

Universidad de Alcalá  
Escuela Politécnica Superior

MASTER EN INGENIERÍA INDUSTRIAL  
Especialidad en Robótica y Percepción

**Trabajo Fin de Máster**

Detección y clasificación de personas en secuencias de  
mapas de profundidad cenitales.

**Autor:** Luis Javier González Chico

**Tutor/es:** Carlos Andrés Luna Vázquez

ESCUELA POLITECNICA  
SUPERIOR

2017



UNIVERSIDAD DE ALCALÁ  
Escuela Politécnica Superior

MÁSTER EN INGENIERÍA INDUSTRIAL  
Especialidad en Robótica y Percepción

Trabajo Fin de Máster  
Detección y clasificación de personas en secuencias de  
mapas de profundidad cenitales.

**Autor:** Luis Javier González Chico

**Tutor/es:** Carlos Andrés Luna Vázquez

**TRIBUNAL:**

**Presidente:** Pedro Gil Jiménez

**Vocal 1º:** Cristina Losada Gutiérrez

**Suplente:** Javier Macias Guarasa

**FECHA:** SEPTIEMBRE 2017

**CALIFICACIÓN:**





## *Agradecimientos:*

*Este trabajo está dedicado a todos los familiares, amigos, compañeros y personal docente que ha formado parte de estos dos años de estudios, pero tiene una dedicación especial a:*

*Isabel Barceló, mi prometida y el amor de mi vida. Por haberme aguantado estos dos duros años de Máster, superando desde una "simple" mudanza hasta una fractura de tibia y peroné y 3 operaciones por ahora con una sonrisa a mi lado.*

*A mis padres y a mi hermana por sus palabras y frases de ánimo, desde aquí o desde allí, no importa la distancia, siempre se agradecen.*

*A mi abuela que siempre me recuerda que yo no quería estudiar y ahora ya me dice "Ya está bien de estudiar, ¿no?", te prometo que me voy a tomar un descanso, por ahora.*

*Alberto Ceballos, "el Cebas", si no fuera por tu culpa, no estaría hoy aquí, cada vez queda menos para tener funcionando LUJAL S.A..*

*Los modelos, Aida, Almudena, Carlos, David, Esteban y Rocío por prestarme un rato de su valioso tiempo para realizar los vídeos, sin vosotros no hubiese sido posible.*



# CONTENIDO

---

Índice de tablas . . . . .	9
Índice de figuras . . . . .	10
Índice de ecuaciones. . . . .	11
Lista de acrónimos . . . . .	12
Resumen en castellano . . . . .	13
Resumen en inglés. . . . .	13
Palabras clave. . . . .	13
Resumen extendido . . . . .	14
1 Introducción . . . . .	17
2 Bases Teóricas. . . . .	18
2.1 Estado del arte . . . . .	18
2.2 Descripción teórica del trabajo. . . . .	20
2.2.1 Proceso Off-line. . . . .	21
2.2.2 Proceso On-line. . . . .	30
3 Descripción experimental. . . . .	34
3.1 Adquisición de secuencias de mapas de profundidad . . . . .	34
3.1.1 Procedimiento de configuración del entorno de trabajo. . . . .	34
3.1.2 Software de adquisición de datos de profundidad. . . . .	36
3.2 Creación de la base de datos. . . . .	37
3.2.1 Entorno y población a estudio. . . . .	38
3.2.2 Codificación del nombre de las secuencias . . . . .	38
3.2.3 Diversidad de secuencias. . . . .	40
3.2.4 Base de datos de secuencias. . . . .	41
3.3 Clasificación . . . . .	45
3.3.1 Generación de clases modelo . . . . .	45
3.3.2 Comprobación de clases modelo. . . . .	46
3.4 Filtrado de las medidas de profundidad . . . . .	47
3.5 Reducción de la zona de trabajo . . . . .	51
3.5.1 Acotación de la región de interés de análisis. . . . .	53
4 Análisis de resultados . . . . .	61
4.1 Diferenciación doble. . . . .	61
4.2 Diferenciación múltiple . . . . .	62
4.3 Diferenciación individual. . . . .	62
5 Conclusiones . . . . .	64
6 Líneas de trabajo futuras . . . . .	65

7	Posibles aplicaciones. ....	66
8	Pliego de condiciones. ....	67
9	Presupuesto. ....	68
10	Bibliografía. ....	69
11	- Anexo 1 – Resultados sin filtrar. ....	70
12	- Anexo 2 – Resultados filtrados. ....	73

