

Universidad de Alcalá

Escuela Politécnica Superior

Grado en Ingeniería Electrónica de Comunicaciones



Trabajo Fin de Grado

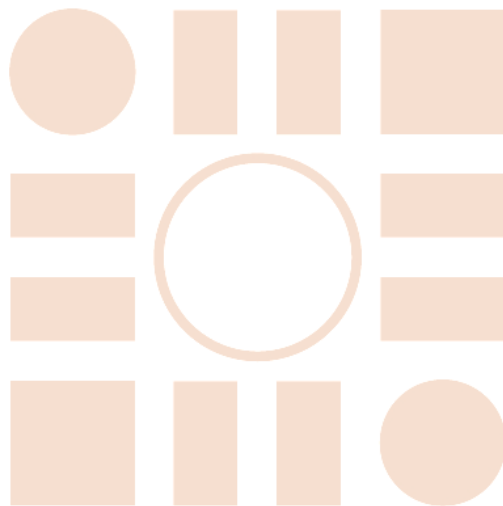
Re-identificación de personas utilizando mapas de profundidad e imágenes infrarrojas capturadas por una cámara de tiempo de vuelo (ToF)

Autor: Miguel Sebastián Pérez

Tutor/es: Carlos Andrés Luna Vázquez

2019/2020

Universidad de Alcalá
Escuela Politécnica Superior



ESCUELA POLITECNICA
SUPERIOR

UNIVERSIDAD DE ALCALÁ
Escuela Politécnica Superior

Grado en Ingeniería Electrónica de Comunicaciones

Trabajo Fin de Grado

**Re-identificación de personas utilizando mapas de
profundidad e imágenes infrarrojas capturadas por una
cámara de tiempo de vuelo (ToF)**

Autor: Miguel Sebastián Pérez

Tutor/es: Carlos Andrés Luna Vázquez

TRIBUNAL:

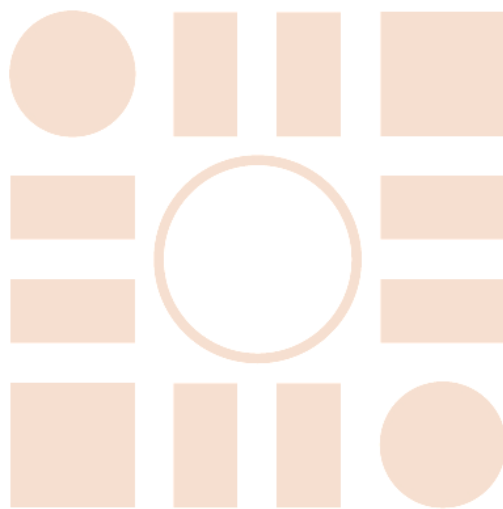
Presidente: Manuel Mazo Quintas

Vocal 1º: Carlos Andrés Luna Vázquez

Vocal 2º: Ernesto Martín Gorostiza

FECHA: 10/06/2020

El estudio y la preparación puede que haga gente intelectual. Pero la humildad
y sencillez hace personas “Sabias”
Arturo Hdez. L



ESCUELA POLITECNICA
SUPERIOR

Agradecimientos

Con este documento se cierra un ciclo en mi vida, un ciclo cargado de muchos vaivenes tanto emocionales como físicos. Nunca pensé que esta etapa en mi vida fuera a estar cargada de tanta evolución personal, gracias a ella he aprendido a conocer realmente el valor de las cosas y como la lucha y la constancia son valores primordiales que una persona debe tener para conseguir sus objetivos. Ahora me encuentro aquí, redactando mi TFG, lo que supone mi último escoyo para poder ser graduado en Ingeniería Electrónica de Comunicaciones.

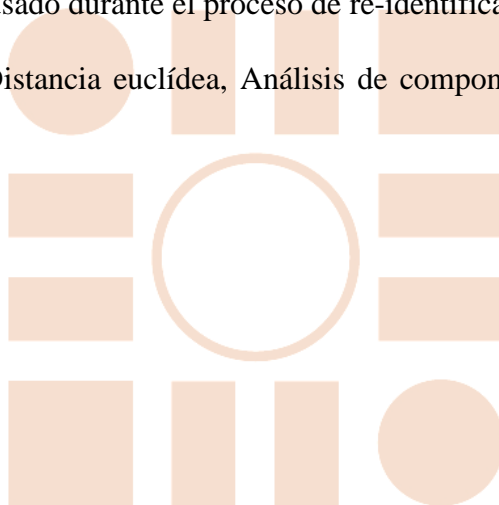
Quiero agradecer la ayuda inestimable de mi tutor Carlos Andrés Luna, sin él, la realización de este trabajo no hubiera sido posible, ya que me ha aportado todo su conocimiento y me ha facilitado cualquier información necesaria, como así ha estado siempre que le he pedido opinión. También, al grupo GEINTRA como a todas las personas que me han facilitado el área inteligente para poder tomar las muestras necesarias para la realización de este trabajo. No quiero terminar este capítulo de agradecimientos sin dar las gracias a mi familia y amigos, que tanto para lo bueno como para lo malo siempre han estado ahí aguantándome en mis días insoportables en los que estaba saturado de estudiar y estudiar y no llegar a entender nada. Estaré eternamente agradecido a cada una de las personas que forman y han formado parte de mi vida y me han ayudado a conseguir lo que me he propuesto.

ESCUELA POLITECNICA
SUPERIOR

Resumen

La finalidad de este trabajo es el desarrollo de un algoritmo capaz de realizar la identificación y re-identificación de personas, a partir de unas imágenes en profundidad e intensidad, obtenidas con un sensor de tiempo de vuelo (ToF) ubicado en posición cenital. El uso de este tipo de sensores posibilita identificar un individuo cuando pasa por la escena que observa un sensor y re-identificar cuando vuelva a pasar por la misma o cuando pase por la escena observada por otro sensor, todo ello salvaguardando la privacidad del individuo. Además, se implementó una interfaz de usuario para facilitar la comunicación entre usuario-máquina. La detección se realiza mediante el análisis de las secuencias de intensidad y profundidad, de las cuales se extrae un conjunto o vector de características de cada persona presente en las mismas. Este vector es almacenado en una base de datos, para ser usado durante el proceso de re-identificación.

Palabras Clave: ToF, Distancia euclídea, Análisis de componentes principales (PCA), Histograma.



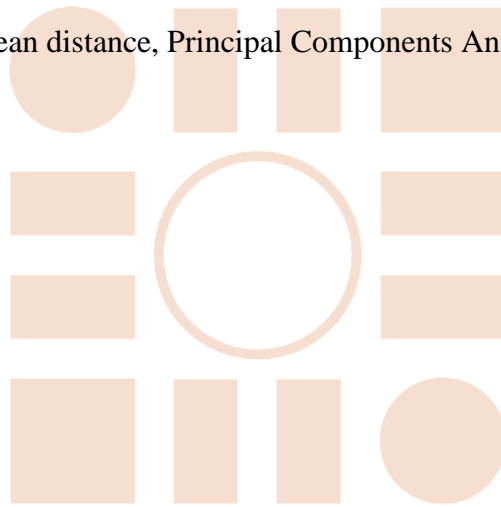
ESCUELA POLITECNICA
SUPERIOR

Abstract

The main objective of this work is the creation of a app capable of re-identifying people using only Depth and Intensity images obtained by a Time of Flight (ToF) sensor placed in zenithal position. This app is equipped with a user interface to facilitate communication between machine-user. This interface will have several options to choose, depending on the user's need, the options will be detailed later in the document.

Detection is carried out by analyzing the intensity and depth sequences, where there are a series of components that each of the sequences detects. These components are stored in the database. For re-identification, the stored components are used, to compare them with the sequence to be analyzed with different used techniques.

Keywords: ToF, Euclidean distance, Principal Components Analysis (PCA), Histogram.



ESCUELA POLITECNICA
SUPERIOR

Índice

Contenido

Resumen	8
Abstract.....	9
Índice	10
Índice de Figuras	12
Índice de tablas	14
Lista de Acrónimos.....	15
1.Introducción.....	16
1.1 Introducción.....	16
1.2 Sistema Propuesto	17
2. Estudio Teórico	19
2.1 Introducción.....	19
2.2 Cámaras de tiempo de vuelo (ToF)	19
2.2.1 Cámara ToF seleccionada.....	20
2.2.1.1 Medida de la distancia	20
2.2.1.2 Precisión	21
2.3.3 Factores que influyen en la medida de la distancia	22
2.4.1 Características generales de la cámara Basler	24
2.4.2 Imágenes en profundidad.....	26
2.5 Ventajas que nos ofrece la cámara Basler respecto a otras	26
2.2 Análisis de componentes principales (PCA)	27
2.2.1 Procedimiento matemático	27
3. Desarrollo y Diseño	29
3.1 Introducción.....	29
3.2 Bloque de adquisición de imágenes.....	30
3.3 Bloque Cierre de la cámara	30
3.4 Bloque Análisis de secuencias.....	31
3.4.1 Etapa de Identificación	31
3.4.4 Etapa de Re-Identificación	38
4. Resultados.....	39
4.1 Introducción.....	39
4.2 Base de datos de secuencias	39
4.3 Resultados Cámara Basler	40
4.4 Resultados Obtenidos aplicando la técnica de PCA.....	41

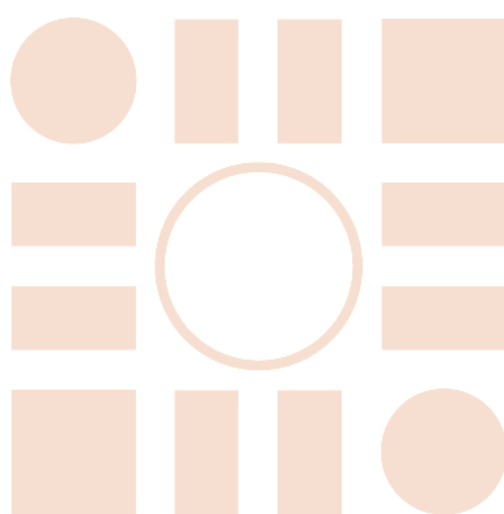
4.5 Resultados finales obtenidos.	42
5.Conclusiones y trabajos futuros.....	50
5.1. Conclusión personal.	50
5.3 Trabajos futuros.....	50
6. Bibliografía.....	51
Anexo A.....	52
A.1 Manual de Usuario.....	52
A.1.1 Requisitos del Sistema.....	52
A.1.2 Estructura de los directorios	52
A.1.3 Uso de la Aplicación.....	55
Anexo B.....	58
B.1 Pliego de condiciones	58
B.1.1 Hardware.....	58
B.1.2 Software.....	58
Anexo C.....	59
C.1 Presupuesto	59



ESCUELA POLITECNICA
SUPERIOR

Índice de Figuras

Figura 1.1 Diagrama de bloques de la solución propuesta.....	17
Figura 2.1 Parte delantera y trasera de la cámara Basler.....	20
Figura 2.2. Principio de funcionamiento de la medida de distancia con la cámara Basler	20
Figura 2.3 Precisión en las medidas	21
Figura 2.4 Efectos de dispersión de luz.....	22
Figura 2.5 Efecto Multipath.....	23
Figura 3.1 Bloques del sistema.....	29
Figura 3.2 Proceso de adquisición de las imágenes	30
Figura 3.3 Imágenes de Profundidad.....	30
Figura 3.4 Imágenes de Intensidad.....	30
Figura 3.5 Diagrama general de bloques del detector	31
Figura 3.6 Esquema de la posición de la cámara.....	32
Figura 3.7 Esquema de altura de interés.....	33
Figura 3.8 División en secciones.....	35
Figura 3.9 Ejemplo de obtención de las medidas para la cabeza y los hombros	35
Figura 3.10 Histograma de intensidad de la Seq-DFJ-D02.....	37
Figura 3.11 Histograma de la Seq-DFJ-D02 de todas las secciones	37
Figura 3.12 Comparativa de histogramas, (a) Histograma correspondiente a seq-DFJ-D02; (b) Histograma correspondiente a seq-DFJ-D01	38
Figura 3.13 Comparativa de histogramas; (a) Histograma correspondientes a seq-MSP-D02; (b) Histograma correspondiente a seq-DFJ-D01	38
Figura 4.1 Gráfico de los resultados obtenidos con el uso de la PCA.....	43
Figura 4.2 Punto donde se extraen los valores	44
Figura 4.3- Imagen de intensidad representando los diferentes valores de intensidad encontrados en el pelo.....	47
Figura 4.4 Comparativa de resultados	51
Figura A.1 Directorios utilizados en la aplicación	52
Figura A.2 Pantalla inicial de la interfaz de usuario.....	55
Figura A.3 Segunda pantalla de la interfaz de usuario	55
Figura A.4 Tercera pantalla de la interfaz de usuario.....	56



ESCUELA POLITECNICA
SUPERIOR

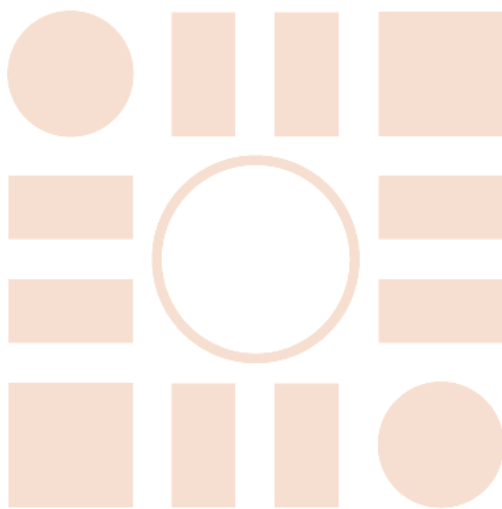
Índice de tablas

Tabla 2.1 Características Basler ToF640-20gm_850nm.	5
Tabla 4.1 Lista de Secuencias a evaluar	41
Tabla 4.2 Comparativa de las alturas reales con las obtenidas a través de la cámara ToF	42
Tabla 4.3 Estructura de las tabla de resultados.....	45
Tabla 4.4 Tabla de aciertos sin aplicar las prioridades para las secuencias	46
Tabla 4.5 Tabla de aciertos aplicando la técnica de 6 histogramas y reducción de parámetros de profundidad	48
Tabla 4.6 Comparativa de resultados de profundidad aplicando o no los componentes extras.....	49
Tabla 4.7 Tabla de aciertos aplicando las prioridades para las secuencias	50
Tabla C.1 Costes de Hardware	59
Tabla C.2 Costes de Software.....	59
Tabla C.3 Costes de mano de obra	59
Tabla C.4 Coste total	60

ESCUELA POLITECNICA
SUPERIOR

Lista de Acrónimos

Acrónimo	Definición
ToF	Referido a cámara de tiempo de vuelo
PCA	Análisis de componentes principales
ROI	Regiones de Interés



ESCUELA POLITECNICA
SUPERIOR

1.Introducción

1.1 Introducción

El uso de sistemas basados en visión artificial ha ido aumentando considerablemente en los últimos años, debido a la disminución del coste de los mismos y al surgimiento de nuevas tecnologías y sensores, que permiten cada vez más mejorar las prestaciones en diferentes aplicaciones. Por ello, desde pequeñas empresas a grandes consorcios están destinando grandes recursos en el desarrollo de nuevas aplicaciones, que utilicen este tipo de sistemas.

En muchas aplicaciones, como pueden ser en sistemas de vigilancia y seguridad de espacios y/o locales, seguimiento de personas para estudios de mercado o movilidad, etc, se requiere de algoritmos identificación y re-identificación de personas. La re-identificación de personas consiste en detectar que un individuo que está siendo captado por un sensor ya ha sido previamente observado por dicho sensor o por otro, y poder diferenciarlo del resto de los usuarios [1].

