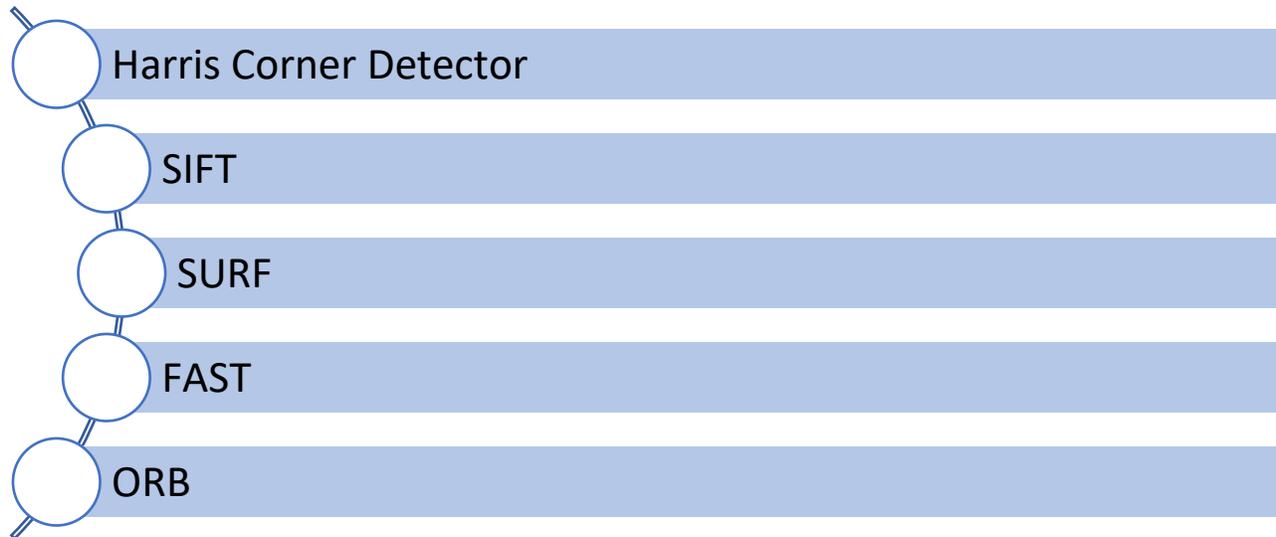
Un punto importante que definir para el banco de imágenes fue el formato que tendría la imagen a profundidad ya que la cámara ToF utilizaba un visualizador de fábrica y era difícil poder interpretar y trabajar la imagen fuera de ese software. Tras una breve investigación de la herramienta se encontró una forma de trabajar directamente con el lenguaje Python para poder manipular la información del sensor y poderla almacenar.

Para realizar la toma de capturas de rostros se buscó un pequeño estudio de fotografía que consiste en un fondo (pantalla verde, blanca y negra), dos focos de estudio con paraguas, un tripié y una silla. La iluminación no es un problema para este proyecto de investigación ya que, como se mostrará a continuación, la luz no afecta a la información capturada por una cámara ToF, sin embargo, utilizamos focos de estudio con paraguas para obtener una imagen de dos dimensiones con luz uniforme para ser utilizada en algún otro proyecto de la universidad.

Se realizó una investigación con la finalidad de identificar y comparar a los procesos de reconocimiento facial basados en la extracción de características, en la figura 3.2 se muestran los principales extractores de características involucrados en el reconocimiento facial:

Figura 3.2

Principales extractores de características en el reconocimiento facial



Nota: Elaboración propia.

3.2 Desarrollo de las soluciones (Propuesta)

Gracias a la información que se recuperó en la revisión de la literatura se diseñaron las siguientes estrategias para poder cumplir con los objetivos del presente proyecto.

3.2.1 Banco de imágenes.

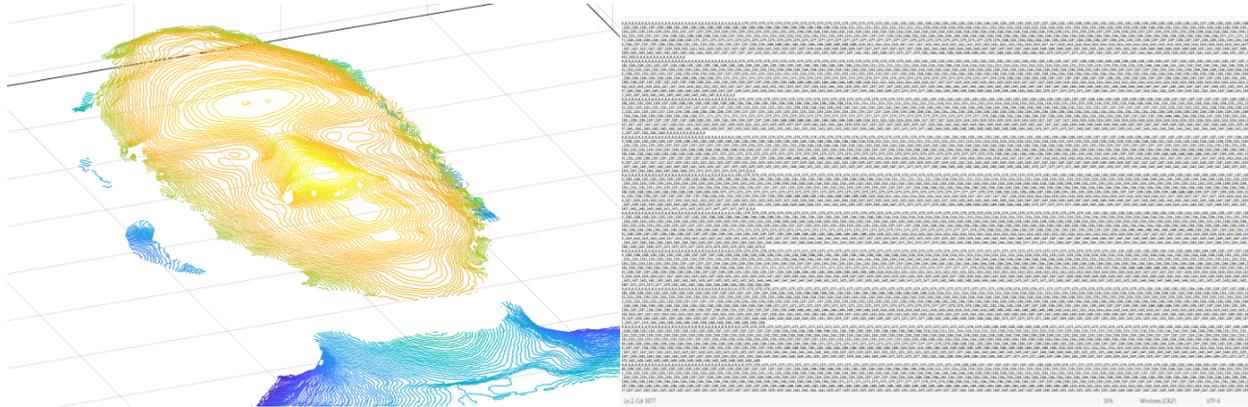
Como se mencionó anteriormente, identificar el formato en el cual se guardarán los rostros capturados en imágenes a profundidad generaba un conflicto debido a que la herramienta complicaba la captura de dicha información y se limitaba a ser un visualizador. Existen distintos formatos que permiten trabajar con una nube de puntos como por ejemplo el

formato PLY (Polygon File Format) u OBJ (Wavefront OBJ), sin embargo, no se eligieron ya que al ser usados se notaron limitaciones como la accesibilidad al archivo (ya que se requerían librerías adicionales) o se tenía un incremento en tamaño del archivo, el cual aumenta los tiempos de captura.

Para dar solución a lo anterior, se propuso trabajar con la imagen como un archivo de texto, donde se guarda una imagen como una matriz de píxeles, donde cada píxel no representa un color, sino una distancia con respecto a la cámara medida en milímetros. Cada píxel resulta estar separado por comas y cada fila de la matriz está separada por saltos de línea. Esto ayuda al rendimiento del programa ya que mejora mucho la velocidad de guardado de la imagen a profundidad. En la Figura 3.3 se muestra una comparación de la imagen vista a través de una gráfica y esa misma imagen en su formato de archivo de texto.

Figura 3.3

Comparativa del rostro a profundidad, visualización 3D (izquierda) y archivo de texto (derecha)



Nota: Elaboración propia.