## ÍNDICE DE TABLAS

---

Tabla 1 - Codificación de sujetos .....	38
Tabla 2 - Codificación de secuencia.....	39
Tabla 3 - Codificación de complementos .....	39
Tabla 4 - Ejemplo de codificación.....	39
Tabla 5 - Secuencias seleccionadas como archivos modelo .....	41
Tabla 6 - Secuencias de pruebas.....	42
Tabla 7 - Tabla de presentación de resultados modificación 2.....	50
Tabla 8 - Vectores características modelos .....	57
Tabla 9 - Comparativa de vectores .....	58
Tabla 10 - Coeficiente de correlación de Pearson.....	59
Tabla 11 - Comparativa resultados con y sin complemento.....	61
Tabla 12 - Tasa de acierto diferenciación doble.....	61
Tabla 13 - Comparativa resultados agrupados por tamaños.....	62
Tabla 14 - Tasa de acierto diferenciación múltiple.....	62
Tabla 15 - Comparativa resultados diferenciación individual.....	62
Tabla 16 - Tasa de acierto diferenciación individual.....	63
Tabla 17 - Comparativa resultados agrupados.....	63
Tabla 18 - Tasa de acierto agrupada.....	63
Tabla 19 - Precio Hardware.....	68
Tabla 20 - Precio Software .....	68
Tabla 21 - Precio Mano de Obra .....	68
Tabla 22 - Precio Total .....	68
Tabla 23 - Ejemplo de resultados.....	71
Tabla 24 - Tabla de presentación de resultados modificación 1.....	72
Tabla 25 - Ejemplo de resultados.....	75

## ÍNDICE DE FIGURAS

---

Figura 0-1 - Ejemplo de siluetas de sujetos con diferentes complementos.....	16
Figura 1-1 - Diagrama en bloques a realizar en el TFM.....	17
Figura 2-1 - Descripción proceso de clasificación.....	20
Figura 2-2 - Descripción proceso off-line.....	21
Figura 2-3 - Posición de la cámara.....	22
Figura 2-4 - Altura de interés.....	23
Figura 2-5 - Píxeles agrupados por subregiones.....	25
Figura 2-6 - Matriz de máximos por subregiones.....	26
Figura 2-7 - Agrupación de máximos de subregiones colindantes.....	27
Figura 2-8 - Direcciones radiales de búsqueda.....	27
Figura 2-9 - Densidad de píxeles por nivel.....	29
Figura 2-10 - Descripción proceso on-line.....	31
Figura 3-1 - Proceso de obtención de secuencias de mapas de profundidad..	34
Figura 3-2- Interfaz gráfica modificada.....	37
Figura 3-3 - Sombrero tipo 1.....	40
Figura 3-4 - Sombrero tipo 2.....	40
Figura 3-5 - Sombrero tipo 5.....	40
Figura 3-6 - Sombrero tipo 7.....	40
Figura 3-7 - Sombrero tipo 8.....	40
Figura 3-8 - Sombrero tipo 9.....	40
Figura 3-9 - Entorno de trabajo.....	43
Figura 3-10 - Complemento tipo 1.....	43
Figura 3-11 - Complemento tipo 2.....	43
Figura 3-12 - Complemento tipo 3.....	44
Figura 3-13 - Complemento tipo 4.....	44
Figura 3-14 - Complemento tipo 5.....	44
Figura 3-15 - Complemento tipo 6.....	44
Figura 3-16 - Complemento tipo 7.....	44
Figura 3-17 - Complemento tipo 8.....	44
Figura 3-18 - Complemento tipo 9.....	44
Figura 3-19 - Clasificación de secuencias.....	45

Figura 3-20 - Representación de trayectorias seguidas por los sujetos. ....	46
Figura 3-21 - Secuencia cenital sin filtrar. ....	47
Figura 3-22 - Secuencia cenital filtrada. ....	48
Figura 3-23 - Comparativa filtrado. ....	48
Figura 3-24 - Trayectoria de sujeto bajo cámara ToF. ....	51
Figura 3-25 - Acotación de datos. ....	52
Figura 3-26 - Sujeto centrado en la secuencia. ....	52

## ÍNDICE DE ECUACIONES

---

(1). ....	22
(2). ....	22
(3). ....	22
(4). ....	23
(5). ....	24
(6). ....	24
(7). ....	25
(8). ....	26
(9). ....	26
(10). ....	28
(11). ....	29
(12). ....	31
(13). ....	32
(14). ....	32
(15). ....	32
(16). ....	32
(17). ....	33
(18). ....	33