La finalidad de este trabajo es la creación de un método capaz de realizar un proceso de identificación y re-identificación de personas que entran y salen de un establecimiento, a partir de imágenes de profundidad e imágenes de intensidad obtenidas por un sensor de tiempo de vuelo (ToF) ubicado en posición cenital. La alternativa desarrollada es capaz de identificar y re-identificar a personas salvaguardando su identidad. Esto se logra, gracias a que con el uso de un sensor ToF solo se obtiene información relativa a la forma y a la intensidad de la luz infrarroja reflejada por las diferentes partes de la persona, en este proyecto solamente se utiliza información de la cabeza y los hombros. Para ello, se ha escogido la ubicación del sensor en posición cenital, debido a que además de evitar los errores de oclusión, ayuda a preservar mucho más la identidad de la persona.

Aunque en un principio, la detección, identificación y la re-identificación se realizaría con un clasificador basado en la técnica de análisis de componentes principales (PCA), debido a que los resultados obtenidos mediante las pruebas realizadas no fueron los esperados, se han implementado otras alternativas que mejoran dichos resultados, como se verá en el Capítulo 4.

Para la validación de los resultados se ha creado una base de datos de imágenes de profundidad e intensidad, a partir de un conjunto de individuos con características biométricas diferentes.

En lo que sigue el trabajo se estructura de la siguiente forma: en el apartado 1.2 se muestra de forma general el sistema propuesto. En el apartado 2 se realiza un análisis de las bases teóricas del funcionamiento de la cámara Basler como las características que posee la cámara. En el apartado 3 se detalla el desarrollo y diseño que se ha realizado en este TFG. En el apartado 4 se exponen los resultados obtenidos con las diferentes técnicas usadas. En el apartado 5 se exponen las conclusiones sacadas tras la realización del trabajo. En el apartado 6 se exhibe la bibliografía consultada para este proyecto. Y para finalizar se exponen tres anexos. En el anexo A se detalla un manual de usuario para facilitar el uso del programa. En el anexo B se detalla un pliego de condiciones. En el anexo C se representan mediante tablas el presupuesto requerido para la realización de este trabajo.

1.2 Sistema Propuesto

El sistema propuesto consta de dos etapas, la etapa de identificación, cuando el individuo entra al local y la re-identificación que se ejecuta cuando sale. En la [Figura 1.1](#) se muestran los diferentes bloques o módulos que conforman cada una de las etapas, que se describen a continuación [2]:

- **Etapa de identificación:** Esta etapa se encarga de realizar el cálculo de los componentes principales de la persona a analizar, para posteriormente crear una clase donde almacena dichas componentes para ser usadas por la etapa de re-identificación posteriormente.
- **Etapa de re-identificación:** Lleva a cabo la re-identificación de personas a partir de imágenes de intensidad y profundidad obtenidas de una secuencia grabada. Para ello, es necesario que previamente se haya realizado la etapa de identificación, donde se obtienen las clases con los componentes principales.

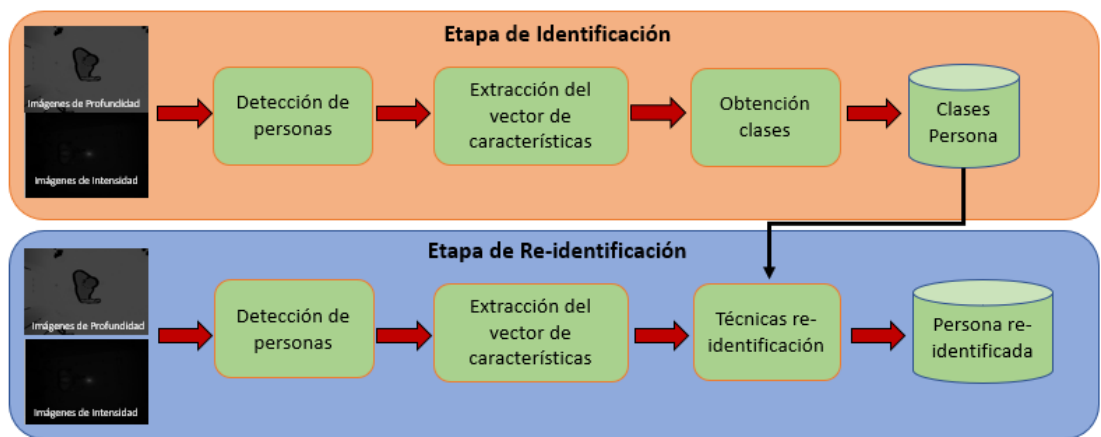


Figura 1.1. Diagrama de bloques de la solución propuesta.

Cabe destacar que los bloques de detección de personas y extracción del vector de características son comunes a ambas partes, por lo que sólo se explican una vez.

1. **Obtención de las imágenes:** Una vez lanzada la aplicación, el usuario escoge la función de obtención de las imágenes ([Véase Anexo A manual de usuario](#)), obteniéndose simultáneamente las imágenes de intensidad y las de profundidad.
2. **Detección de personas:** Consiste primeramente en la reducción del ruido y estimación del valor de los píxeles no válidos en las imágenes de profundidad. Posteriormente, se procede a la extracción de los máximos existentes en las secuencias de profundidad, determinando si coinciden con personas.
3. **Extracción del vector de características:** Consiste en la extracción del vector de características donde se ha almacenado un conjunto de características físicas

de la persona analizada. Este vector se utiliza tanto para la etapa de re-identificación como para la creación de la clase que identifica a la persona.

4. **Creación de la Clase:** Una vez finalizada la secuencia grabada, se crea la clase con las características detectadas por la secuencia, para así, ser capaces de identificar a la persona. Para la creación de clases, como para la posterior clasificación se emplea el análisis de componentes principales (PCA) [1].
5. **Re-Identificación:** El vector de características es extraído tras la detección, este vector es introducido en un clasificador que compara a la persona detectada con las personas para las que se obtuvieron las clases.

El sistema propuesto consta de una interfaz de usuario, que permite, tanto hacer la captura de las imágenes cuando el usuario lo requiera, como también poder detectar y re-identificar a personas, a través de las imágenes tomadas previamente por la cámara y/o las almacenadas en la base de datos.

En los capítulos posteriores se explican con más detalle la funcionalidad del sistema propuesto.

ESCUELA POLITECNICA
SUPERIOR

2. Estudio Teórico

2.1 Introducción

En este apartado, primeramente, se analiza de forma general la funcionalidad de las cámaras ToF, a continuación, se detallan las características de la cámara usada y las ventajas que ofrece en diferencia con las otras técnicas actuales usadas para las cámaras ToF. Finalmente, se procede a la explicación del método de análisis de componentes principales (PCA), por ser el inicialmente planteado para el desarrollo del proyecto.

2.2 Cámaras de tiempo de vuelo (ToF)

Los sensores de tiempo de vuelo destacan por generar imágenes de profundidad en las que cada píxel contiene la distancia entre el punto de la escena y el plano en el que se encuentra el sensor. Además, generalmente ofrecen imágenes de intensidad obtenidas mediante iluminación infrarroja. Es por ello por lo que este tipo de sensores han revolucionado el mercado de los sistemas de visión artificial. A continuación, se exponen una lista de beneficios que aportan las cámaras ToF:

- Proporciona una imagen con datos 2D y 3D en un solo disparo.
- Resolución VGA para imágenes 3D.
- Fácil de integrar.

Las cámaras de tiempo de vuelo se pueden utilizar en diversos sectores en el ámbito empresarial, como:

- **Automatización logística:** Las cámaras se pueden utilizar para el embalaje de envíos, por ejemplo, llenado de cajas, etiquetado, asistencia robot de estantería, etc.
- **Automatización y robótica en las empresas:** Búsqueda de objetos, recogida de objetos, objetos de montaje, detección de objetos dañados.
- **Medicina:** Monitorización de pacientes, en sistemas como resonancias magnéticas.
- **Vehículos autónomos:** Identificación de objetos y personas del entorno.

2.2.1 Cámara ToF seleccionada.

La cámara inicialmente seleccionada para la realización de este proyecto fue la **Basler ToF640-20gm_850nm**, debido entre otras características, a su resolución espacial, a su robustez y distancias de trabajo. En la [Figura 2.1](#) se muestran las imágenes de la parte delantera y trasera de esta cámara.



Figura 2.1. Parte delantera y trasera de la cámara Basler.

2.2.1.1 Medida de la distancia

La cámara Basler ToF utiliza la técnica de luz pulsada, utilizando un conjunto de Leds infrarrojos. Para medir distancias envía los pulsos y mide el tiempo que tardan los pulsos en reflejarse en los objetos y regresar al sensor. El sensor convierte la carga eléctrica generada por la energía luminosa en una información de distancia, que se almacena en los píxeles correspondientes. En la [Figura. 2.2](#) se muestra un esquema del principio de funcionamiento de la medida de distancia con dicha cámara.

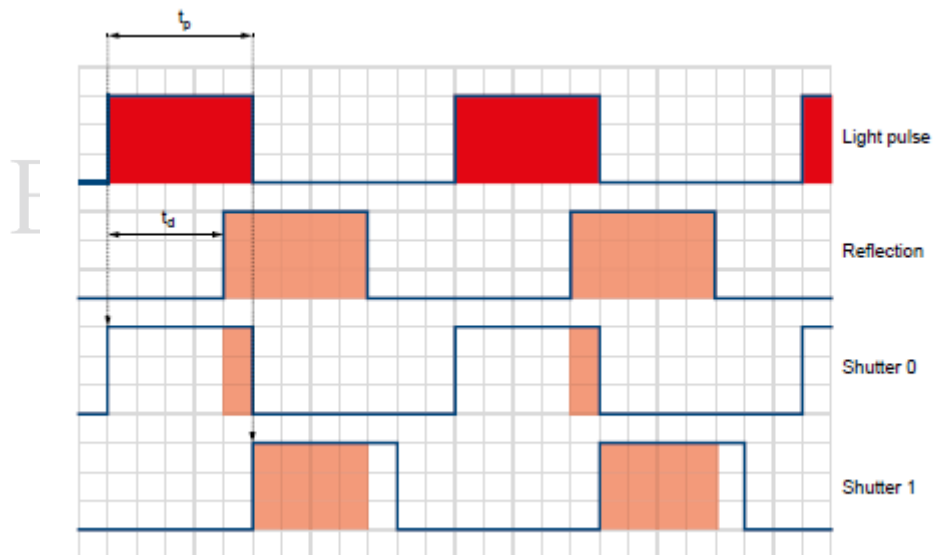


Figura 2.2. Principio de funcionamiento de la medida de distancia con la cámara Basler [3].

Cuando se activa la cámara, todos los LED's se encienden simultáneamente durante un periodo de tiempo específico (t_p) y la energía reflejada en el objeto se recoge en cada píxel, utilizando dos ventanas de obturación durante el mismo t_p . El obturador 0 (**Shutter 0**) se abre cuando comienza el pulso de luz y el obturador 1 (**Shutter 1**) se abre cuando termina el pulso de luz [3]. Este proceso se repite consecutivamente hasta que finaliza el tiempo de integración.

La expresión del cálculo de la distancia es:

$$d = \frac{c}{2} t_p \left(\frac{S_1}{S_0 + S_1} \right) \quad (2.1)$$

Donde:

- c = Velocidad de la luz
- d = Distancia
- t_p = periodo en el que el pulso de luz está activo
- S_0 = La carga eléctrica acumulada durante la primera ventana de obturación S_0
- S_1 = La carga eléctrica acumulada durante la segunda ventana de obturación S_1 .

2.2.1.2 Precisión

Las cámaras Basler ToF están calibradas para un rango de medición de 0.5 a 5 metros. El rendimiento de la cámara se puede caracterizar en los siguientes criterios:

- **Precisión absoluta:** La precisión absoluta es la diferencia media entre la distancia medida y la verdadera distancia. Para lograr un resultado lo más preciso posible, las cámaras Basler ToF se calibran en la fábrica y se prueban para garantizar una precisión confiable.
- **Repetibilidad:** Es una medida estadística definida como la posibilidad de obtener resultados consistentes al replicar un estudio con un conjunto distinto de datos.

El fabricante muestra en [3] que las cámaras Basler ToF tienen una precisión de $\pm 1\text{cm}$ y la repetibilidad de 8 mm. La Figura 2.3 muestra el error en las medidas a diferentes distancias. Sin embargo, estas medidas se han realizado bajo unas condiciones determinadas: sin luz ambiente, a 22 °C de temperatura, con el offset corregido, y solo haciendo 25 medidas con los 40x40 pixeles centrales del sensor. Esto hace que el error sea mucho mayor bajo otras condiciones.

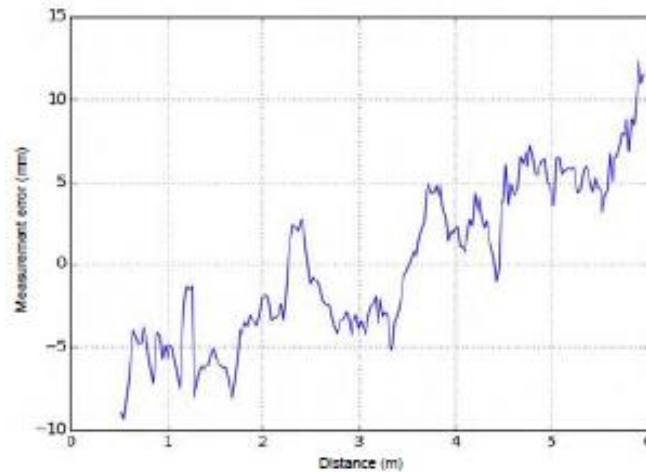


Figura 2.3. Medición de precisión para la cámara Basler ToF [3].

2.3.3 Factores que influyen en la medida de la distancia

2.3.3.1 Luz ambiente

La medición de distancia para la cámara Basler ToF depende de la reflexión de la luz emitida por la cámara, cualquier luz externa a esta, puede ser perjudicial para los resultados obtenidos.

Para poder reducir la influencia de la luz ambiente, se utiliza un filtro de paso banda óptico que permite pasar la luz que pertenezca al mismo espectro que la fuente de luz incorporada por la cámara, eliminando cualquier fuente de luz artificial. El problema principal se encuentra en la luz solar, debido a que es activa en todo el espectro, dependiendo del escenario que encontremos se toman una serie de medidas adicionales para proteger la cámara de sobreexposiciones.

2.3.3.2 Dispersión de luz

Uno de los factores para tener en cuenta por las cámaras ToF es la dispersión de luz, que puede ocurrir, debido a reflexiones no deseadas dentro de la lente de la cámara o detrás de ella y/o debido a la reflexión de la luz en superficies brillantes en la zona iluminada, creando un error en la medición. La Figura 2.4 ilustra este efecto.

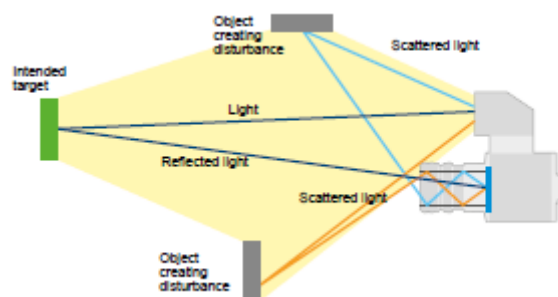


Figura 2.4. Efectos de dispersión de luz [3].

Para combatir el efecto de dispersión de luz, Basler ha desarrollado las siguientes medidas:

- Una selección minuciosa de las lentes a utilizar.
- Uso de un parasol.
- Apertura limitada de la fuente de luz.

2.3.3.3 Múltiples Reflexiones (Multipath)

Debido a la presencia de otros objetos, puede darse el caso de obtener una medida incorrecta. La [Figura 2.5](#) muestra el efecto multicamino (multipath), este efecto es un fenómeno producido cuando las señales llegan al destino por dos o más caminos, y en diferentes tiempos [\[7\]](#).