Tener la imagen de profundidad en un archivo de texto permite una gran compatibilidad con las herramientas usadas ya que es directamente interpretable como una matriz de valores que representan distancias. En la Figura 3.3 se puede ver que se recrea exitosamente el rostro a partir del archivo de texto.

Algunos bancos de imágenes expuestos anteriormente cuentan con un número determinado de imágenes, la base de datos ORL tiene 400 imágenes de 40 personas diferentes, la CASIA NIR-VIS contiene un número variante de imágenes de 725 individuos y la RGB-D-T contiene aproximadamente 45 mil imágenes de 51 personas diferentes. La base de datos creada para este proyecto obtuvo 486 imágenes por individuo capturado y se capturó un total de 50 individuos, dando un total de 24,300 imágenes.

El banco de imágenes generado en este proyecto fue denominado CCPR (Colección a Color y a Profundidad de Rostros) y las imágenes que fueron capturadas para este proyecto fueron obtenidas bajo una cesión de derechos por parte de los individuos involucrados donde se brinda el consentimiento informado, se detalla la privacidad y confidencialidad de las imágenes. Esto siguiendo las recomendaciones de bioética en el proceso de investigación mencionadas por Outomuro & Mirabile (2015).

Tabla 3.1

Características de la CCPR

Características	Valores
Tipo de imágenes	<i>Imágenes a RGB (mitad de las muestras)</i> <i>Imágenes a profundidad (mitad de las muestras)</i>
Resolución de las imágenes	<i>Imágenes a RGB – 640x480 píxeles</i> <i>Imágenes a profundidad - 640x480 píxeles</i>
Formato de archivos	<i>Imágenes a RGB – JPG</i> <i>Imágenes a profundidad – Archivos de texto</i>
Peso aproximado de archivo individual	<i>Imágenes a RGB – 300 KB aproximadamente.</i> <i>Imágenes a profundidad – 1,500 KB aproximadamente.</i>

Nota: Elaboración propia.

3.2.2 Extracción de características

A las imágenes a tratar se les realizará un preprocesamiento para extraer las características centradas en el rostro, por lo que se quitarán los valores de distancia con el fondo de la imagen, así mismo se hará lo posible por minimizar la presencia del cuello y pecho de los individuos.

Para poder generar el grafo de identidad se extraerán las características de la imagen a profundidad con el extractor de características Harris-Corner, este método ayuda a mostrar los cambios abruptos en color, por lo que, al ser aplicado en nuestra imagen con valores de distancia, nos brindará características relativas a las dimensiones topológicas del rostro.

3.2.3 Generación del grafo de identidad

Este proyecto plantea que, a través de la obtención de características representativas de un rostro, se pueden generar un grafo de identidad que se asocie a un rostro particular y gracias a esto poder lograr un reconocimiento facial usando una imagen de profundidad.

Se propone que las características sean representativas de áreas faciales que permitan determinar las particularidades de los rostros y gracias a esto se pueda diferenciar entre los rostros de la base de datos.

3.3 Implementación

3.3.1 Banco de imágenes

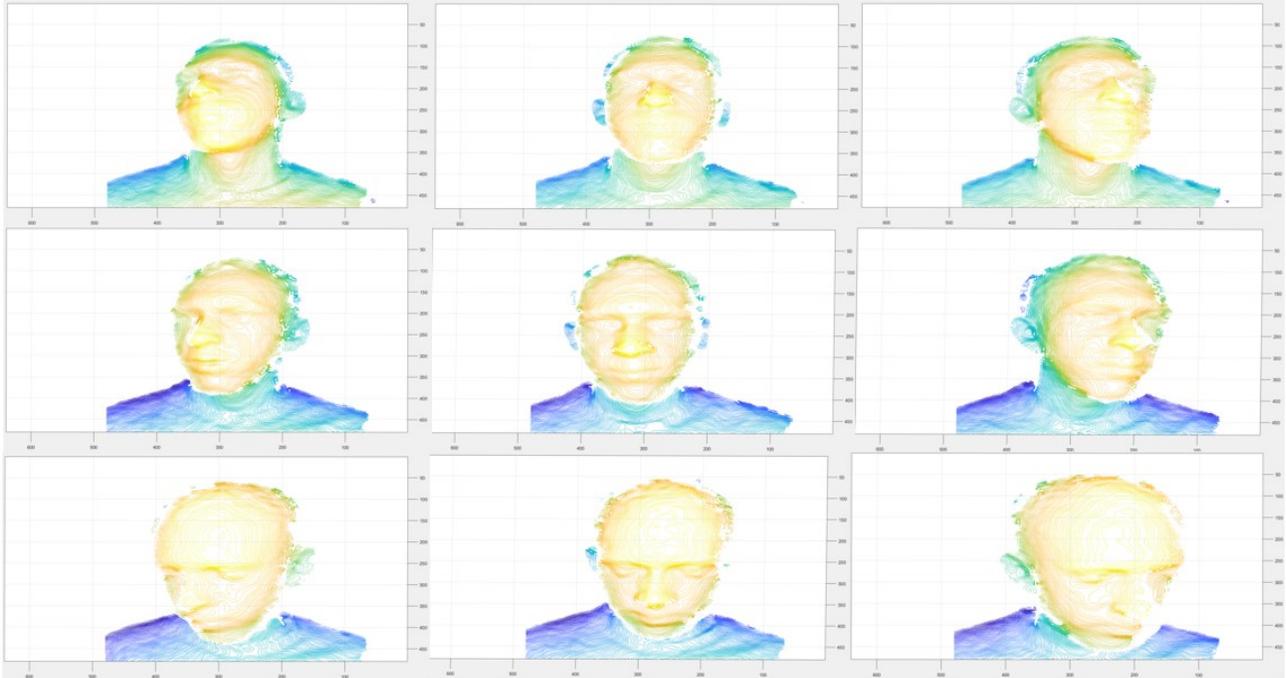
El banco de imágenes fue implementado pensando en su uso en este proyecto y cualquier otro que requiera imágenes RGB o de profundidad, por lo que se capturaron las imágenes bajo las siguientes normas:

- La distancia del individuo con respecto al sensor queda entre los 40 cm – 68 cm.
- La altura de la cámara queda a 127 cm a distancia del suelo
- La altura de la base de la silla queda a 98 cm a distancia del suelo
- El fondo verde queda a una distancia de 107 cm con respecto a la cámara
- Las imágenes, en ambos formatos, cuentan con una resolución 640 x 480 píxeles.
- Se capturaron 9 perspectivas diferentes:
 - Frontal, frontal con inclinación a la derecha y frontal con inclinación a la izquierda
 - Frontal superior, inclinación a la superior derecha, inclinación a la superior izquierda.
 - Frontal inferior, inclinación a la inferior derecha, inclinación a la inferior izquierda.
- Para cada imagen de profundidad se toma también una imagen RGB simultánea.