## LISTA DE ACRÓNIMOS

---

SDK	<i>Software Development Kit.</i> Kit de Desarrollo de Software.
PCA	<i>Principal Component Analysis.</i> Análisis de Componentes Principales.
ToF	<i>Time of Flight.</i> Tiempo de Vuelo.
Roi	<i>Region of Interest.</i> Región de Interés.
API	<i>Application Programming Interface.</i> Interfaz de programación de aplicaciones.
IDE	<i>Integrated Development Environment.</i> Entorno de Desarrollo Integrado



## RESUMEN EN CASTELLANO

---

El objetivo final de este proyecto es la detección de personas y la clasificación de algunos complementos que las mismas puedan llevar. Para ello se usan los mapas de profundidad del entorno, obtenidos con una cámara de tiempo de vuelo situada en posición cenital. De dichos mapas de profundidad, a partir de las áreas que componen la parte superior del cuerpo (cabeza y hombros), se obtiene un vector de características para cada individuo. Estos vectores característicos son usados tanto en el entrenamiento del clasificador, como en el proceso de clasificación *online*, que en este trabajo se realiza mediante el análisis de componentes principales (PCA).

## RESUMEN EN INGLÉS

---

The main objective of this project is the detection of persons and the classification of some complements that can wear. For that depth maps are used obtained from a time of flight camera placed in top position. From those depth maps, and specially the upper body area (head and shoulders), a characteristic vector is obtained for each individual. Those vectors are used for training the classificatory and also in the online classification process which is done by the principal components analysis. (PCA)

## PALABRAS CLAVE

---

- Cámaras de tiempo de vuelo.
- Kinect 2.
- PCA.
- Detección de personas.
- Clasificación de personas.

## RESUMEN EXTENDIDO

---

En los últimos años se están instalando una gran cantidad de cámaras para detectar, clasificar e identificar personas en diversos entornos de la vida cotidiana, como pueden ser estaciones de tren y metro, terminales aéreas, hospitales, centrales nucleares, fabricas, etc.

La legislación española prohíbe grabar a los ciudadanos sin su consentimiento previo, haciendo muy complicado la aplicación de estas tecnologías.

Para la realización de este trabajo y poder respetar estos parámetros establecidos por la ley sobre la privacidad del ciudadano, se van a utilizar sensores de profundidad basados en tiempo de vuelo (ToF, Time of Flight). Los sensores de este tipo, ubicados en posición cenital, debido a su principio de funcionamiento, permiten obtener un mapa de profundidad del entorno dónde no es posible la identificación de la persona.

Otra ventaja de este tipo de sensores es la reducción del coste computacional a la hora de obtener información 3D de la escena.

Existen algunos trabajos en la literatura dónde se realiza la detección y conteo de personas de forma robusta, y no invasiva a través de cámaras ToF con visión cenital de la escena, pero no se han encontrado trabajos dónde se clasifiquen los complementos que una persona pueda llevar.

En este trabajo se pretenden desarrollar nuevos algoritmos de detección y clasificación de las personas en secuencias de mapas de profundidad, que permitan mejorar los resultados de trabajos previos. Para ello se implementarán procedimientos de clasificación y caracterización de personas, atendiendo a forma de peinado y uso de complementos (gorras, sombreros, etc.).

### Objetivos y campo de aplicación

El principal objetivo es desarrollar algoritmos de detección, seguimiento y clasificación de personas respetando la privacidad de las mismas. Para ello, sólo se utilizará la información de profundidad que proporcionan las cámaras de tiempo de vuelo (ToF).

El campo de aplicación es muy amplio y se pueden citar algunas de las posibles aplicaciones:

- Vigilancia.
- Marketing.
- Control de accesos.
- Análisis de flujo de personas.
- Marketing dinámico (Digital signage).
- Análisis de comportamiento de personas.
- Análisis de tipos de personas.

## Descripción del trabajo

El objetivo final de este proyecto pretende clasificar por grupos, según sus características físicas, a las personas que se encuentren atravesando una puerta utilizando cámaras de tiempo de vuelo (ToF). Esto hace de la escena un entorno de dimensiones controladas, por dónde los sujetos van a pasar siguiendo dos trayectorias posibles, entrada y salida.

Una vez detectada la presencia de una persona, mediante el análisis de la información de profundidad entregada por una cámara ubicada en posición cenital, se analizan sus características físicas para proceder a clasificarla.