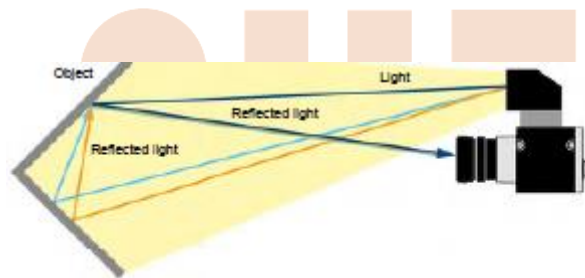


Figura 2.5. Efecto Multipath [\[3\]](#).

2.3.3.4 Temperatura

La temperatura tiene una gran influencia a la hora de realizar mediciones de distancia fiables, es por ello por lo que encontramos diferentes aspectos a ser considerados:

- El rango de funcionamiento de la cámara va desde los 0°C hasta los 50°C.
- La cámara requiere un tiempo de calentamiento de aproximadamente 20 minutos para lograr una temperatura de funcionamiento estable.
- La cámara requiere de una temperatura estable. Es importante evitar espacios donde haya gran cantidad de fluctuaciones de temperatura.
- Si es posible mantener la cámara fría a través de un ventilador.

2.3.3.5 Reflectividad de los objetos

La reflectividad de los objetos también influye en la precisión de las medidas. Por ello hay dos aspectos que deben ser considerados:

Tipo de Reflexión

Pueden ocurrir dos tipos de reflexión:

- **Reflexión difusa:** Este tipo de reflexión ocurre en objetos de color mate. La reflexión es uniforme, lo que significa que los pulsos de luz se reflejan igualmente en todos los ángulos. Este tipo de reflexión es preferible respecto a otras, debido a que la intensidad de luz reflejada de nuevo en el sensor no está influenciada por el ángulo.
- **Reflexión especular:** Este tipo de reflexión ocurre en objetos brillantes y materiales transparentes. El pulso de luz reflejado sigue un camino diferente al pulso de luz emitido por la cámara que puede conducir a múltiples reflexiones. Otro de los problemas que pueden existir con este tipo de reflexión es que, la luz reflejada llegue directamente al sensor llegando a saturar la información.

Color del Objeto

Los objetos brillantes proporcionan mejores resultados ya que, en términos generales, pueden reflejar más luz, mientras que con objetos oscuros se absorbe cierta cantidad de luz que no será devuelta al sensor. Otro aspecto importante para considerar es la distancia al objeto. Mientras que los objetos brillantes generalmente producen unos resultados más precisos, si se colocan cerca de la cámara reflejan mucha luz produciendo una sobresaturación.

2.4.1 Características generales de la cámara Basler

La siguiente [Tabla 2.1](#) muestra las características principales que proporciona la cámara Basler.

ESCUELA POLITECNICA
SUPERIOR

Specification	Basler ToF Camera
Resolution (H x V pixels)	640 x 480
Sensor Type	Panasonic MN34902BL
Optical Size	1/4"
Lens	3.6 mm
Field of View (H x V)	57° x 43°
Max. Frame Rate (at full resolution in Standard processing mode)	20 fps
Mono/Color	Mono
Wavelength	850 nm, ± 30 nm
Non-ambiguity Range	0–13.320 m (default channel)
Absolute Accuracy and Repeatability	Absolute Accuracy: ± 1 cm * Repeatability (1 σ): 8 mm *
Drift with Temperature	0.7 mm/K For detailed information, see Section 6.2.3 on page 36 .
External Light Disturbance	above 15 kLux
Communication Interface	Gigabit Ethernet (1000 Mbit/s)
Image Components	<ul style="list-style-type: none"> Range map Intensity image Confidence map
Pixel Formats	<ul style="list-style-type: none"> Coord3D_C16 Available for the range map Mono16 Available for the intensity image Confidence16 Available for the confidence map RGB8 Available for the range map Coord3D_ABC32f (point cloud) Available for the range map
Synchronization	Free run, via hardware or software trigger
Exposure Time Control	Programmable via the camera API, Auto Mode
Camera Power Requirements	24 VDC (± 10 %), supplied via the camera's 12-pin connector
Max Power Consumption	15 W avg 29 W peak (1.2 A @ 24 VDC)
Cooling	Passive, no fan
Enclosure Rating	IP30
Size (L x W x H)	141.9 mm x 61.5mm x 76.4 mm
Weight	≈ 450 g
Lens Adapter	n/a (Integrated lens)
Multi-Camera Setup	Via Multichannel feature or trigger synchronization
Conformity	CE (includes RoHS), GenICam, GigE Vision, FCC The EU Declaration of Conformity is available on the Basler website: www.baslerweb.com
Eye Safety	EN 62471:2008 RG 1
Software	Basler ToF Driver (version 1.1.0 or higher) Available for Windows (x86, x64) and Linux (x86, x64)

Tabla 2.1. Características Basler ToF640-20gm_850nm [3].

2.4.2 Imágenes en profundidad

Cabe destacar un aspecto importante sobre las imágenes en profundidad de la cámara utilizada. Al tener la cámara en posición cenital los valores por profundidad obtenidos se corresponden con el eje Z. El problema es, que dichas medidas obtenidas por la cámara no son medidas métricas, si no, que son adimensionales. Por tanto, tanto en la etapa de re-identificación como en la de entrenamiento, se ha realizado una ecuación que transforma el dato a medidas métricas, que es necesario para el desarrollo de este trabajo. El cálculo de la distancia del píxel se obtiene con la ecuación (2.1).

$$Distancia(mm) = DepthMin_{parameter} + \frac{(PixelValue * (DepthMax_{parameter} - DepthMin_{parameter}))}{65535} \quad (2.1)$$

Donde los parámetros:

- **DepthMax_parameter:** Tiene un valor de 13320.
- **DepthMin_parameter:** Tiene un valor de 0.
- **Valor del Píxel:** Depende de la distancia.

2.5 Ventajas que ofrece la cámara Basler respecto a otras

En la actualidad se utilizan diferentes técnicas de modulación para las cámaras ToF. Cada una de las técnicas aportan una serie de ventajas respecto a las otras, en este apartado se exponen las ventajas existentes en el uso de la técnica que emplean las cámaras Basler ToF.

La cámara Basler ToF utiliza una modulación pulsada (PM) [4], esta técnica aporta las siguientes ventajas:

- Disminuye la influencia de la iluminación del fondo.
- Alta relación Señal/Ruido.
- Permite realizar medidas a larga distancia.

2.2 Análisis de componentes principales (PCA)

El análisis de componentes principales es un método estadístico que permite simplificar la complejidad de espacios con muchas dimensiones a la vez que conserva su información, consiguiéndose encontrar una serie de patrones o características dentro de un conjunto de datos.

2.2.1 Procedimiento matemático

Este procedimiento requiere una etapa de entrenamiento en la cual se obtiene el modelo o clase de un conjunto de datos.

Se parte de un conjunto de N vectores de dimensión P , a los cuales se les quieren extraer las características comunes. Estos datos se agrupan en la matriz T :

$$T = \begin{pmatrix} X_{1,1} & \cdots & X_{N,1} \\ \vdots & \ddots & \vdots \\ X_{1,P} & \cdots & X_{N,P} \end{pmatrix} \quad (2.2)$$

Una vez calculada la matriz T , se calcula la media como se observa en 2.3.

$$\mu = \frac{1}{N} \sum_{n=1}^N X_n = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_P \end{pmatrix} \quad (2.3)$$

A continuación, se obtiene una matriz cuyas columnas se corresponden con las desviaciones de la media de los vectores X_n , se puede ver a continuación en 2.4.

$$B = X - h\mu^T, h[i] = 1, i = 1, \dots, N \quad (2.4)$$

A partir de B se obtiene la matriz de covarianza. Si los datos son reales, la matriz de covarianza se puede calcular como 2.5:

$$C = \frac{1}{N} B * B^T = \begin{pmatrix} cov(X_1, X_1) & \cdots & cov(X_1, X_p) \\ \vdots & \ddots & \vdots \\ cov(X_p, X_1) & \cdots & cov(X_p, X_p) \end{pmatrix} \quad (2.5)$$

Donde $cov(X_1, X_p)$ es la covarianza entre las componentes X_i y X_j .

Como último paso, se obtienen los autovalores de la matriz de covarianza C , en la cual se busca la matriz V formada por los autovectores que diagonaliza la matriz de covarianza C :

$$V^{-1}CV = D \quad (2.6)$$

donde D es la matriz diagonal compuesta por los autovalores de C .

Una vez calculados los autovalores se crea la matriz de transformación U , que permite el cambio de sistema de coordenadas original al creado por PCA. Para conseguirlo, en las columnas de U se colocan los autovectores ($eig_m, m = 1, \dots, P$) en dependencia del autovalor asociado, de siguiendo el orden de mayor a menor y. La matriz U también no

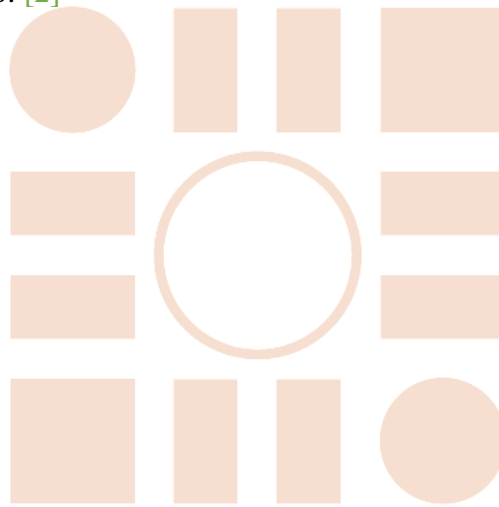
permite reducir el número de dimensiones de los datos, por, la matriz U se forma con el número de autovectores m asociados a los autovalores de mayor valor.

$$U = (eig_1, \dots, eig_m) \quad (2.7)$$

De este modo, la matriz U permite proyectar cualquier vector v sobre el nuevo espacio transformado PCA mediante 2.8, obteniendo el vector proyectado v_T .

$$v_T = U^T * (v - \mu) \quad (2.8)$$

La matriz de transformación (U), y el vector medio (μ) definen el modelo de cada una de las personas a re-identificar, por tanto, cada vez que entra una persona se debe crear y almacenar dicho modelo. [2]



ESCUELA POLITECNICA
SUPERIOR

3. Desarrollo y Diseño

3.1 Introducción

En este capítulo se aborda lo concerniente al desarrollo del sistema. Primero se establece una visión general del mismo, para así posteriormente ir detallando cada bloque/función desarrollada.

En el sistema se distinguen tres bloques principales, el primero de ellos hace referencia a la adquisición de los datos con la cámara ToF. El segundo contiene la parte del cierre de dicha cámara y a la finalización del programa. Y el tercer y último bloque se corresponde con el análisis de las secuencias tomadas o guardadas en la base de datos. Para ello, se ha desarrollado una interfaz de usuario para facilitar el manejo y estudio del sistema, en el Anexo A se explica con más detalle el funcionamiento del interfaz de usuario creado. La [Figura 3.1](#) muestra los bloques del sistema.

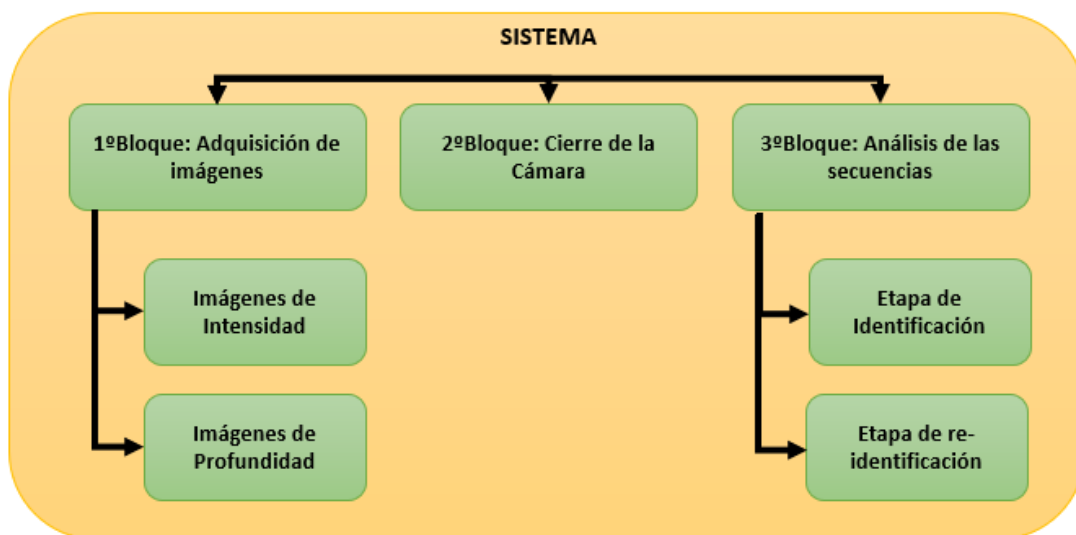


Figura 3.1. Bloques del sistema.

Como se observa en la [Figura 3.1](#) los tres bloques principales del sistema están interconectados entre sí, pudiendo acceder a cada uno de ellos cuando se requiera, facilitando una vez más el manejo de esta interfaz. A continuación, se procede al estudio de cada uno de estos bloques y sub-bloques.

3.2 Bloque de adquisición de imágenes

Este bloque realiza el proceso de adquisición de las imágenes y el posterior almacenamiento de ellas en la base de datos generada para este trabajo. Como se puede ver en la [Figura 3.2](#) se representa el diagrama de bloques. Donde primeramente se realiza la creación de dos ficheros, un fichero referente a intensidad y el otro a profundidad. Cuando se ha realizado la creación de estos, se almacenan en ellos las respectivas muestras de intensidad y profundidad capturadas por la cámara. Para concluir, estos ficheros son guardados en la base de datos que se ha creado para este TFG.



Figura 3.2. Proceso de adquisición de las imágenes.

En la [Figura 3.3](#) se muestran ejemplos de capturas de las imágenes de profundidad adquiridas por la cámara Kinect, mientras que la [Figura 3.4](#) se muestra las imágenes correspondientes de intensidad.

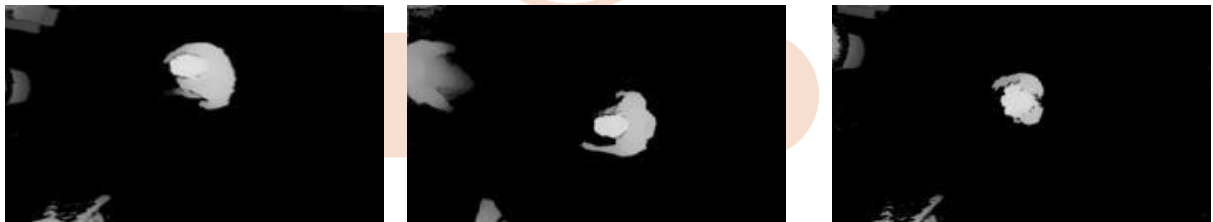


Figura 3.3. Imágenes de Profundidad.



Figura 3.4. Imágenes de Intensidad.

3.3 Bloque Cierre de la cámara

La finalidad de este bloque consiste en el cierre de la cámara y la finalización del programa a través del interfaz creado, el usuario puede acceder a esta funcionalidad cuando lo requiera.

3.4 Bloque Análisis de secuencias

Este bloque es el encargado de realizar:

- Etapa de Identificación.
- Etapa de Re-identificación

A continuación, se detalla en profundidad cada una de las dos etapas.

3.4.1 Etapa de Identificación

Para esta etapa se ha hecho uso del detector desarrollado por [12], [13] y [1]. A continuación, se detalla el funcionamiento de dicho detector. La Figura 3.5 muestra los diferentes bloques del mismo.