Las 9 perspectivas capturadas tienen como propósito el mostrarnos la topología completa del rostro para tomar esas características en cuenta. En la Figura 3.4 se muestra un ejemplo de las perspectivas obtenidas en el banco de imágenes.

Figura 3.4

Perspectivas de los rostros capturaos

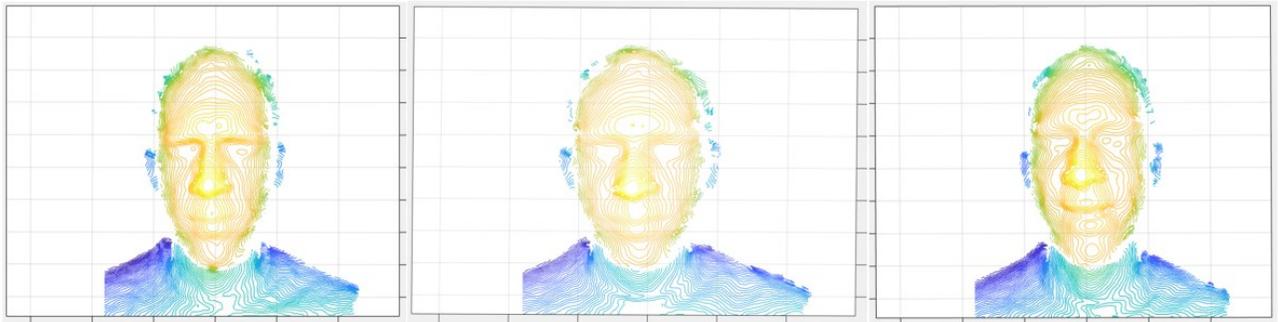


Nota: Elaboración Propia.

Aparte de las perspectivas, el banco de imágenes captura diferentes expresiones de los individuos, los individuos son capturados con una expresión neutral o seria, con una expresión sonriente y por último, una expresión con el ceño fruncido como se muestra en la Figura 3.5.

Figura 3.5

Expresiones capturadas en el banco de imágenes. (Enojado, serio y sonriente de izquierda a derecha)

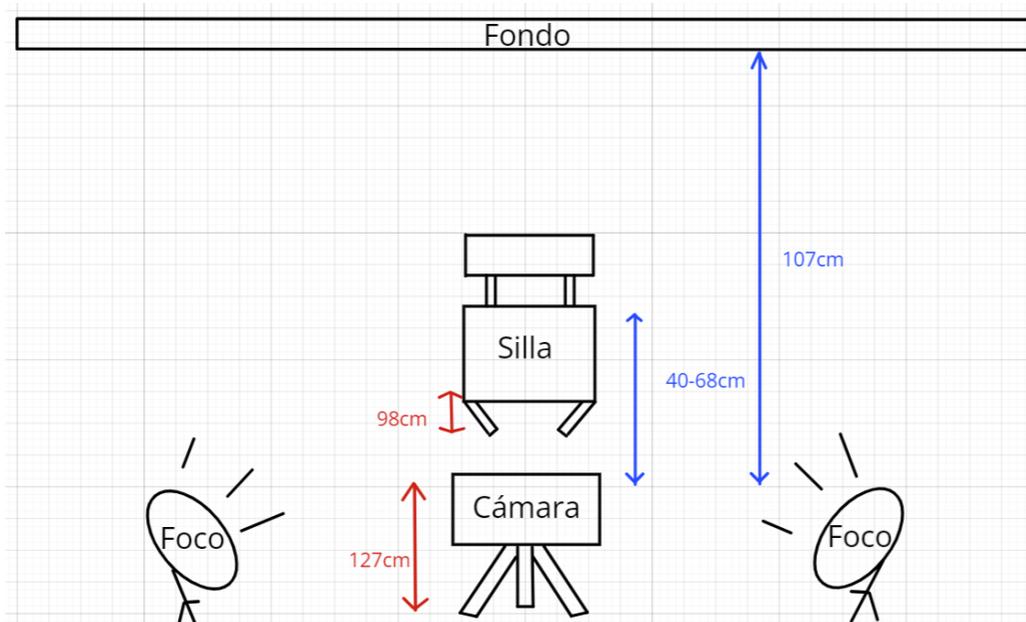


Nota: Elaboración propia.

Las medidas previamente señaladas pueden diferir ya que en la captura de imágenes se tuvo que ajustar un poco la toma debido a las características de altura, postura y fisionomía de cada individuo. En la Figura 3.5 se muestra un bosquejo del diagrama que representa el entorno de captura de las imágenes.

Figura 3.6

Diagrama de elementos involucrados en la captura de imágenes



Nota: Elaboración Propia.

Para poder realizar las capturas se utilizó la herramienta del sensor y el lenguaje Python para poder acceder al hardware y realizar la captura continua de los fotogramas (tanto en RGB como a profundidad) hasta que se guardará la toma por medio de una entrada de teclado. Al inicio surgió la preocupación ya que el proceso de captura resultó ser, en práctica, tedioso y caro en cuanto a tiempo invertido. Debido a lo anterior se tomaron dos medidas para mejorar el proceso de captura:

1. Cambiar el formato de la captura en RGB, ya que al inicio del proyecto se guardaba en un archivo de texto al igual que la imagen de profundidad, sin embargo, era mucho

más tardado debido a que se guardaba el valor de tres canales de color por píxel, incrementando el espacio de memoria requerido y el tiempo de procesamiento.