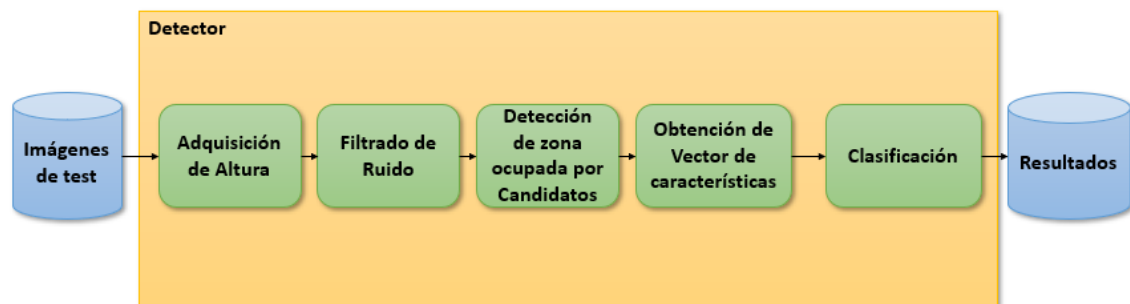


Figura 3.5. Diagrama general de bloques del detector [1].

3.4.1 Secuencias de Profundidad

3.4.1.1 Adquisición de Altura

Una vez adquiridas las imágenes de intensidad y profundidad respectivamente, se procede al uso del detector. La primera función que va a desempeñar es el cálculo de la altura. Dado que la cámara ToF se encuentra situada en una posición cenital, su eje óptico es perpendicular al plano del suelo. Por tanto, la altura de la persona es la diferencia entre la altura de la cámara respecto al suelo y la altura medida (3.1).

$$H_{m,n} = H_{camara} - Z_{m,n} \quad (3.1)$$

En la Figura 3.6 se puede observar el sistema de coordenadas definido.

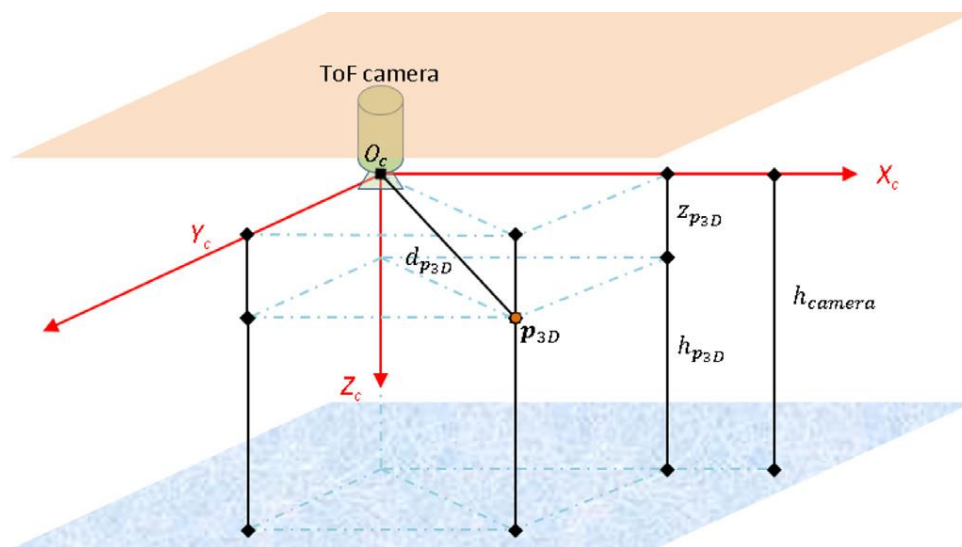


Figura 3.6. Esquema de la posición de la cámara [5].

3.4.1.2 Reducción de Ruido

Uno de los problemas fundamentales de la cámara ToF es el alto ruido que presenta en las matrices medidas. Este ruido es especialmente significativo si hay objetos en movimiento en la escena, lo que produce gran cantidad de medidas no válidas. Para reducir el ruido y las de mediciones no válidas en las imágenes de profundidad, se ha implementado un algoritmo de reducción de ruido que incluye dos etapas. En la primera, se hace una estimación, mediante interpolación de los píxeles vecinos, de los píxeles que la cámara proporciona como mediciones no válidas. En la segunda etapa se procede a la eliminación del ruido en la secuencia captada utilizando un filtro de media.

Esta eliminación de ruido solo se ha implementado para las imágenes de profundidad, ya que este error no afecta a los valores de intensidad.

3.4.1.3 Estimación de Regiones de Interés (ROIs)

Para nuestro sistema se ha establecido unas regiones de interés, ya que estas son las que proporcionan toda la información válida para el proceso de identificación y re-identificación. Por ello, están formadas por la cabeza, el cuello y los hombros del sujeto a analizar estableciéndose que la diferencia entre la cabeza y los hombros no debe ser mayor que una altura, nombrada en nuestro caso $h_{interés}$, dando un valor de 40cm en base a consideraciones antropométricas [6]. En la Figura 3.7 se representa un esquema de la altura de interés.

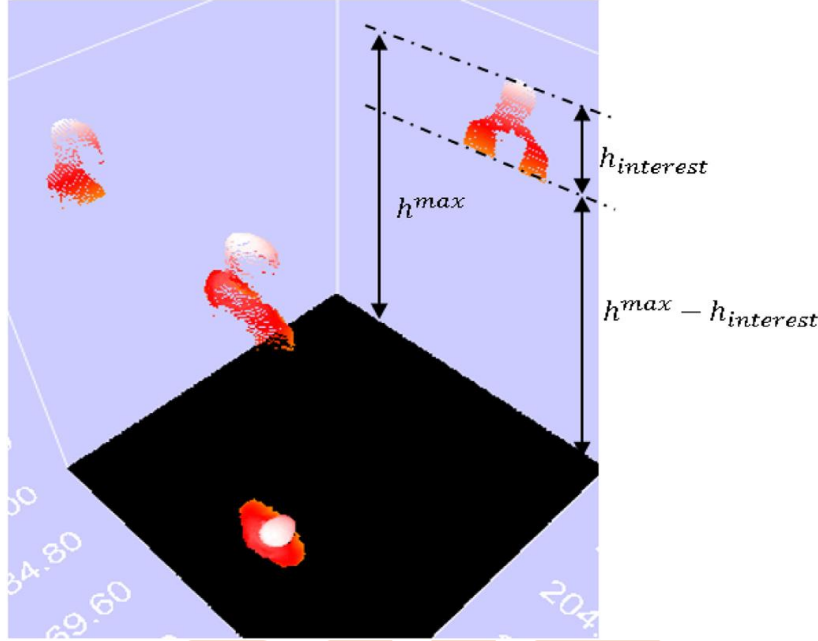


Figura 3.7. Esquema de altura de interés [5].

3.4.1.3.1 Detección de máximos locales

Este sub-bloque se utiliza para comprobar cuáles de las regiones de la matriz hallada puede corresponder a personas. Primero se realiza una división del plano de la imagen en subregiones con unas dimensiones de $\mathbf{D} \times \mathbf{D}$ píxeles, de este modo el número total de estas queda de la siguiente forma $\mathbf{N}_r \times \mathbf{N}_c$, siendo:

$$N_r = \frac{M}{D} \quad (3.2)$$

$$N_c = \frac{N}{D} \quad (3.3)$$

Donde:

- M: Alto de la imagen, igual a 424
- N: Ancho de la imagen igual a 512
- D: Dimensión de la Subregión, igual a 20

Después de realizar el cálculo de las subregiones y el tamaño de cada una de ellas, se produce el cálculo de los máximos de cada uno de las SR.

$$\mathbf{H}^{maxSR} = \begin{bmatrix} h_{1,1}^{maxSR} & \dots & h_{1,N_c}^{maxSR} \\ \vdots & \ddots & \vdots \\ h_{N_r,1}^{maxSR} & \dots & h_{N_r,N_c}^{maxSR} \end{bmatrix}$$

Cada valor contemplado en la matriz \mathbf{H}^{maxSR} es considerado como un posible candidato para ser una persona, se considera como válido si se supera la altura mínima ($h_{pmin} = 1m$) y además es mayor o igual que los máximos de las subregiones contiguas. Si se

cumplen las condiciones nombradas anteriormente, se considera que se ha detectado a una persona. Dado que es muy probable que dos SR's vecinas pertenezcan al mismo sujeto, serán sustituidos por el que tenga el valor más alto que representará a los otros.

3.4.1.3.2 Ajuste de las regiones de Interés

Para poder determinar las subregiones que pertenecen a la ROI, se ha diseñado un algoritmo capaz de diferenciar a una persona de su entorno [1], a través de una serie de condiciones. Si estas condiciones se cumplen las subregiones analizadas se agregan a la ROI. Para seleccionar las SR's que pertenecen a la ROI, se establecen 8 niveles radiales de área de vecindad local y 8 direcciones radiales alrededor de cada SR. Las condiciones que deben cumplir las áreas vecinas deben ser:

- El máximo analizado debe encontrarse dentro de la región de interés.
- El valor del máximo asociado a la SR debe ser menor que el de la SR anterior en esa misma dirección y mayor que el de la siguiente SR, con el fin de poder separar a personas que se encuentren próximas en la imagen [1].

3.4.1.4 Extracción de Características

Una vez calculada la ROI, el objetivo es poder determinar la presencia de una persona. Por tanto, se ha diseñado un algoritmo, que es capaz de extraer un conjunto de características en base a los valores hallados del sujeto detectado. Estas características son almacenadas en un vector, que a partir de ahora lo referenciaremos como “vector de características”, el cual contiene información de atributos físicos de las personas y debe permitir discriminar entre diferentes individuos.

La creación del vector de características es realizada con toda la información que aportan las imágenes de profundidad. Este vector tiene un total de 23 componentes $\theta = (\theta_1, \theta_2, \dots, \theta_{23})$ relacionadas con la fisionomía. La primera de todos los componentes hace referencia a la altura de la persona detectada (mm). Las siguientes 18 corresponden al número de puntos en diferentes franjas de 2 cm en las que se dividen los puntos incluidos en la ROI, la Figura 3.8 muestra la distribución de las franjas de 2cm. [5].

Es posible que encontremos gran cantidad de ruido introducido por el sensor, para ello se ha aplicado un algoritmo para la reducción del mismo, éste se asegura que la primera franja introducida en el vector pertenece realmente a la cabeza de la persona. De modo que se ha determinado cuál de las tres primeras componentes tiene un mayor número de medidas.

$$S_{h_{max}} = \arg \max(\theta_1, \theta_2, \theta_3) \quad (3.4)$$

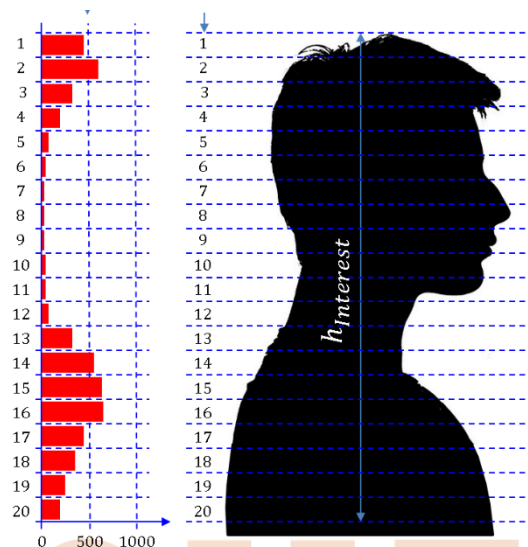


Figura 3.8. División en secciones [1].

Tras determinar cuál de los tres primeros componentes corresponde a la parte superior de la cabeza, se rellena el vector con los siguientes 18 restantes, dándose así un vector de 19 componentes [5]. Para reforzar la información y obtener unos mejores resultados se han añadido una serie de características extras formando un vector de 23 componentes, estas son, la altura de los hombros al vector, datos sobre la forma de la cabeza, como la medida del eje mayor de la elipse que la conforma y la relación entre el eje mayor y el menor de la misma, como los hombros tienen una forma parecida se ha intentado obtener las mismas características. La Figura 3.9 muestra un ejemplo de algunos de los componentes extras añadidos, en concreto se corresponden con las medidas del eje mayor y menor tanto para la cabeza como para los hombros.

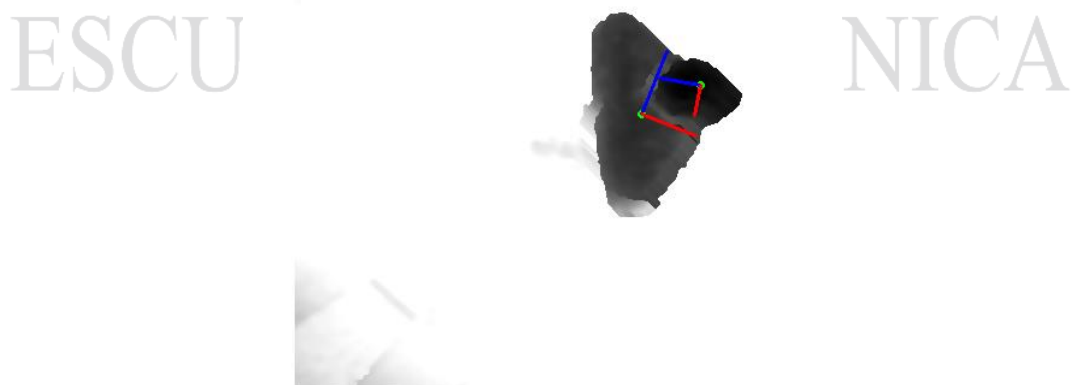


Figura 3.9. Ejemplo de obtención de las medidas para la cabeza y los hombros.

3.4.2 Secuencias de Intensidad

Finalizada la explicación para la extracción del vector de características para las imágenes de profundidad, se procede a la explicación sobre la extracción de los componentes referidos a las imágenes de intensidad.

Para cada sección se almacenan en un vector los valores de intensidad de cada pixel perteneciente a esa franja de alturas. De este modo, se obtienen un total de 20 vectores, uno por cada sección. El tratamiento de estos vectores se hace a través de histogramas, en concreto se crean 20 de estos, uno para cada sección de la persona. Para el cálculo de estos, se ha usado la función CalcHist() de la herramienta OpenCV [8].

La Figura 3.10 representa el histograma hallado para los valores de intensidad correspondientes a la sección 0, mientras que en la Figura 3.11 se representa el histograma con todos los valores de intensidad para todas las secciones* [11] .

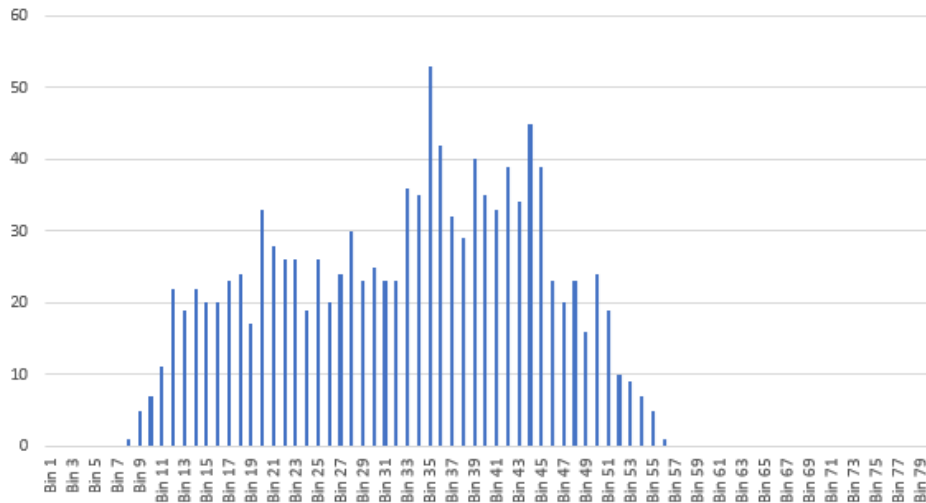


Figura 3.10. Histograma de intensidad de la Seq-DFJ-D02.