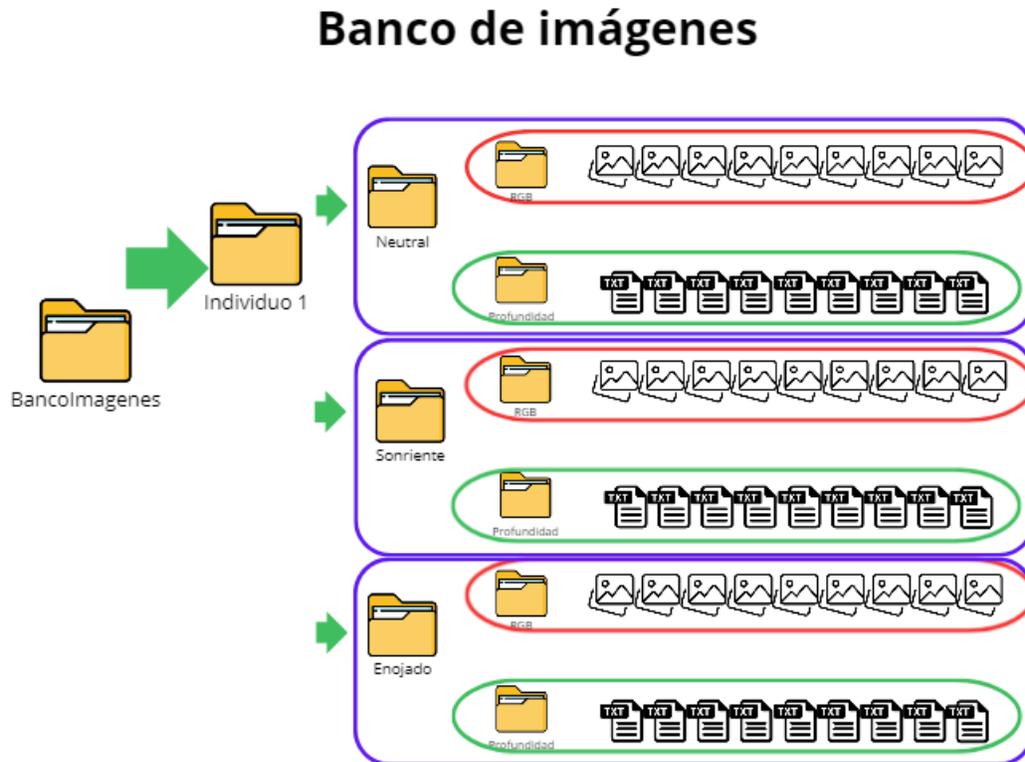
2. La creación de un sistema que permitiera realizar la captura de las imágenes en ambos formatos y las organizara en sus debidas carpetas e involucra la mínima interacción con el usuario.

Ambos puntos mejoraron en gran medida el proceso de captura cuyo tiempo inicial de captura rondaba los 90 minutos por individuo y requería mucha interacción del usuario. Con el primer punto se logró obtener un tiempo aproximado de 30 minutos por captura de individuo y con la realización del sistema se obtuvo un tiempo aproximado de 10 minutos de captura por individuo y con una muy baja intervención por parte del usuario en todo el proceso de captura.

Este sistema fue denominado SCRCP (Sistema Capturador de Rostros a Color y a Profundidad) y ayuda a crear las carpetas organizadas para los diferentes individuos que serán capturados. Sólo requiere que el usuario ingrese por teclado un input para confirmar la captura de las diferentes poses que serán tomadas de manera consecutiva. Este sistema resulta especialmente útil ya que el usuario no necesita una instrucción especializada para poder capturar imágenes. En la Figura 3.7 se muestra la distribución de carpetas y archivos capturados por individuo.

Figura 3.7

Distribución de carpetas por individuo capturadas con el SCRCP



Nota: Elaboración propia.

El sistema permite ver el flujo constante de fotogramas y capturar a través de una lectura de teclado. El sistema proporciona una visualización de los datos que se están capturando con la cámara ToF tanto en formato RGB como en profundidad. En la Figura 3.8 se puede apreciar ambas perspectivas.

El SCRCP pasó por múltiples etapas para mejorar el desempeño de este y para automatizar el uso de este. En la Tabla 3.1 se muestran las iteraciones de desarrollo, así como el impacto de cada iteración.

Tabla 3.2

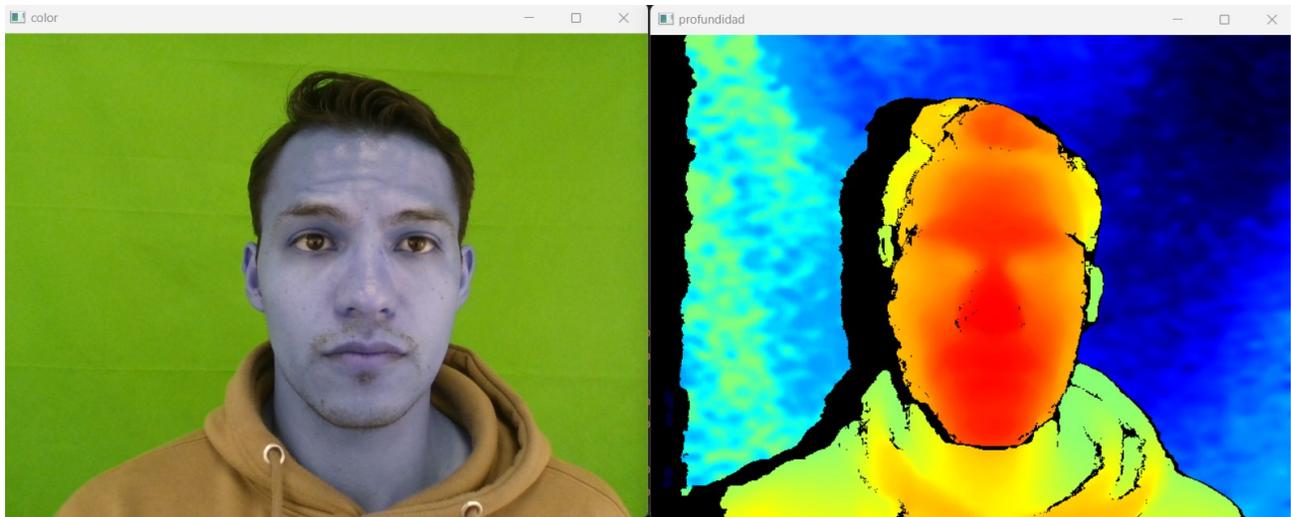
Iteraciones del sistema SCRCP

Número de poses capturadas	9 poses: De frente, frente derecha, frente izquierda, arriba frente, arriba izquierda, arriba derecha, abajo frente, abajo derecha y abajo izquierda
Número de expresiones capturadas	3 expresiones: Neutral, sonriente y enojado
Número capturas por pose	20 imágenes por pose: 10 a profundidad y 10 a color. Cada captura separada por 0.2 segundos
Número total de capturas por individuo	540: 20 fotos por pose, cada pose tiene 3 expresiones
Resolución de las imágenes en el CCPR	Ambos formatos comparten la misma resolución, 640 x 480 píxeles

Nota: Elaboración Propia

Figura 3.8

Visualización de las imágenes capturadas con la cámara ToF

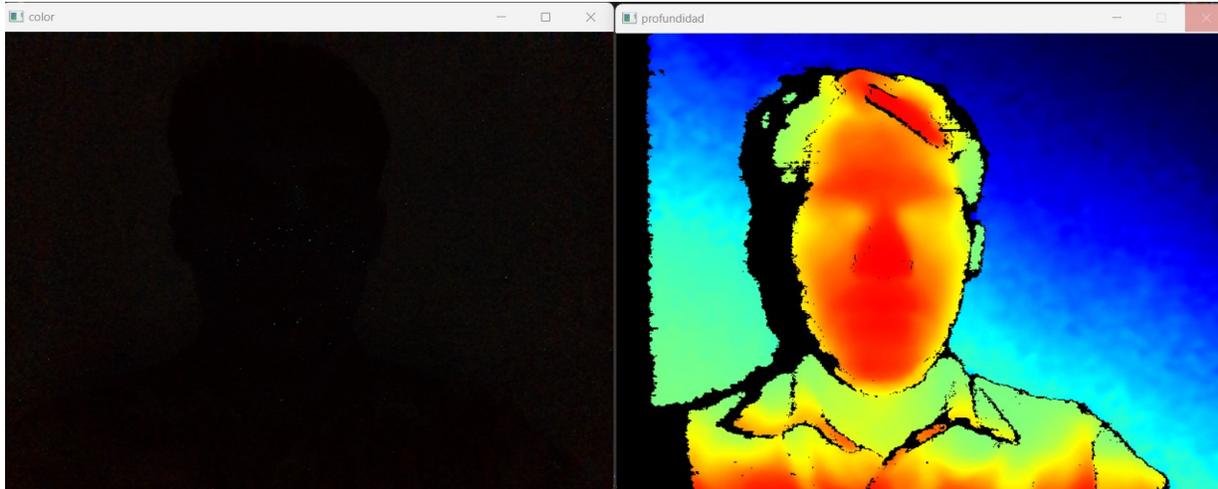


Nota: Elaboración propia.

Anteriormente se mencionó que las imágenes capturadas por la cámara ToF no se ven afectadas por la iluminación, En la Figura 3.9 se puede apreciar la vista de la cámara ToF para capturar la imagen RGB y la imagen a profundidad en un ambiente sin luz, se puede ver que la imagen RGB se ve completamente oscura mientras que la imagen a profundidad no se ve afectada en su lectura.

Figura 3.9

Perspectiva de la imagen RGB (izquierda) y la imagen a profundidad (derecha) sin luz con la cámara ToF



Nota: Elaboración propia

3.3.2 Extracción de Características

Para realizar la extracción de características de las imágenes a profundidad usaremos el extractor Harris Corner. Al ser aplicado a la imagen de profundidad, las características resaltan los cambios de profundidad. Esto permite que el reconocimiento facial sea independiente de la iluminación que recibe el individuo capturado. En la Figura 3.10 se muestra una comparación donde la imagen de la izquierda exhibe las características obtenidas con Harris Corner en una imagen RGB con iluminación moderada y a la derecha podemos ver las características obtenidas con una imagen bajo la misma perspectiva, pero con menor iluminación. De esta forma podemos ver que Harris Corner es sensible a los

cambios de iluminación. Sin embargo, como ya hemos visto en la figura anterior, nuestra imagen a profundidad no muestra cambios con la iluminación.

Figura 3.10

Harris Corner en imágenes RGB con diferente iluminación



Nota: Elaboración propia.

3.4 Producción de documentación

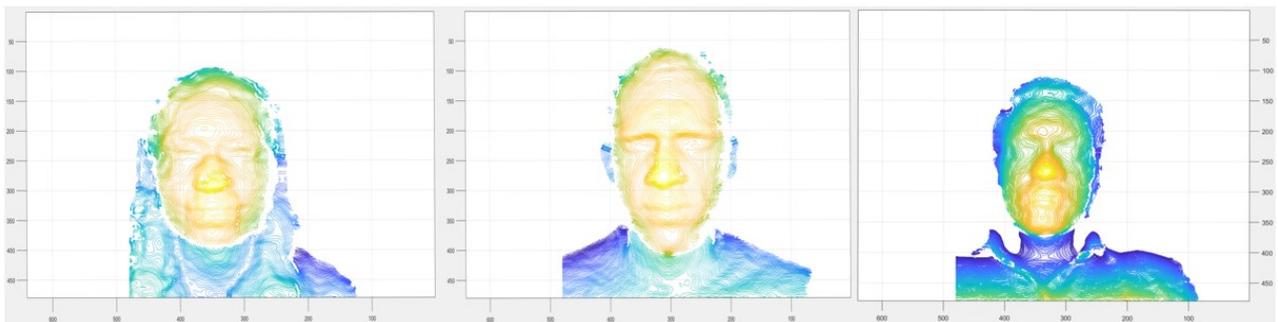
El SCRCP fue concluido, documentado y registrado ante el INDAUTOR (Instituto Nacional del Derecho de Autor) como programa de computación con el número de registro 03-2023-103012383100-01. De la misma manera, el banco de imágenes CRCP fue ingresado para su registro ante el INDAUTOR.

IV. Resultados

En la Figura 4.1 podemos visualizar una muestra de archivos de texto del banco de imágenes interpretado como matrices de puntos de distancia. De esta manera, los cambios detectados entre los píxeles son relativos a un cambio de distancia entre dichos puntos y no de color como en una imagen RGB. Estas muestras nos ayudarán a demostrar el proceso realizado en este proyecto para lograr la creación del grafo de identidad. Estas imágenes cuentan con un preprocesamiento que nos permite eliminar el fondo de la imagen y así centrarnos en los puntos representativos al rostro, esto se logró gracias a un simple filtro que descarta aquellos puntos que superan una distancia cercana a la distancia entre la cámara y el fondo.