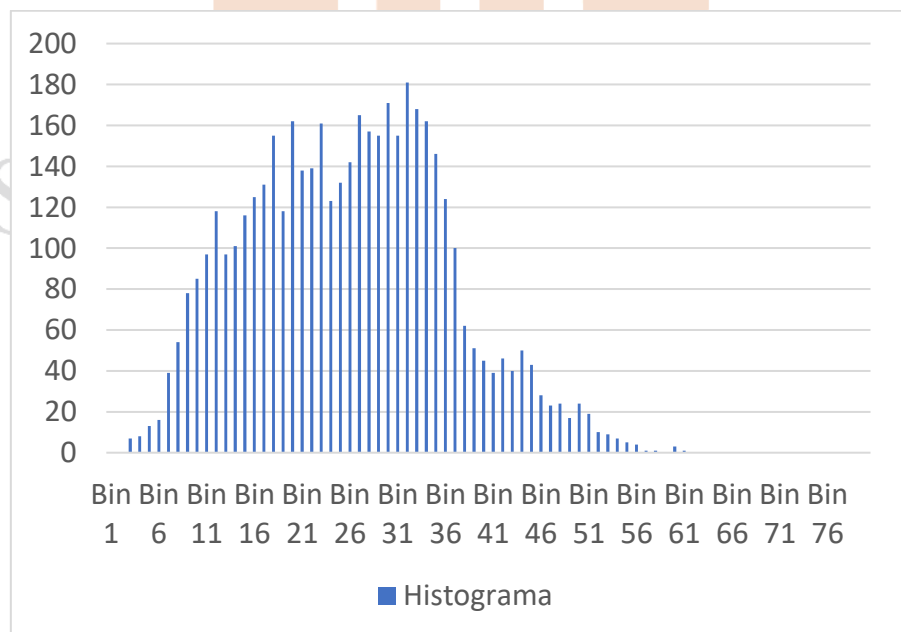


Figura 3.11. Histograma de la Seq-DFJ-D02 de todas las secciones.

*bin= hace referencia al número de subdivisiones que se encuentran en el histograma, cada bin tiene un rango de valores, por ejemplo, el primer bin comprende valores desde 0 hasta 100 y así respectivamente. De modo que se han establecido 80 subdivisiones para el cálculo del histograma

En las siguientes imágenes se muestra una pequeña comparativa entre histogramas, en la [Figura 3.12](#) se puede ver como ambos histogramas son bastante parecidos, esto es debido a que pertenecen a la misma persona realizando sentidos contrarios. Mientras que en la [Figura 3.13](#), se observa una mayor diferencia entre estos, ya que se corresponden a personas diferentes, de aquí la variación entre uno y otro. Con esta pequeña comparativa es fácil ver el potencial en la identificación de personas utilizando esta técnica para evaluar los valores de intensidad obtenidos por cada individuo, esta técnica permite diferenciar en la mayoría de los casos entre diferentes sujetos, aportando una información útil para el proceso de re-identificación.

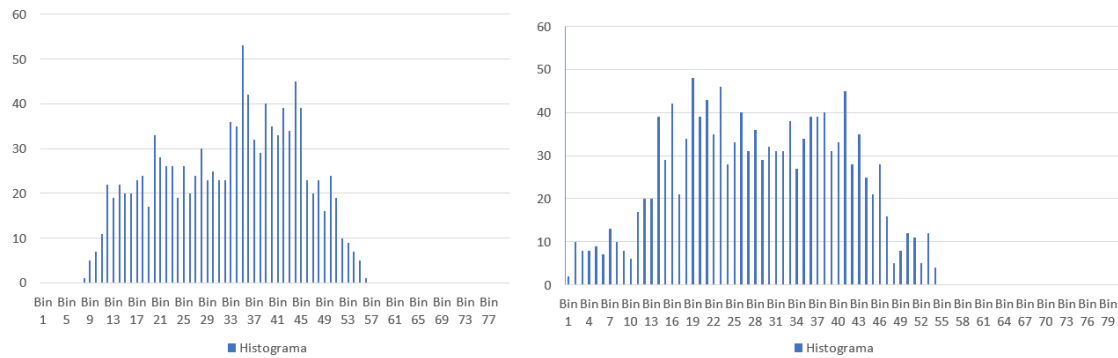


Figura 3.12. Comparativa de histogramas, (a) Histograma correspondiente a seq-DFJ-D02; (b) Histograma correspondiente a seq-DFJ-D01.

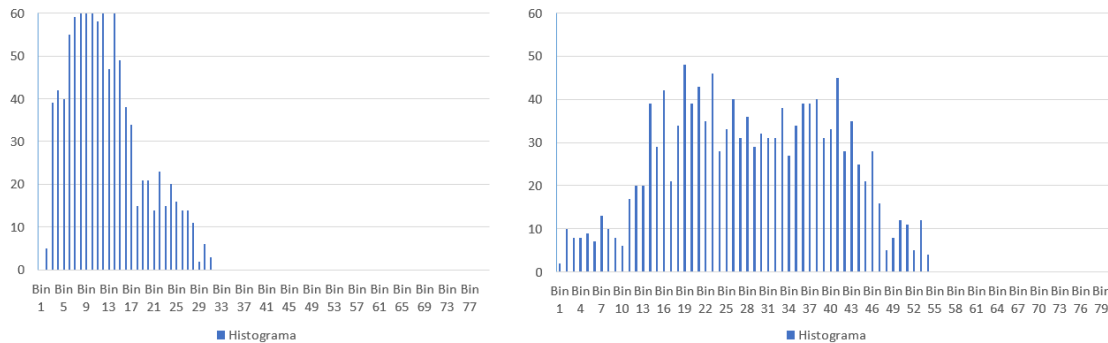


Figura 3.13. Comparativa de histogramas; (a) Histograma correspondientes a seq-MSP-D02; (b) Histograma correspondiente a seq-DFJ-D01.

A partir de los valores máximos de cada histograma, se obtienen los componentes de intensidad del vector de características. Dicho vector finalmente queda compuesto por 43 componentes, las primeras 23 se sacan de las imágenes de profundidad, como ya se ha comentado anteriormente Y los 20 restantes se corresponden con los valores de intensidad hallados mediante los histogramas.

Una vez creado el nuevo vector de características se procede al cálculo de la PCA, para posteriormente pasar a la etapa de re-identificación.

3.4.4 Etapa de Re-Identificación

La re-identificación utiliza un clasificador basado en PCA. Por ello, el vector asociado a cada nueva persona detectada se compara con cada una de las clases previamente creadas y almacenadas en el proceso de identificación. Para realizar la comparación, en primer lugar se proyecta el nuevo vector a clasificar (α) al espacio transformado PCA de dimensión inferior (ecuación (3.5)), y posteriormente se recupera al espacio original (ecuación (3.6)).

$$\alpha_p = \mathbf{U}^T(\alpha - \mu) \quad (3.5)$$

$$\alpha_r = \mathbf{U} \cdot \alpha_p + \mu \quad (3.6)$$

En caso de que el vector de entrada sea similar a los utilizados para crear la clase, el error de recuperación (definido como la distancia euclídea entre el vector original α y el recuperado α_r) debe ser pequeño. Así, se considera que la nueva persona pertenece a la clase para la cual el error de recuperación sea menor, asignándole la identidad correspondiente.

El número de componentes seleccionadas (m) influye de forma significativa en el error de recuperación. Por este motivo, en el presente trabajo se ha realizado un estudio experimental, por lo que los resultados mostrados se han obtenido considerando únicamente 3 componentes.

ESCUELA POLITECNICA
SUPERIOR

4. Resultados

4.1 Introducción

En este capítulo se detallan los resultados obtenidos tras la realización de este trabajo, el cual, se puede separar en tres partes. La primera parte correspondiente a la cámara Basler Tof, la segunda parte referente con la implementación de la técnica de PCA una vez incluida la cámara Kinect 2. Por último, la tercera parte se ven los resultados obtenidos a través de la técnica de distancia euclídea implementada.

4.2 Base de datos de secuencias

Para evaluar el funcionamiento de las diferentes alternativas, se ha creado una base de datos con secuencias de personas caminando en una dirección y en la dirección contraria, simulando la entrada y salida a un local. Dichas secuencias han sido grabadas en el Espacio Inteligente GEINTRA de la Universidad de Alcalá.

A continuación, se muestra en la [Tabla 4.1](#) una lista con todas las secuencias almacenadas para la realización de este trabajo. A estas secuencias se les ha asignado un código de identificación (**PX**) para cada una de ellas, para así, facilitar la muestra de los resultados.

Secuencia	Sexo	Código Secuencia
Seq-MSP	Hombre	P1
Seq-CLV	Hombre	P2
Seq-ISP	Mujer	P3
Seq-JVD	Hombre	P4
Seq-TOP	Hombre	P5
Seq-BAP	Hombre	P6
Seq-JTM	Hombre	P7
Seq-ADC	Hombre	P8
Seq-AMC	Mujer	P9
Seq-DCP	Hombre	P10
Seq-DFJ	Hombre	P11
Seq-FST	Hombre	P12
Seq-ADR	Hombre	P13
Seq-NRS	Mujer	P14
Seq-ARM	Hombre	P15
Seq-CTT	Mujer	P16
Seq-MLT	Mujer	P17
Seq-PHJ	Mujer	P18
Seq-EGR	Hombre	P19
Seq-MSR	Mujer	P20
Seq-JSD	Hombre	P21

Tabla 4.1. Lista de secuencias almacenadas en la base de datos.

Encontramos un total de 21 secuencias almacenadas en la base de datos.

Para cada persona encontramos dos secuencias, una se corresponde a la entrada y otra a la salida. La nomenclatura utilizada es: **Seq-XXX-D0Y**:

- **XXX**: Hace referencia a las iniciales de cada una de la persona evaluada.
- **D0Y**: Hace referencia al sentido del movimiento de la persona, donde, Y puede tener un valor de 0 o 1, dependiendo si la secuencia se corresponde a la entrada o a la salida.

4.3 Resultados Cámara Basler

En un primer momento, como ya se ha visto a lo largo de esta memoria, se optó por la utilización de esta cámara para el desarrollo del proyecto. No obstante, durante la implementación del mismo, se comprobó que el funcionamiento de esta cámara en condiciones de entorno reales, dista mucho de las especificaciones proporcionadas por el fabricante. Mediante experimentos, se determinó que solamente mide correctamente la profundidad de los objetos con elevada reflectividad, esto se traduce en una disminución considerable de la altura de la persona, debido a la baja reflectividad del pelo. También, se observa una gran cantidad de píxeles erróneos sobre esa zona. En la [Tabla 4.2](#) se muestran ejemplos de las medidas de la altura con dicha cámara,

Secuencia	Medida Obtenida	Medida Real
Seq-CLV	162cm	182cm
Seq-DFJ	158cm	174cm
Seq-MSP	155cm	172cm
Seq-ISP	132cm	155cm

Tabla 4.2. Comparativa entre las alturas reales de las personas y las obtenidas con la cámara Basler.

Como se puede ver en la [Tabla 4.2](#), los resultados obtenidos distan mucho de las medidas reales. Imposibilitando la utilización de la cámara Basler para el desarrollo final de este trabajo. Por ello, se decidió el uso de la cámara Kinect 2, que aportaba mayor fiabilidad y ya había sido utilizada en diferentes proyectos desarrollados por el grupo de investigación GEINTRA.

4.4 Resultados Obtenidos aplicando la técnica de PCA

En este apartado se muestran los resultados obtenidos utilizando la técnica PCA implementada con el vector de 43 componentes, donde se incluye la información de profundidad e intensidad. Se comprobó que los resultados eran muy similares a los obtenidos en [1]. En nuestro caso, de las 21 secuencias, se re-identificó correctamente a 16 personas, mientras que los 5 restantes se confundían con otras personas. La Figura 4.1 muestra un gráfico con los resultados obtenidos empleando la técnica de PCA.



Figura 4.1. Gráfico de los resultados obtenidos con el uso de la PCA.

Estos resultados corroboran que con la implementación de valores de intensidad utilizando la técnica PCA no se produce ninguna mejoría con respecto a los resultados obtenidos en [1]. La no mejora puede ser por el escaso número de muestras que se han podido almacenar en la creación de la PCA. Otra de las posibles causas ha sido la no implementación de todo el histograma, es decir, solo se ha almacenado en el vector de características los valores máximos de cada histograma, en vez de guardar todos los valores de cada uno de estos, lo que proporciona una mayor cantidad de información extra para mejorar los resultados.

4.5 Resultados finales obtenidos.

En este apartado se muestran las modificaciones realizadas al procedimiento inicialmente propuesto, con el fin de mejorar los resultados de la re-identificación.

La primera modificación consiste en la creación de solo 2 histogramas de intensidad, cuando la persona de la secuencia se encuentra en el centro de la escena, [Figura 4.2](#) por ello sólo se coge una única imagen de profundidad e intensidad. No se realizan modificaciones en la extracción de los 23 componentes de profundidad.

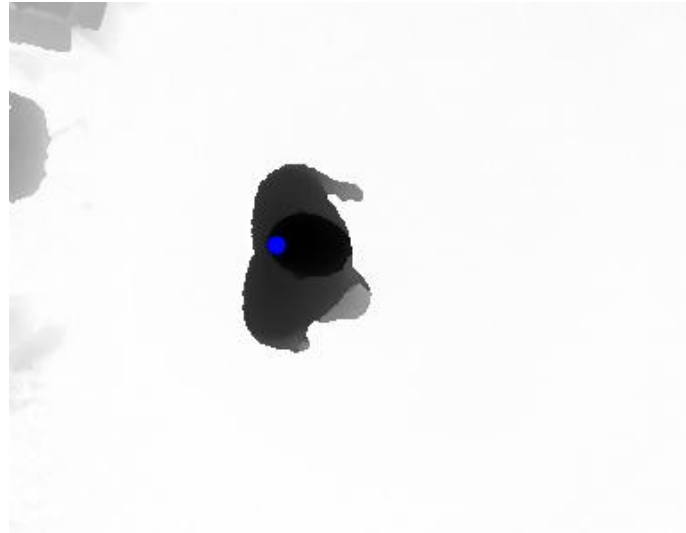


Figura 4.2. Punto donde se extraen los valores.

Sólo se ha necesitado el uso de dos histogramas de intensidad, debido a que, como se muestra en la [Figura 3.8](#), las secciones intermedias proporcionan pocos valores válidos, de modo que, el primer histograma está compuesto por las 5 primeras secciones (aproximadamente hasta los ojos). Y el segundo por las 5 últimas (referente a la parte de los hombros), que son las que nos proporcionan mayor cantidad de información. Los histogramas generados siguen la misma estructura utilizada que en el Capítulo 4.4.

El procedimiento de clasificación se basa en determinar la diferencia que existe entre los vectores de profundidad y los histogramas de intensidad de la persona que sale y los de cada una de las personas que han entrado. Esta diferencia se calcula mediante la distancia euclídea.

El error de profundidad se normaliza a un rango de 0 a 3. En el caso de los errores de intensidad, se calculan por separado para cada histograma. Para finalizar se procede al cálculo del error total, este error sigue la siguiente distribución:

$$Error\ Total = Pr1 * ErrorProfundidad + Pr2 * ErrorHistograma2 + Pr3 * ErrorHistograma1 \quad (4.3)$$

Donde:

- Pr: Indica el grado de prioridad definido para cada error.

Se ha hecho uso de un control de prioridades debido a que los resultados obtenidos sin este control son peores de los esperados.

La estructura que se muestra en las siguientes tablas de resultados es [Tabla 4.3]:

	P_j	P_{j+1}	..	P₂₁
P_i	$E_{(i,j)}$	$E_{(i,j+1)}$..	$E_{(i,21)}$
P_{i+1}	$E_{(i+1,j)}$
..
P₂₁	$E_{(21,21)}$

Tabla 4.3. Estructura de las tablas de resultados.

- **P_j**: Muestra las diferentes secuencias almacenadas a comparar con la secuencia escogida (P_i).
- **P_i**: Muestra la secuencia a comparar con las secuencias de almacenadas (P_j).
- $E_{(i,j)}$: Muestra el error correspondiente entre la secuencia a comparar con la correspondiente secuencia almacenada

En la Tabla 4.4, se muestran los resultados, indicando las secuencias correctas y las secuencias erróneas, sin aplicar ningún tipo de prioridad.

ESCUELA POLITECNICA
SUPERIOR

	P 1	P 2	P3	P4	P5	P6	P7	P8	P9	P 10	P 11	P 12	P 13	P 14	P 15	P 16	P 17	P 18	P 19	P 20	P 21
P 1	2.98	7.72	6.33	8.70	3.55	7.52	8.15	3.62	5.23	9.50	6.73	6.26	5.76	7.72	7.54	5.38	7.04	7.55	7.69	6.88	4.74
P 2	8.19	7.9	10.6	7.7	8.85	6.38	7.46	8.99	8.03	12.3	10.5	8.09	12.4	12.1	11.4	9.22	10.6	11.1	11.9	9.72	7.09
P 3	5.77	7.64	2.53	10.0	4.53	6.76	6.56	4.51	4.42	9.08	6.51	6.38	6.42	4.94	7.22	4.16	4.82	5.46	5.60	4.82	6.40
P 4	7.79	6.11	8.83	6.48	8.50	5.74	6.63	7.25	7.07	10.4	10.6	4.87	10.4	9.38	9.50	6.82	7.81	7.98	8.83	7.07	5.59
P 5	2.91	8.12	6.94	9.35	3.16	7.46	8.10	4.21	5.43	9.70	7.29	6.95	5.24	8.33	8.51	5.46	7.45	7.86	8.14	6.72	5.63
P 6	7.39	7.65	7.55	7.02	7.9	4.75	4.60	7.87	5.88	9.75	7.95	7.34	9.75	9.46	10.1	6.25	7.92	8.62	9.19	6.87	6.72
P 7	7.21	8.32	7.18	7.11	6.67	5.10	2.99	7.71	5.33	8.81	6.93	7.53	9.58	9.13	10.0	6.34	7.59	8.18	8.79	6.69	6.41
P 8	3.92	7.72	6.68	8.68	4.24	6.95	7.4	3.0	5.57	9.43	7.51	5.72	6.63	7.90	7.92	5.51	6.81	7.36	7.82	6.92	4.74
P 9	4.99	8.03	6.57	9.16	5.09	6.20	7.20	5.34	4.02	9.33	8.04	6.21	7.12	8.47	8.34	5.56	7.13	8.68	8.08	6.35	4.47
P 10	8.58	5.27	7.34	7.87	8.18	5.47	6.18	7.29	6.28	8.08	8.09	4.63	10.7	7.23	7.81	6.61	6.25	7.80	6.84	5.32	7.04
P 11	7.25	8.76	7.12	11.5	6.86	6.52	7.33	7.58	6.84	9.01	2.29	8.97	9.39	9.04	9.27	6.59	8.37	9.76	9.2	7.68	8.94
P 12	6.01	6.24	7.16	7.75	6.46	5.24	5.79	5.12	5.15	9.36	9.31	4.00	8.74	8.14	8.10	5.54	6.67	6.48	7.71	6.22	3.50
P 13	5.43	10.3	6.13	11.0	4.95	9.62	9.41	5.59	6.98	10.3	8.20	8.63	2.87	6.84	9.76	6.03	6.34	5.67	6.44	6.99	7.8
P 14	7.45	6.51	4.48	8.91	6.25	7.39	7.31	5.61	5.38	9.67	8.85	5.52	6.95	4.26	7.50	4.36	4.46	4.68	4.68	3.62	6.80
P 15	6.86	4.28	5.71	6.15	6.52	6.95	7.56	5.43	5.69	9.19	9.08	4.04	9.76	5.22	5.06	6.17	5.32	7.04	5.49	4.98	5.45
P 16	5.21	5.97	4.26	8.88	4.40	4.85	5.43	4.24	3.70	8.51	6.44	4.98	6.99	6.50	7.61	1.90	4.86	5.60	6.50	3.98	5.17
P 17	6.68	6.31	4.20	8.43	5.79	5.84	5.80	4.86	4.55	8.57	7.73	4.83	6.96	5.09	7.37	3.50	2.14	3.91	4.79	3.56	5.65
P 18	6.57	8.06	4.61	9.15	5.73	7.15	6.72	4.99	5.19	8.97	8.18	6.15	4.73	5.32	8.55	4.07	4.41	2.59	4.26	4.40	6.01
P 19	7.33	6.55	4.78	8.95	6.31	6.84	6.79	5.06	5.28	9.27	8.52	4.97	6.37	5.05	7.56	4.10	3.7	3.74	4.19	4.19	6.38
P 20	4.08	8.00	6.56	10.4	3.93	7.20	7.28	5.58	4.05	10.4	7.99	7.16	7.55	8.54	7.93	6.78	7.69	8.23	8.63	7.02	5.09
P 21	4.57	6.75	5.88	8.69	4.48	5.81	6.13	4.11	3.35	9.58	8.63	4.76	7.35	7.51	7.22	5.40	5.94	6.10	7.28	5.49	2.61

Tabla 4.4.- Tabla de aciertos sin aplicar las prioridades para las secuencias.

Como se ha visto en la [Tabla 4.4](#) anterior, los resultados sin aplicar ningún tipo de prioridad no son buenos, debido a que se obtiene una probabilidad de acierto menor que la aplicada con la técnica de PCA. Esto es producido a que en algunas secuencias se obtienen resultados erróneos en profundidad, en intensidad o en ambas. Los malos resultados en profundidad suelen estar relacionados con la diferente posición de la cabeza, es decir, la persona para una dirección se puede encontrar completamente erguida, mientras que para la otra dirección puede tener la cabeza agachada. En cuanto a los malos resultados para intensidad, vienen derivados principalmente del primer histograma, que

hace referencia a las 5 primeras secciones que están compuestas principalmente por el pelo que es una fuente de error, dependiendo de:

- Si la persona lleva tinte de pelo o no, el tinte de pelo es un factor que influye en gran medida para los valores de intensidad, debido a que el uso de estos productos químicos altera el cabello produciendo gran variabilidad de valores de intensidad y produciendo reflexiones desiguales entre las diferentes partes del pelo.
- El tipo de pelo es otro de los factores para tener en cuenta, a medida que se ha realizado la evaluación de los resultados se ha corroborado la influencia de este en los valores de intensidad. Una persona con el pelo más fino produce mejores resultados que una persona de cabello grueso como también el color de pelo influye en las medidas de intensidad.
- El tipo de peinado es otro de los factores que influyen en los resultados. Una persona con el pelo rizado obtiene mejores resultados que esa misma persona con el cabello liso debido a que las reflexiones producidas no son las mismas.

La [Figura 4.3](#) muestra una imagen de intensidad donde se puede observar que existen diferentes valores de intensidad en la parte del pelo.

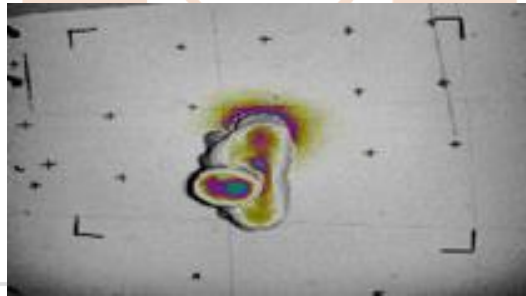


Figura 4.3. Imagen de intensidad representando los diferentes valores de intensidad.

Otra de las técnicas que se ha realizado para este trabajo, ha sido la creación de 6 histogramas de intensidad. Cada uno de éstos está comprendido por 3 secciones, las cuales pertenecen a la región de interés. El primer histograma está compuesto de las 3 primeras secciones de la ROI, y así hasta completar el 6 y último histograma que abarca desde la franja 16 hasta la 18. Con esta técnica se han desechado las dos últimas secciones de la región de interés para contener la misma información en cada uno de los histogramas creados. También se optó por eliminar los componentes extras de profundidad, debido a que en casos puntuales daban malos resultados, por tanto, se optó por eliminarlas para evaluar los resultados obtenidos, exceptuando la altura de la persona, dado que es un parámetro primordial que proporciona mucha información a la hora de realizar el análisis.

En la [Tabla 4.5](#), se muestran los resultados obtenidos aplicando la técnica de creación de 6 histogramas de intensidad y eliminación de los componentes extras, obteniéndose 12 secuencias correctas de 21, lo que se traduce en un porcentaje de acierto del 57%,

	P 1	P 2	P3	P4	P5	P6	P7	P8	P9	P 10	P 11	P 12	P 13	P 14	P 15	P 16	P 17	P 18	P 19	P 20	P 21
P 1	4.70	10.9	7.66	11.8	6.47	10.1	9.87	4.35	7.67	12.6	9.77	7.66	11.2	9.74	8.87	8.19	9.79	11.2	11.7	10.7	6.13
P 2	9.70	10.4	12.5	15.0	11.2	8.80	9.87	11.2	10.6	16.6	15.2	10.8	17.4	15.0	13.7	12.2	13.3	14.2	16.2	12.6	9.45
P 3	6.56	9.31	4.07	11.8	6.71	7.97	7.82	5.12	6.02	11.1	8.66	6.90	11.4	6.59	7.97	5.78	7.06	8.88	8.74	7.06	6.98
P 4	10.4	8.93	10.4	7.95	12.2	9.19	9.33	8.91	10.2	13.8	15.0	6.51	16.2	10.4	11.1	9.23	9.51	10.5	11.1	10.9	7.36
P 5	3.85	10.4	8.33	14.1	3.63	8.42	9.17	6.11	5.97	14.2	10.3	9.85	11.3	11.3	10.3	8.68	10.4	12.1	12.6	8.51	8.06
P 6	9.27	10.9	9.62	15.8	9.52	7.21	6.61	10.5	8.14	14.4	11.3	10.6	15.3	12.8	13.1	9.24	11.2	12.5	13.9	9.24	9.75
P 7	8.98	11.5	8.95	15.0	9.49	7.19	4.04	9.67	7.84	12.6	9.12	10.1	14.9	12.0	12.5	8.46	10.2	11.7	12.8	9.39	8.73
P 8	5.59	10.5	7.96	11.5	6.91	9.42	9.03	3.69	7.55	12.5	10.4	7.42	12.4	9.98	9.15	8.21	9.31	10.9	11.7	10.6	6.28
P 9	5.82	10.5	7.18	13.2	6.57	7.83	7.98	6.38	4.92	12.8	10.4	8.22	12.6	10.4	9.13	8.02	9.46	10.5	10.9	8.97	6.10
P 10	11.2	8.40	8.54	10.0	11.8	8.79	8.48	8.18	9.47	10.7	11.8	5.10	16.2	7.76	9.15	9.62	8.62	11.1	9.59	8.60	8.27
P 11	10.7	14.1	9.67	17.0	11.1	11.0	9.67	10.6	10.4	13.9	3.65	12.7	15.4	12.6	12.8	10.6	12.3	14.6	14.1	11.7	12.7
P 12	7.95	9.27	8.74	10.4	9.58	7.98	7.92	5.85	7.70	12.4	12.8	5.20	14.5	9.98	9.50	8.20	9.01	9.87	11.2	10.1	4.77
P 13	10.3	15.5	10.8	16.7	10.7	14.6	14.2	10.6	12.2	15.7	13.3	13.5	5.31	12.1	14.6	10.7	10.5	9.89	13.3	13.4	12.6
P 14	13.6	11.3	10.2	14.4	13.0	13.0	13.5	12.0	11.8	16.5	15.3	11.4	15.1	9.93	13.8	10.6	10.1	10.2	8.38	9.80	12.8
P 15	10.8	7.97	9.51	6.32	11.9	11.3	12.0	8.07	10.6	12.6	14.4	6.18	16.2	7.77	8.29	9.23	8.63	10.7	8.38	9.96	7.98
P 16	7.83	9.42	6.47	12.2	8.24	7.99	8.20	6.89	6.93	12.7	10.4	8.13	12.1	9.51	10.5	3.52	7.11	8.02	10.7	8.03	8.03
P 17	10.5	9.85	7.58	11.4	10.6	9.66	9.67	8.81	8.66	13.3	12.1	8.60	11.3	8.63	11.2	6.42	3.57	5.32	8.69	8.28	9.32
P 18	11.3	12.9	9.09	14.0	11.3	11.9	11.3	10.0	10.1	14.5	13.1	11.2	10.2	10.4	13.3	7.90	7.42	4.32	9.75	10.6	10.82
P 19	11.7	10.5	8.97	11.7	11.5	11.2	11.3	9.39	10.3	14.3	13.8	9.3	13.3	9.44	11.9	8.03	7.34	7.23	6.86	9.46	10.7
P 20	5.55	10.4	8.79	13.7	6.29	8.89	9.85	7.49	6.56	14.2	11.8	9.5	12.8	11.6	10.0	9.30	10.3	11.6	12.7	9.75	6.96
P 21	5.96	8.06	7.92	11.9	6.71	7.01	8.05	6.26	5.61	13.4	12.83	6.13	13.2	10.1	8.84	8.07	9.15	10.2	11.1	8.17	4.24

Tabla 4.5.- Tabla de aciertos aplicando la técnica de 6 histogramas y reducción de parámetros de profundidad.

La principal fuente de error para los valores de profundidad viene derivada de las secciones de la persona, ya que, para un porcentaje muy elevado de las secuencias, estos valores son muy malos, haciendo imposible equilibrarlos con los buenos resultados de intensidad. Otro factor menos significativo es la eliminación de los componentes extras, éstos para la gran mayoría de las secuencias proporcionan buenos resultados, con su eliminación los errores se agravan en gran medida. En cuanto a los errores producidos por los valores de intensidad, éstos se corresponden con la parte del pelo como la técnica

detallada anteriormente. En la [Tabla 4.6](#) se muestran ejemplos de los resultados obtenidos con y sin los componentes extras, donde se contempla que con la adición de estos se obtienen buenos resultados en profundidad para la re-identificación, mientras que, sin ellos, la re-identificación sería incorrecta para las secuencias que se ven la tabla.

Secuencia Analizada	Secuencia Evaluada	Resultado con los componentes extras	Resultado sin los componentes extras
Seq-JSD-D02	Seq-BAP-D01	1.65	0.427
	Seq-JSD-D01	0.46	0.959
	Seq-JTM-D01	1.26	0.419
Seq-JTM-D02	Seq-JTM-D01	0.34	0.392
	Seq-MSP-D01	0.59	0.3
	Seq-BAP-D01	0.79	0.308
Seq-BAP-D02	Seq-BAP-D01	0.51	0.34
	Seq-TOP-D01	0.92	0.32
	Seq-AMC-D01	0.88	0.488

Tabla 4.6.– Comparativa de resultados de profundidad aplicando o no los componentes extras.

En la [Tabla 4.7](#) se muestran los resultados obtenidos estableciendo una prioridad del 50% al error de profundidad, un 40% al del histograma 2 y un 10% al del histograma 1. Los resultados aplicando un control de prioridades mejoran considerablemente, ya que se consigue reducir los grandes errores y potenciar los buenos resultados. Con este control se ha obtenido entre un 85-90% de acierto, consiguiendo ser la mejor técnica para utilizar en la identificación y re-identificación de personas. En cuanto a las 3 secuencias fallidas, cabe destacar que tanto para profundidad e intensidad se obtienen errores muy elevados haciendo imposible el equilibrio de estos.