Figura 4.1

Imágenes a profundidad preprocesadas

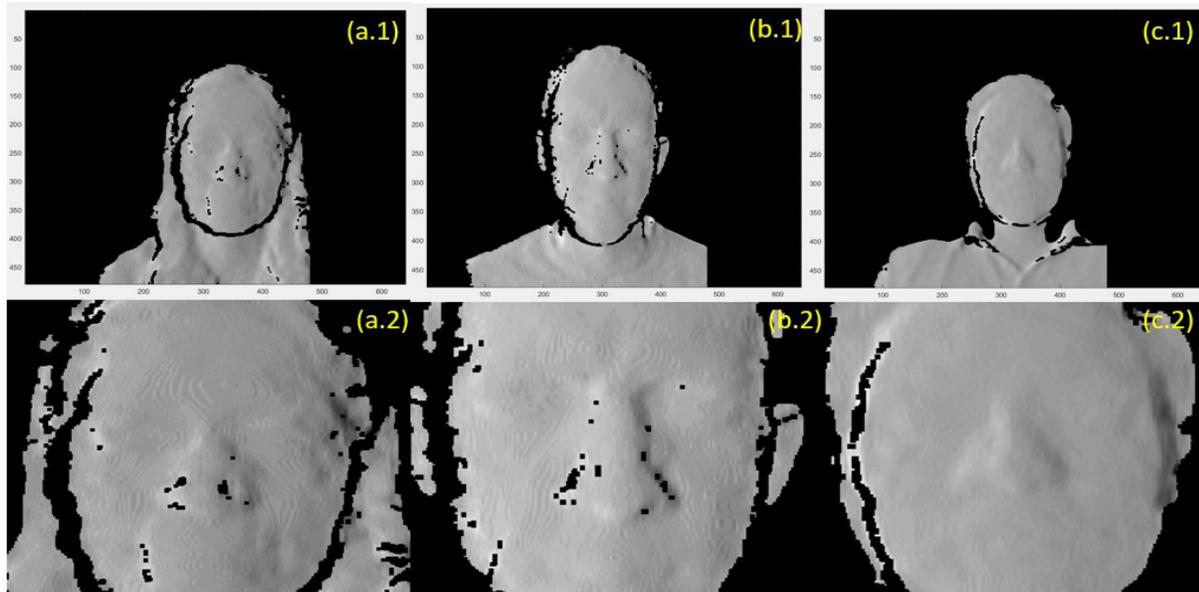


Nota: Elaboración propia.

Posterior al preprocesamiento, se generaron imágenes que permiten ver la topología del rostro, como se puede ver en la Figura 4.2, esto tiene el objetivo de brindar una clara idea de las capas que componen al rostro, así como su profundidad.

Figura 4.2

Topología del rostro en imágenes a profundidad



Nota: Elaboración propia.

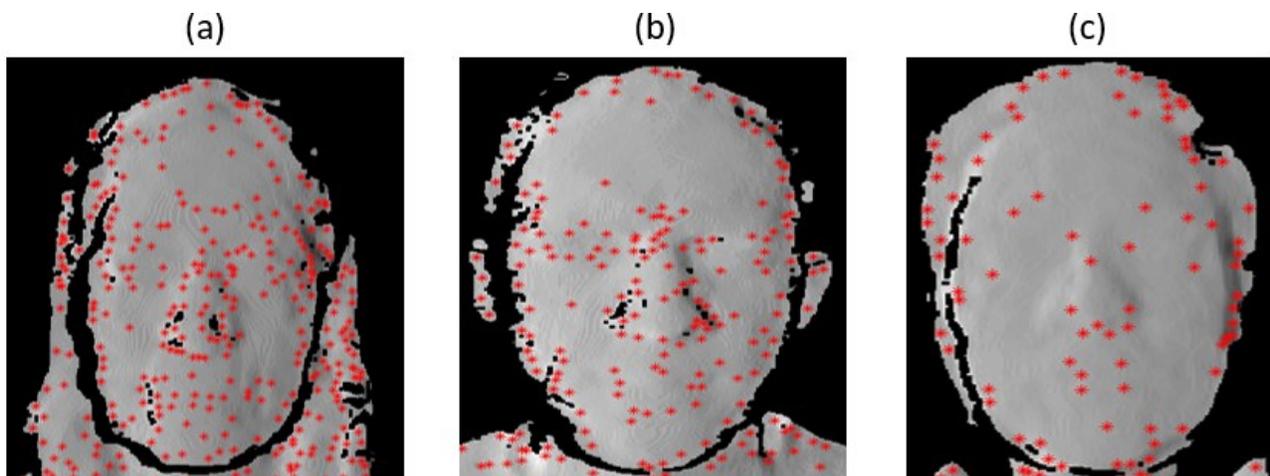
En la Figura 4.2 se pueden apreciar capas que nos ayudan visualizar la topología individual de cada rostro, de izquierda a derecha tenemos rostros diferentes y en la parte inferior (a.2, b.2, c.2) de la imagen tenemos la misma imagen de la parte superior (a.1, b.1, c.1) pero con ligero aumento. En las imágenes a.2, b.2 y c.2 se pueden observar detalles topológicos del rostro que serían omitidos en una imagen RGB, por ejemplo, la parte superior derecha de

la frente en la imagen a.2 presenta un crecimiento ligeramente inclinado a la derecha mientras que en la imagen b.2 se aprecia un abultamiento mayor en cada ceño a comparación de las imágenes a.2 y c.2 (a pesar de ser capturas con una expresión neutral) y, finalmente, en la imagen c.2 se presenta una falta de profundidad en el área del ceño y la nariz a comparación de las imágenes a.2 y b.2.

Lo anterior demuestra que cada rostro presenta topologías únicas que podrían ser omitidas si no se analiza la profundidad en imágenes. Entonces, las características que sean generadas a partir de estas topologías permiten identificar zonas únicas del rostro analizado. En la Figura 4.3 se muestran las características extraídas de los mismos rostros con el descriptor Harris-Corner.

Figura 4.3

Características obtenidas de las imágenes a profundidad

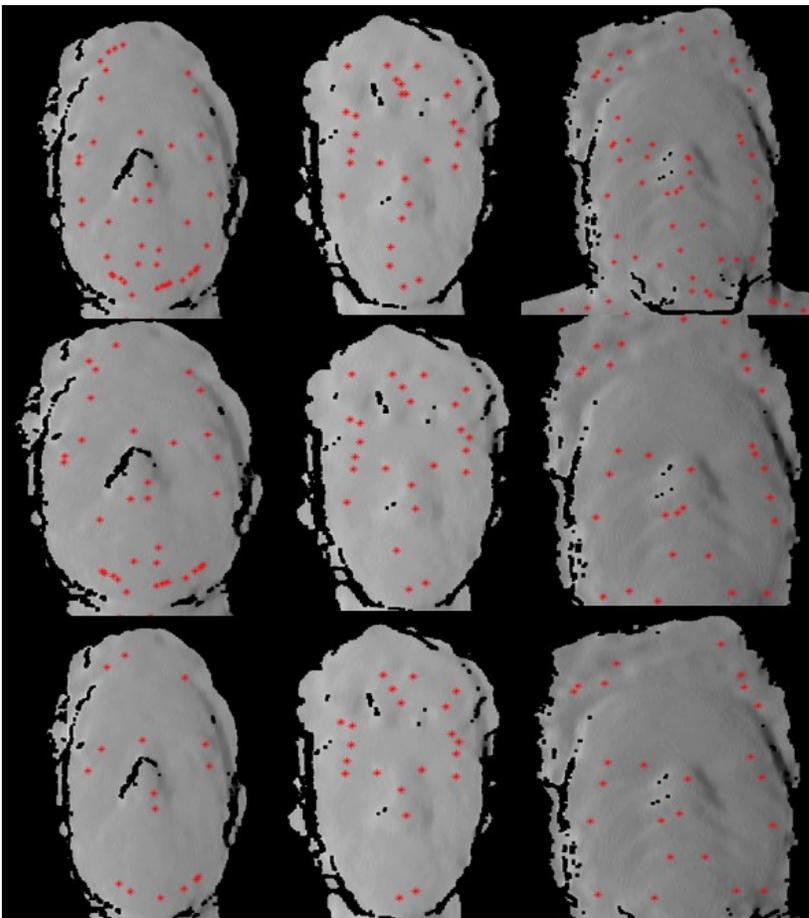


Nota: Elaboración propia.

En la Figura 4.3 se pueden observar las características obtenidas con Harris Corner, también se puede ver que las características se presentan en las zonas que topológicamente muestran puntos de interés de dicho rostro. Sin embargo, existen puntos que llegan a ser redundantes, por lo que se realiza un barrido con la finalidad de dejar aquellos puntos más significativos. Como se muestra en la Figura 4.4.

Figura 4.4

Comparación de características obtenidas

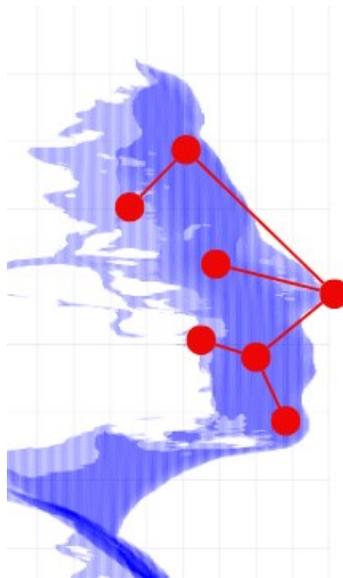


Nota: Elaboración propia.

De las características obtenidas, se analizó que el punto central, y más cercano al sensor, serían las características obtenidas del punto referente a la nariz del rostro, lo que presento una oportunidad de obtener de manera rápida un origen para el grafo que se genera con las características del rostro. En la Figura 4.5 se muestra gráficamente lo anterior.

Figura 4.5

Diseño de la construcción del Grafo Identidad



Nota: Elaboración propia.

Cada característica extraída se convierte en un nodo en el grafo, y se establecen conexiones basadas en la proximidad y relación espacial entre estas características. Este grafo resultante se convierte en una huella facial única para esa persona.

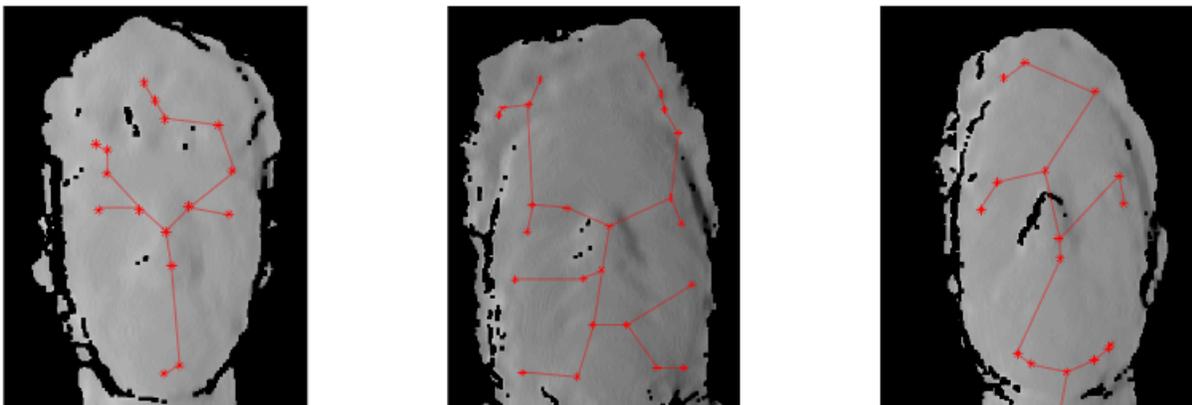
Para autenticar la identidad, el sistema compara la huella facial del individuo en tiempo real con la huella almacenada en la base de datos mediante el uso de técnicas de isomorfismo. El enfoque se basa en la aplicación de isomorfismo para analizar la similitud entre los grafos

generados a partir de las características faciales. El isomorfismo se utiliza para establecer correspondencias estructurales entre los nodos y bordes de ambos grafos, determinando si existen coincidencias suficientes.

En el proceso de autenticación biométrica, el isomorfismo se convierte en una herramienta clave para identificar patrones y relaciones específicas entre las características faciales. Al considerar la disposición relativa de nodos que representan rasgos como ojos, nariz y boca, el sistema puede determinar la autenticidad de una huella facial mediante la búsqueda de correspondencias isomórficas. En la Figura 4.6 se muestra una representación gráfica del grafo generado a partir de las características.

Figura 4.6

Grafos generados a partir de las características



Nota: Elaboración propia.

V. Conclusiones

En resumen, el proyecto de reconocimiento facial 3D mediante el uso de imágenes a profundidad ha representado un acercamiento diferente a los sistemas de identificación biométrica convencionales. La implementación de esta tecnología plantea una nueva manera de abordar el problema abierto de procesamiento de imágenes, esto debido que al usar a la profundidad se puede superar desafíos presentes en los métodos convencionales de reconocimiento facial basados en imágenes 2D.

A lo largo de esta investigación, hemos explorado y desarrollado un enfoque que capitaliza las ventajas de la captura de imágenes tridimensionales, permitiendo una representación con más información que permite conocer la geometría de los detalles del rostro, generando así las características faciales. El uso de imágenes a profundidad ha ofrecido una mayor resistencia a las variaciones de iluminación, cambios posturales y posibles alteraciones en la apariencia facial como lo sería el maquillaje o vello en el rostro.

Los resultados han mostrado que una interpretación topológica para el reconocimiento facial incorpora más información que un análisis de colores en una imagen bidimensional. Se reconoce la necesidad de continuar investigando y perfeccionando este enfoque. Futuros desarrollos podrían centrarse en la obstrucción del rostro en los algoritmos, la selección de las características que representen más el área de interés facial, el cambio del descriptor

de características, incluso el desarrollo de un descriptor que se adapte al análisis topológico de la imagen.

El desarrollo de este proyecto abre más preguntas que desembocan en trabajos futuros con la finalidad de explorar más el reconocimiento facial por medio del análisis topológico del rostro.