A modo de resumen, en la [Figura 4.4](#) se muestra un gráfico con la comparativa de resultados obtenidos utilizando las diferentes técnicas descritas en el capítulo 4.

	P 1	P 2	P3	P4	P5	P6	P7	P8	P9	P 10	P 11	P 12	P 13	P 14	P 15	P 16	P 17	P 18	P 19	P 20	P 21
--	-----	-----	----	----	----	----	----	----	----	------	------	------	------	------	------	------	------	------	------	------	------

P 1	0.84	1.64	1.90	1.79	1.17	1.94	2.18	0.96	1.38	1.95	2.12	1.12	2.02	2.07	1.41	1.50	1.95	2.33	5.81	1.73	1.01
P 2	1.57	1.43	2.69	2.36	1.82	1.49	2.15	2.04	1.81	2.62	2.13	1.86	3.21	2.99	2.25	2.14	2.76	3.24	6.80	2.27	1.74
P 3	1.73	2.00	0.57	2.74	1.35	1.94	1.74	1.36	1.25	2.08	1.15	1.57	1.94	1.16	1.58	1.39	1.34	1.76	4.87	1.26	1.88
P 4	1.65	1.72	2.48	1.23	2.04	1.97	2.29	1.79	1.97	2.22	2.66	1.44	2.56	2.49	2.15	1.78	2.26	2.47	6.1	1.95	1.49
P 5	0.92	1.61	1.92	1.90	0.95	1.76	2.07	1.12	1.35	1.91	2.01	2.18	1.90	2.14	1.64	1.30	1.95	2.30	5.8	1.49	1.36
P 6	1.64	1.45	1.82	2.63	1.46	1.04	1.30	1.91	1.30	1.85	1.57	1.70	2.55	2.26	2.05	2.28	1.94	2.53	6.00	1.41	1.92
P 7	1.75	1.89	1.84	2.87	1.52	2.29	0.59	2.01	1.33	1.59	1.35	1.90	2.67	2.26	2.13	1.58	2.01	2.58	6.00	1.54	2.08
P 8	1.06	1.78	1.92	1.81	1.23	1.93	2.07	0.78	1.48	1.88	2.20	1.04	2.08	2.05	1.48	1.57	1.86	2.26	7.75	1.73	2.09
P 9	1.42	1.85	1.94	1.94	1.49	1.65	2.16	1.34	1.26	1.85	1.19	2.27	2.09	2.22	1.65	1.44	1.99	2.21	5.75	1.58	1.38
P 10	2.07	1.82	2.20	2.48	2.05	1.62	1.68	1.96	1.78	1.67	1.80	1.58	2.90	2.23	1.98	1.87	2.12	2.71	6.09	1.80	2.06
P 11	1.86	1.90	2.15	2.93	1.83	1.27	1.62	1.12	1.67	1.82	0.7	2.16	2.89	2.59	2.18	1.83	2.36	3.05	2.54	1.91	2.39
P 12	1.27	1.57	2.00	1.66	1.54	1.77	2.01	1.13	1.50	1.82	2.43	0.87	2.19	2.05	1.59	1.50	1.90	2.09	5.71	1.69	0.93
P 13	1.86	2.51	1.55	2.61	1.71	2.58	2.46	1.58	1.89	2.22	2.61	1.85	0.92	1.49	2.19	1.53	1.38	1.20	5.03	1.57	2.06
P 14	1.94	1.94	1.09	2.52	1.61	2.20	2.01	1.61	1.51	2.23	2.55	1.59	1.65	0.89	1.82	1.35	1.25	1.42	4.78	1.04	1.96
P 15	1.36	1.36	1.60	1.80	1.37	2.02	1.98	1.21	1.39	2.22	2.48	1.10	2.50	1.57	1.03	1.82	1.71	2.27	5.56	1.61	1.17
P 16	1.28	1.41	1.18	2.18	1.08	1.39	1.51	1.24	0.97	1.65	1.80	1.18	1.94	1.65	1.58	0.51	1.39	1.90	5.49	0.84	1.50
P 17	1.79	1.92	1.25	2.36	1.58	1.92	1.71	1.45	1.44	1.86	2.23	1.40	1.78	1.34	1.73	1.18	0.58	1.43	5.20	1.07	1.76
P 18	2.08	2.41	1.46	2.48	1.88	2.34	2.11	1.62	1.86	2.01	2.56	1.70	1.08	1.38	2.18	1.35	1.32	0.74	4.60	1.34	2.02
P 19	1.91	1.99	1.23	2.45	1.66	2.10	1.88	1.42	1.55	2.00	2.40	1.41	1.40	1.10	1.74	1.32	0.97	1.09	0.82	1.22	1.83
P 20	0.70	0.97	1.19	1.54	0.69	1.01	0.73	1.01	0.56	1.30	1.05	1.10	1.72	1.33	0.79	1.21	1.23	1.91	6.32	1.07	0.93
P 21	1.13	1.54	1.58	1.99	1.14	1.72	1.88	0.91	1.06	1.97	2.30	0.92	2.07	1.82	1.19	1.50	1.59	2.02	5.59	1.38	0.84

Tabla 4.7.- Tabla de aciertos aplicando las prioridades para las secuencias.

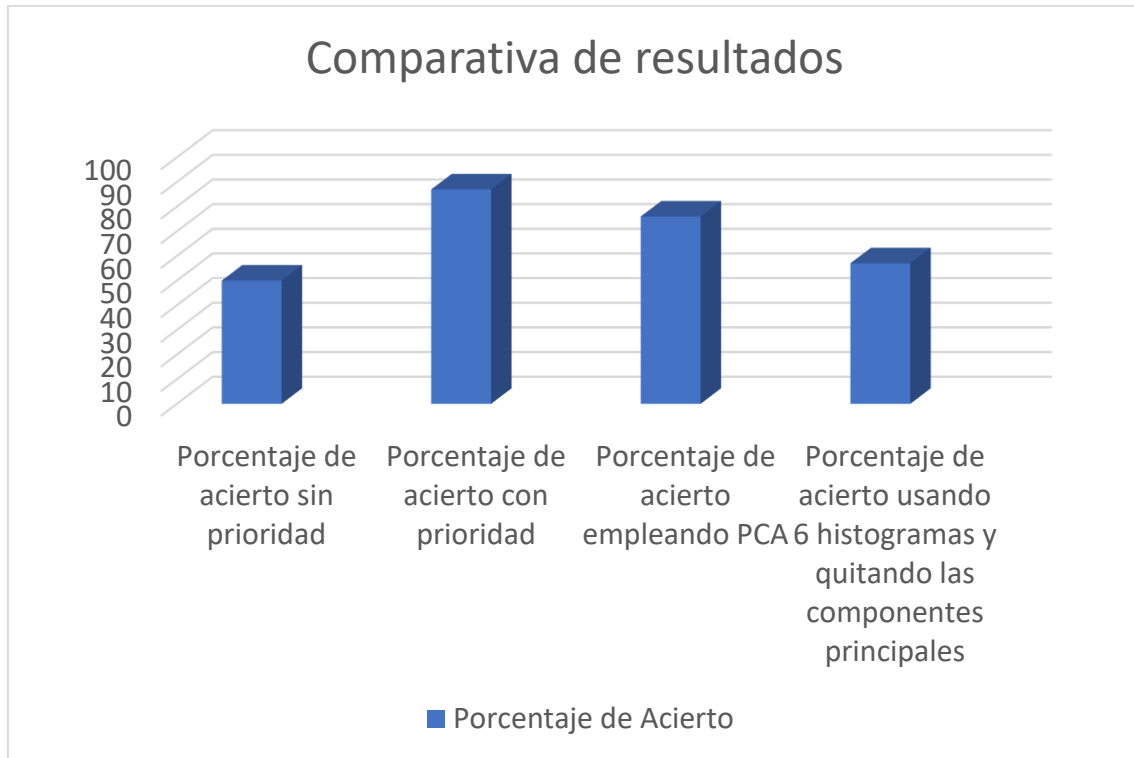
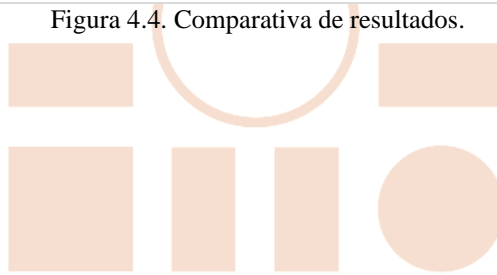


Figura 4.4. Comparativa de resultados.



ESCUELA POLITECNICA
SUPERIOR

5. Conclusiones y trabajos futuros.

En este trabajo se han desarrollado diferentes procedimientos para lograr una re-identificación lo más óptima posible. En todos, el resultado supera el 50% de acierto. No obstante, la técnica que mejores resultados mostró, es la basada en el cálculo de distancias euclídeas aplicando un control de prioridades para secuencias de intensidad y profundidad, obteniéndose entre un 85% y 90% de acierto.

Además, se ha constatado que aplicar un control de prioridades a los diferentes vectores de características mejora sustancialmente los resultados.

5.1. Conclusión personal.

Antes de empezar la realización de este trabajo, tenía apenas una pequeña noción sobre el tema de visión artificial y el procesamiento de imágenes. Ahora, una vez terminado este trabajo, puedo decir que he aprendido gran cantidad de cosas, abriéndome nuevas vías de trabajo para futuros. Gracias a este trabajo, he podido mejorar sustancialmente mi manejo con el lenguaje de programación C/C++, como también el uso de las librerías de OpenCV. También, he mejorado la manera de afrontar un proyecto, marcar una serie de pautas y conseguir abordar el proyecto con éxito.

5.3 Trabajos futuros

El campo de la identificación y re-identificación está avanzando bastante en estos últimos años, pero, a pesar de esto, aún queda muchísimo por descubrir o desarrollar, este TFG es un pequeño grano para dicho avance. A continuación, se propone una serie de posibles trabajos futuros.

- **Creación de dos modelos PCA's:** Para mejorar los resultados, se propone la creación de un modelo PCA para la secuencia de profundidad y otro para la de intensidad, marcando como prioritario el modelo de profundidad.
- **Utilización de otra cámara:** Se propone la utilización de otra cámara con mayor resolución, lo que daría una mejora de las secuencias tomadas y por conveniente una mejora de los resultados.
- **Utilización de otros Clasificadores:** Se propone la utilización de otros clasificadores más actuales, con los que quizá se pueda mejorar los resultados obtenidos.
- **Integración de PCA y distancia Euclídea:** Se propone la utilización de un modelo PCA para las secuencias de profundidad, hallando su error para cada una de las secuencias, mientras que para las secuencias de intensidad realizar el uso de la distancia euclídea.

6. Bibliografía

- [1] – Sara Luengo, “Re-identificación de personas utilizando únicamente información de profundidad”, 2018, TFG UAH.
- [2] – Sara Luengo Sánchez, Carlos Luna Vázquez, Cristina Losada Gutiérrez, Manuel Mazo Quintas, “Re-identificación de personas utilizando únicamente información de profundidad”. En Memorias de la 19ª Convención científica de Ingeniería y Arquitectura. X Congreso Internacional de Telemática y Telecomunicaciones (CITTEL 2018), 2018, ISBN: 978-959-261-585-4.
- [3] https://www.baslerweb.com/fp-1523277387/media/downloads/documents/users_manuals/AW00133810000_Tof_GigE_UM.pdf .[Último acceso 02/06/2020].
- [4]-Daniel Montagud Cortés, “Estudio de cámaras de tiempo de vuelo (ToF) para aplicaciones en exteriors”, 2018, TFG UAH.
- [5] – Carlos A. Luna, Cristina Losada-Gutierrez, David Fuentes-Jimenez, Alvaro Fernandez-Rincon, Manuel Mazo, Javier Macias-Guarasa, “Robust people detection using depth information from an overhead Time-of Flight camera”. In 12th International Conference on Computer Vision Theory and Applications.,2017, ISBN: 978-989-758-225-7
- [6] - Matzner, S., Heredia-Langner, A., Amidan, B., Boettcher, E., Lochtefeld, D., & Webb, T. (2015). “Standoff human identification using body shape”. In 2015 IEEE international symposium on technologies for homeland security (HST), 2015, ISBN: 978-1-4799-1737-2.
- [7]- “Página web de wikipedia” www.wikipedia.com [Último acceso 28/04/2020]
- [8] – “Página de OpenCV” www.OpenCV.org [Último acceso 15/01/2020]
- [9]- “Página Web de Basler”: <https://www.baslerweb.com/> [Último acceso 2/02/2020]
- [10] – “Página Web de la librería libfreenect2” <https://github.com/OpenKinect/libfreenect2> [Último acceso 15/01/2020]
- [11]- D. Jimenez, D. Pizarro and M. Mazo “Single frame correction of motion artifacts in pmd-based time of flight cameras”. In Image and visión Computing, 2014, <https://doi.org/10.1016/j.imavis.2014.08.014>
- [12] - D. Fuentes Jiménez et al, “Diseño, implementación y evaluación de un sistema de conteo de personas basado en cámaras de tiempo de vuelo”, 2016, TFG UAH.
- [13] - R. García Jiménez et al, “Detección y conteo de personas, a partir de mapas de profundidad cenitales capturados con cámaras ToF”, 2015. TFG UAH.
- [14]- Daniele Liciotti, Marina Paolanti, Emanuele Frontoni, Adriano Mancini, and Primo Zingaretti, “Person Re-identification Dataset with RGB-D Camera in a Top-View Configuration”, Published online 2018 Oct 15. DOI: 10.3390/s18103471

Anexo A

A.1 Manual de Usuario

En este apartado se explica todo el funcionamiento del sistema y la estructura de todo el proyecto realizado para este trabajo.

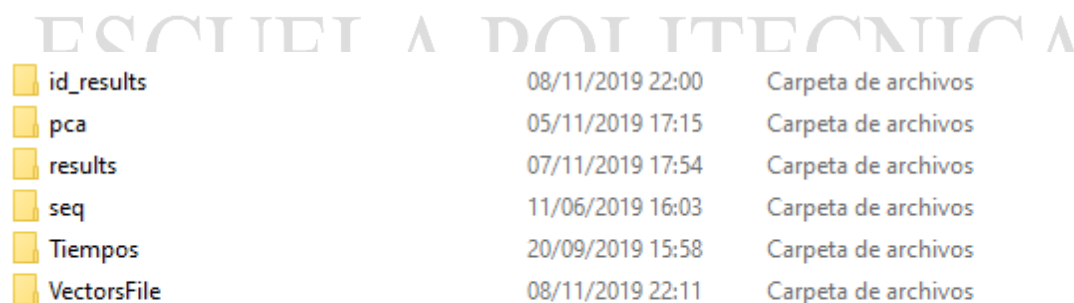
A.1.1 Requisitos del Sistema

Para el uso de la aplicación realizada, se tiene que disponer del sistema operativo de Windows 10, como también, del entorno de desarrollo de Visual Studio 2017 o posteriores, utilizando el lenguaje de programación C/C++. Por tanto, el usuario debe disponer de estas herramientas para un correcto funcionamiento de la aplicación creada. También se ha hecho uso de las siguientes librerías:

- **Librería de OpenCV:** Comúnmente utilizada para todo lo relacionado con la visión artificial, se utiliza para todo tipo de procesamiento de imágenes. OpenCV aporta multitud de funciones hechas, facilitándonos así, el desarrollo de nuestro trabajo.
- **Librería de Basler ToF:** Esta librería, ha sido utilizada para poder utilizar y configurar la cámara Basler, en ella incluía todas las funciones necesarias para la captura de imágenes.