VI. Referencias

- Adjabi, I., Ouahabi, A., Benzaoui, A., & Taleb-Ahmed, A. (2020). Past, Present, and Future of Face Recognition: A Review. *Electronics*, 9(8), 1188. <https://doi.org/10.3390/electronics9081188>
- Bay, H. et al., (2008) "Speeded-up robust features (surf)," *Computer Vision and Image Understanding*, 110(3), pp. 346–359. Obtenido de <https://doi.org/10.1016/j.cviu.2007.09.014>.
- Bowler, L., & Large, A. (2008). Design-based research for LIS. *Library & Information Science Research*, 30(1), 39–46. <https://doi.org/10.1016/J.LISR.2007.06.007>
- Chihaoui, M. et al., (2016) "A survey of 2D face recognition techniques", *Computers*, 5(4), p. 21. Obtenido de <https://doi.org/10.3390/computers5040021>.
- Choi, H.-J., & Lee, S.-H. (2009). Automatic Facial Feature Extraction for 3D Face Recognition. *Journal of Applied Sciences*, 9(20), 3739-3745. <http://koreascience.or.kr/article/JAKO200920237949770.pdf>
- Gupta, S., Castleman, K. R., Markey, M. K., & Bovik, A. C. (2010). Texas 3D Face Recognition Database. 2010 IEEE Southwest Symposium on Image Analysis and Interpretation (pp. 97-100). IEEE.
- Hamouz, M. (20 de mayo de 2014). Feature-based affine-invariant detection and localization of faces. Surrey, Guildford, Inglaterra: ResearchGate.

- Hansard, M. et al., (2021) Time of flight cameras: Principles, methods, and applications, <https://www.researchgate.net/>. Obtenido de https://www.researchgate.net/publication/281885252_Time_of_Flight_Cameras_Principles_Methods_and_Applications.
- He, M. (2021). Research on Face Image Digital Processing and Recognition Based on Data Dimensionality Reduction Algorithm. Computational Intelligence and Neuroscience, 2021, 1-10. <https://doi.org/10.1155/2021/3348225>
- Klare, B.F., Klein, B., Taborsky, E., Blanton, A., Cheney, J., Allen, K., Grother, P., Mah, A., Burge, M., & Jain, A.K. (2015). Pushing the frontiers of unconstrained face detection and recognition: IARPA Janus Benchmark A. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 1931–1939). Boston, MA, USA.
- Kortli, Y., Jridi, M., Al Falou, A., & Atri, M. (2020). Face Recognition Systems: A Survey. Sensors, 20(2), 342. MDPI AG. Obtenido de <http://dx.doi.org/10.3390/s20020342>
- Križaj, J., Štruc, V. and Pavešić, N. (2010) "Adaptation of sift features for robust face recognition", Lecture Notes in Computer Science, pp. 394–404. Obtenido de https://doi.org/10.1007/978-3-642-13772-3_40.
- Madrigal-González, C. A., Ramírez-Madrigal, J. L., Hoyos-Arbeláez, J. C., & Stephen-Fernández, D. (2007). Diseño de un sistema biométrico de identificación usando sensores capacitivos para huellas dactilares. Revista Facultad de Ingeniería Universidad de Antioquia, 21-32.

- Mahouachi, D., & Akhloufi, M. A. (2023). Recent Advances in Infrared Face Analysis and Recognition with Deep Learning. *AI*, 4(1), 199–233. MDPI AG. Obtenido de <http://dx.doi.org/10.3390/ai4010009>
- Martinez, A. M., & Benavente, R. (1998). The AR face database. *CVC Tech. Rep.*, 24, 1-10.
- Min, R., Kose, N. and Dugelay, J.-L. (2014) “KinectFaceDB: A Kinect database for face recognition,” *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 44(11), pp. 1534–1548. Available at: <https://doi.org/10.1109/tsmc.2014.2331215>.
- Ouamane, A. et al., (2014) “Robust multimodal 2D and 3D face authentication using local feature fusion,” *Signal, Image and Video Processing*, 10(1), pp. 129–137. Obtenido de <https://doi.org/10.1007/s11760-014-0712-x>.
- Phillips, P.J., Wechsler, H., Huang, J., & Rauss, P. (1998). The FERET database and evaluation procedure for face recognition algorithms. *Image and Vision Computing*, 16(5), 295-306.
- Prawira, W. et al., (2017) “The detection of 3D objects using a method of a Harris Corner Detector and Lucas-Kanade Tracker based on Stereo Image,” 2017 International Conference on Electrical Engineering and Computer Science (ICECOS). Obtenido de <https://doi.org/10.1109/icecos.2017.8167126>.
- Rublee, E. et al., (2011) “Orb: An efficient alternative to SIFT or surf,” 2011 International Conference on Computer Vision. Obtenido de <https://doi.org/10.1109/iccv.2011.6126544>.

- Ruiz Marín, M., Rodríguez Uribe, J. C., & Olivares Morales, J. C. (2009). Una mirada a la biometría. *Revista Avances en Sistemas e Informática*, 6(2), 29-38.
- Shi, J. and Tomasi, C. (1994) "Good features to track," *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition CVPR-94*. Obtenido de <https://doi.org/10.1109/cvpr.1994.323794>.
- Simón, M.O., Corneanu, C., Nasrollahi, K., Nikisins, O., Escalera, S., Sun, Y., Li, H., Sun, Z., Moeslund, T.B., & Greitans, M. (2016). Improved RGB-D-T based face recognition. *IET Biom*, 5(4), 297-303.
- Soltanpour, S., Boufama, B. and Jonathan Wu, Q.M. (2017) "A survey of local feature methods for 3D face recognition", *Pattern Recognition*, 72, pp. 391–406. Obtenido de <https://doi.org/10.1016/j.patcog.2017.08.003>.
- Stan, L.Z. et al., (2013) The Casia Nir-Vis 2.0 Face Database - cv-foundation.org, Computer Vision Foundation. IEEE Xplore. Available at: https://www.cv-foundation.org/openaccess/content_cvpr_workshops_2013/W13/papers/Li_The_CASIA_NIR-VIS_2013_CVPR_paper.pdf (Accessed: May 1, 2023).
- Strobel, J., Jonassen, D., & Ionas, I. (2008). The evolution of a collaborative authoring system for non-linear hypertext: A design-based research study. *Computers and Education*, 51(1), 67–85. <https://doi.org/10.1016/j.compedu.2007.04.008>
- Tamilselvi, M. and Karthikeyan, S. (2018) A literature survey in face recognition techniques, *A Literature Survey in Face Recognition Techniques*. *International Journal of Pure*

and Applied Mathematics. Obtenido de <https://www.acadpubl.eu/jsi/2018-118-16-17/articles/16/54.pdf>

Zhu, P., Binyuan Hui, H., Zhang, C., Du, D., Hu, Q., Wen, L. (2019, 6 agosto). Multi-view Deep Subspace Clustering Networks. arXiv.org. <https://arxiv.org/abs/1908.01978>

Zhu, X., Yi, D., Lei, Z., Li, S. Z. (06 de diciembre de 2014). Robust 3D Morphable Model Fitting by Sparse SIFT Flow. Estocolmo, Estocolmo, Suecia: IEEE Xplore.