A.1.2 Estructura de los directorios

Como este trabajo se basa en una modernización de un trabajo ya realizado anteriormente [1], se ha realizado la misma estructura de directorios, donde se almacenan todos los datos creados durante la realización de la aplicación. A continuación, la [Figura A.1](#) muestra la estructura utilizada.



id_results	08/11/2019 22:00	Carpeta de archivos
pca	05/11/2019 17:15	Carpeta de archivos
results	07/11/2019 17:54	Carpeta de archivos
seq	11/06/2019 16:03	Carpeta de archivos
Tiempos	20/09/2019 15:58	Carpeta de archivos
VectorsFile	08/11/2019 22:11	Carpeta de archivos

Figura A.1. Directorios utilizados en la aplicación.

A.1.3.1 ID_RESULTS

En esta carpeta se encuentran todos los ficheros correspondientes con los resultados obtenidos, estos archivos son generados en la segunda etapa, ya que se corresponden con la etapa de re-identificación. Estos ficheros nos aportan la información recogida por la re-identificación. Estos ficheros tendrán una extensión. Reid. Estos resultados pueden ser evaluados con cualquier editor de texto

A.1.3.2 PCA

En esta carpeta se encuentran todos los archivos relacionados con la PCA, ambas carpetas se utilizan en las dos etapas, por tanto, son generadas al principio de la aplicación.

- **Classes:** contiene los archivos con las clases creadas. Cada clase está nombrada de acuerdo con la secuencia a partir de la cual se creó. Estos archivos tienen la extensión. **.pca**.
- **Data:** contiene archivos de texto con información sobre las clases creadas. Esta información consiste en la media, los autovalores, los auto vectores y la covarianza. Estos archivos tienen la extensión. **.pcadata**.

A.1.3.3 RESULTS

En esta carpeta se encuentran todos los datos relacionados con el detector, como pueden ser el número de objetos presentes en la imagen detectada, las coordenadas de los centroides, la altura de la persona detectada y si es o no persona.

A.1.3.4 SEQ

En esta carpeta, se encuentran almacenados todos los ficheros con las secuencias de intensidad y profundidad grabadas para luego ser utilizadas. Estas secuencias han sido almacenadas con una extensión **.Pcd**.

*Como se han tenido problemas con el uso de la cámara Basler, y se ha usado otra cámara para la captura de las secuencias, se ha cambiado el formato de dichas secuencias.

- **Secuencias de Intensidad:** extensión **.ir16**
- **Secuencias de Profundidad:** extensión. **.z16**

A.1.3.5 TIEMPOS

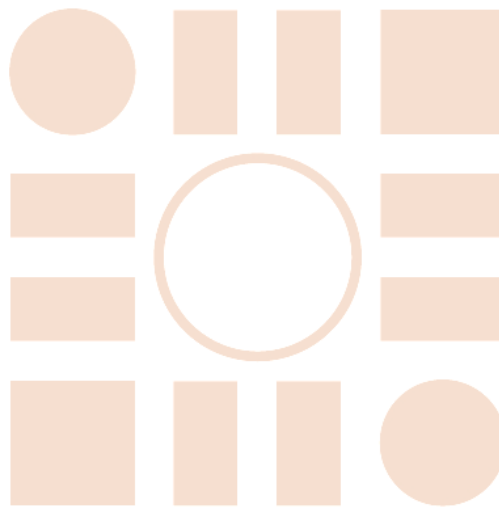
Este directorio almacena los archivos que contienen la información del tiempo de ejecución de los algoritmos del detector. En total hay cinco archivos cuyos nombres se corresponden con la parte del algoritmo de detección cuya información contienen: ClasificadorTime.data, direccionesTime.data, Filter-Time.data, maxDetectorTime.data y vecCaractTime.data. [1]

A.1.3.6 VECTORSFILE

En esta carpeta, se almacenan todos los datos del vector de características de todas las secuencias procesadas. Estos vectores se almacenan, para así, posteriormente generar la clase PCA o para calcular el error, extrayendo la información almacenada en esta carpeta. Estos archivos son binarios, por lo que su extensión es **.data**.

A.1.3.7 PROJECT

En la carpeta Project se encuentran todas las funciones utilizadas en este trabajo, la mayoría de las funciones han sido desarrolladas en otros trabajos ([1], [12],[13]).



ESCUELA POLITECNICA
SUPERIOR

A.1.3 Uso de la Aplicación

A continuación, se detalla los pasos a seguir para un correcto funcionamiento de la aplicación. Una vez compilado y ejecutado el programa, se encuentra la primera pantalla como se observa en la [Figura A.2](#).

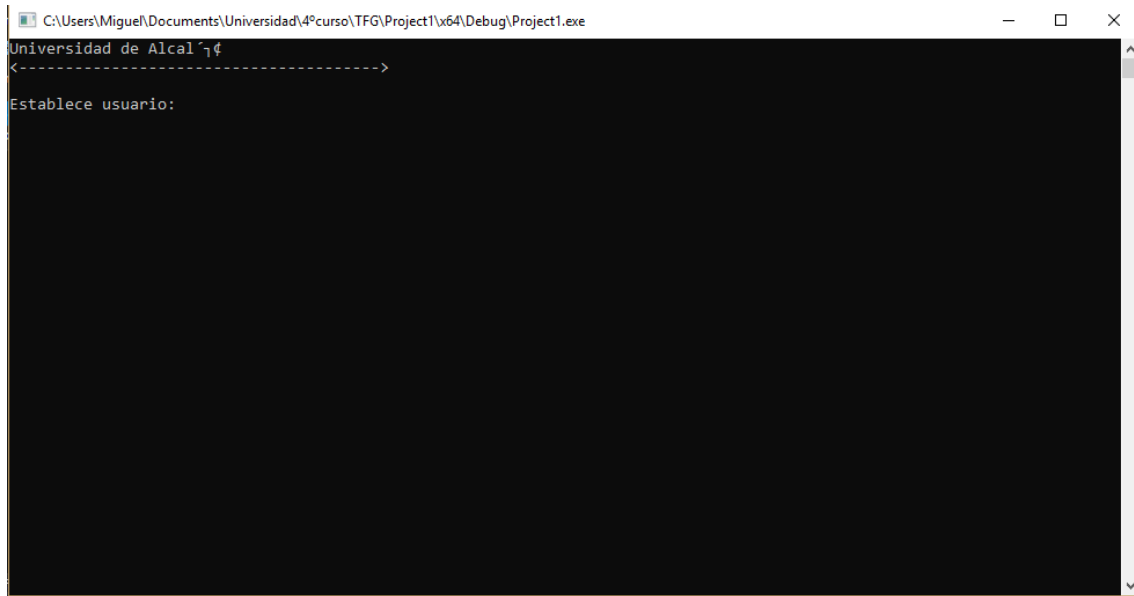


Figura A.2. Pantalla inicial de la interfaz de usuario.

En esta primera pantalla, se pide que se establezca un usuario para la creación de los ficheros tanto de Intensidad como de Profundidad. Una vez establecido el usuario, se procede al paso de la segunda pantalla.

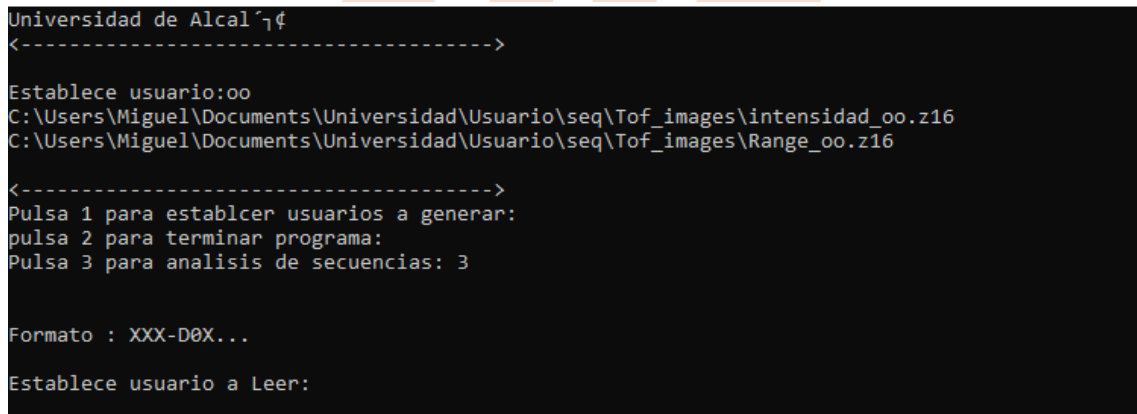


Figura A.3. Segunda pantalla de la interfaz de usuario.

En esta segunda pantalla ([Figura A.3](#)), se observa como aparecen dos rutas. Estas dos rutas hacen referencia a la base de datos donde se ha generado y guardado los ficheros correspondientes al usuario establecido. También se contempla en el interfaz de usuario tres posibles pasos a seguir.

- La primera opción consiste en la creación y grabación de las secuencias de intensidad y profundidad estableciendo un nuevo usuario.
- La segunda opción consiste en el cierre de la cámara y el final del programa.
- La tercera opción consiste en el análisis de las secuencias grabadas y/o almacenadas en la base de datos.

Si el usuario ha establecido la opción de análisis de las secuencias, se da paso a una tercera pantalla, dicha pantalla se muestra en la [Figura A.4](#).



```
Universidad de Alcalá
<----->

Establece usuario:oo
C:\Users\Miguel\Documents\Universidad\Usuario\seq\Tof_images\intensidad_oo.z16
C:\Users\Miguel\Documents\Universidad\Usuario\seq\Tof_images\Range_oo.z16

<----->
Pulsa 1 para establecer usuarios a generar:
pulsa 2 para terminar programa:
Pulsa 3 para analisis de secuencias: 3

Formato : XXX-D0X...
Establece usuario a Leer:
```

Figura A.4. Tercera pantalla de la interfaz de usuario.

En la [Figura A.4](#), se pide que se establezca una secuencia a leer, esta debe seguir el formato que se indica en la interfaz de usuario ([Figura A.4](#)). La secuencia establecida debe encontrarse dentro de la base de datos utilizada para esta aplicación, si se encuentra dentro se procede a una cuarta pantalla en el interfaz de usuario ([Figura A.5](#)).

```
Universidad de Alcalá
<----->

Establece usuario:oo
C:\Users\Miguel\Documents\Universidad\Usuario\seq\Tof_images\intensidad_oo.z16
C:\Users\Miguel\Documents\Universidad\Usuario\seq\Tof_images\Range_oo.z16

<----->
Pulsa 1 para establecer usuarios a generar:
pulsa 2 para terminar programa:
Pulsa 3 para analisis de secuencias: 3

Formato : XXX-D0X...

Establece usuario a Leer:MSP-D02
Analizando imagenes en Profundidad
Analizando imagenes en Intensidad
The sequence file has been opened...

Establece funcionalidad:
Pulsa 0-> Para análisis
Pulsa 1-> Para almacenamiento de datos
Pulse:
```

Figura A.5. Cuarta pantalla de la interfaz de usuario.

La nueva pantalla de la interfaz propone dos nuevas opciones, la segunda opción se basa en el análisis de la secuencia, es decir, es la encargada en realizar la etapa de identificación, mientras que la primera opción realiza la etapa de re-identificación. Cuando una de estas dos funcionalidades acaba, el programa vuelve a la segunda pantalla del interfaz, para que así, el usuario vuelva a realizar la función deseada.

ESCUELA POLITECNICA
SUPERIOR

Anexo B

B.1 Pliego de condiciones

Este anexo presenta el Hardware y Software utilizado en este trabajo.

B.1.1 Hardware

- a. Ordenador portátil ideapad 330S:
 - Sistema Operativo: Windows 10 64-bit.
 - Procesador: Intel® Core™ i5-8250U CPU @ 1.60 GHz 1.80 GHz.
 - Memoria RAM: 8 Gb.
 - Disco Duro: 256 Gb SSD.
- b. GPU:
 - Intel UHD Graphics 620.

B.1.2 Software

Para la realización de este TFG se han empleado los siguientes programas:

- a. **Visual Studio:** Microsoft Visual Studio es un entorno de desarrollo integrado (IDE, por sus siglas en inglés) para Windows, Linux y macOS. Es compatible con múltiples lenguajes de programación, tales como C++, C#, Visual Basic .NET, F#, Java, Python, Ruby y PHP. [7]
- b. **Librería OpenCV:** Biblioteca utilizada principalmente para Visión artificial. [8]
- c. **Librería Basler Tof:** Biblioteca utilizada para configuración de la Cámara. [9]
- d. **Librería Kinect:** Biblioteca utilizada para configuración de la Cámara. [10]

Anexo C

C.1 Presupuesto

En este Anexo, se va a detallar todo el presupuesto usado para la realización de este TFG. En primer lugar, se expone el presupuesto para los materiales Hardware. [\[Tabla C.1.1\]](#)

Material	Cantidad	Subtotal
Ordenador	1	650€
GPU	1	500€
Sensor ToF Kinect	1	177€
Total	1327€	

Tabla C.1. Costes de Hardware.

En segundo lugar, se expone el presupuesto para las herramientas Software. [\[Tabla C.1.2\]](#)

Herramientas	Cantidad	Subtotal
Windows 10	1	40€
Librería OpenCV	1	0€
Visual Studio	1	0€
Librería Basler ToF	1	0€
Librería Kinect	1	0€
Total	40€	

Tabla C.2. Costes de Software.

En tercer lugar, se expone el presupuesto concerniente a la mano de obra. [\[Tabla C.1.3\]](#)

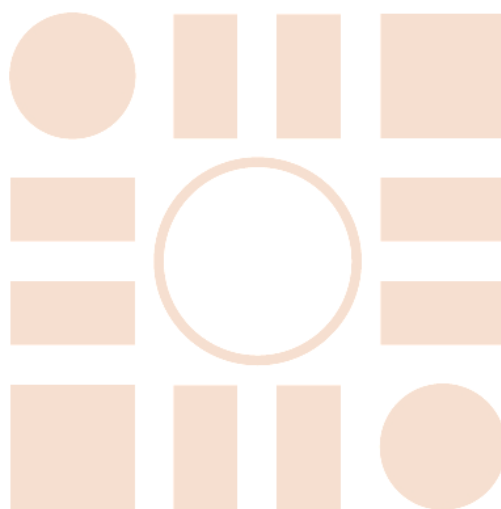
Trabajo	Cantidad(horas)	Precio Unitario(€/h)	Subtotal
Desarrollo Software Costes de Hardware y Software	200	70€	14000€
Documentación	55	10€	550€
Total	14550€		

Tabla C.3. Costes de mano de obra.

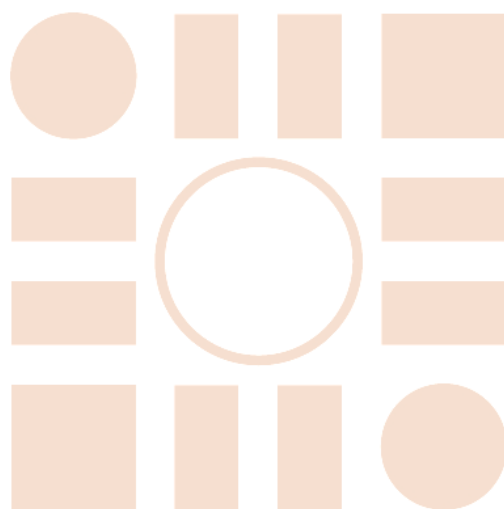
Para finalizar, se muestra el coste total del proyecto. [Tabla C.1.4]

Material	Subtotal
Desarrollo Software	1673€
Desarrollo Software y Documentación	14550€
Total	15673€

Tabla C.4. Coste total.



ESCUELA POLITECNICA
SUPERIOR



ESCUELA POLITECNICA
SUPERIOR



Universidad
de Alcalá