Existen muchas aplicaciones en que puede resultar útil conocer si una persona lleva un complemento determinado (gorra, sombrero, etc.) o no, como pueden ser sistemas de vigilancia, marketing, análisis de comportamientos, etc. Dado el caso particular de este proyecto, sólo se van a tener en cuenta los complementos que los individuos puedan llevar en la cabeza, así como dos tipos de peinados posibles.

Actualmente se está trabajando con cámaras tradicionales (con sensores de intensidad, ya sean a colores o de tonalidad de grises) para realizar tareas de control y marketing tanto en recintos privados como públicos (supermercados, centros comerciales, aeropuertos, estaciones de metro, etc.). En este trabajo se utilizan sensores de profundidad basados en tiempo de vuelo (ToF- del inglés *Time of Flight*). La principal ventaja del uso de este tipo de sensores, en establecimientos de pública concurrencia, frente a los tradicionales, es el respeto a la privacidad de las personas. Esto es posible, debido a que con los sensores ToF no se capturan imágenes de la escena como tal, sino que se obtiene en cada pixel un valor que representa la distancia entre el sensor y un punto de la escena. En el mercado se encuentran distintos tipos de estas cámaras, para este trabajo se optó por usar la cámara *Kinect 2* de *Microsoft* debido a su reducido coste y a que dispone de librerías y software gratuito.

La cámara será instalada en posición cenital delante de una puerta, donde se garantiza el paso de personas de forma individual.

Un problema de las cámaras TOF cuando obtienen un mapa de las medidas de profundidad de cada punto de la escena, es que debido al principio de funcionamiento de las mismas, existe un error aleatorio en cada pixel de dicho mapa, que en escenas sin objetos en movimiento puede rondar los 10 mm. Este error puede resultar mayor en los cambios bruscos de altura de los objetos (bordes), llegando en algunos casos a ser descartada la medida por la cámara, indicando que es un pixel no válido. Por este motivo es necesario implementar un algoritmo de reducción de ruido, que a su vez permita hacer una estimación de los pixeles no válidos.

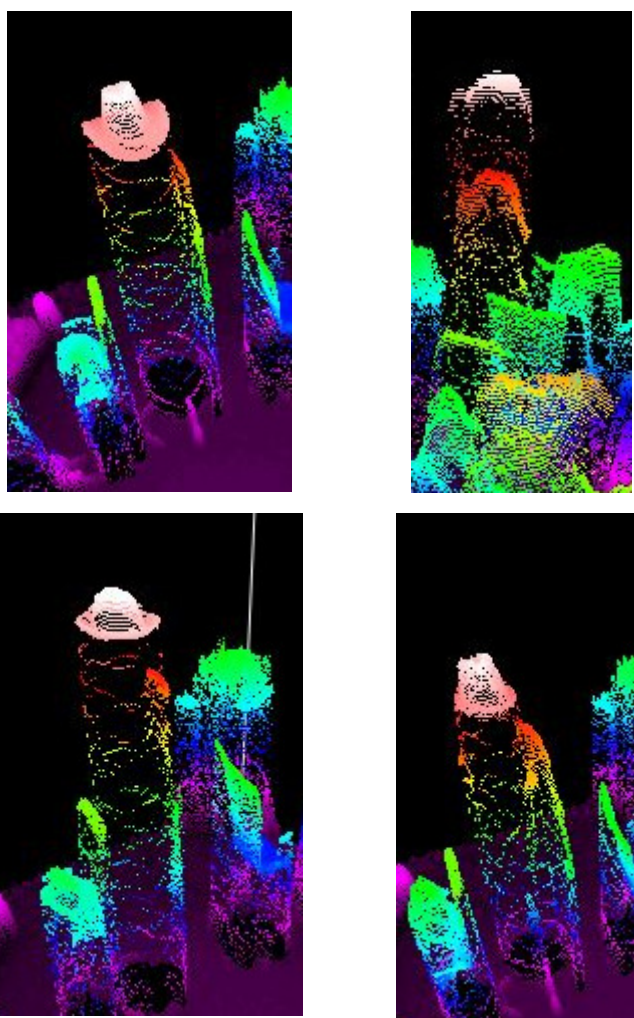
Una vez que se ha reducido el ruido de las secuencias de mapas de profundidad y se han estimado los pixeles no válidos se detectarán las personas, descartando otros objetos presentes en la escena y procederá a clasificar éstas en función de diferentes características físicas.

Las posibles características a identificar van a ser:

- Personas con pelo corto.
- Personas con pelo largo.
- Personas con pelo suelto.
- Personas con coleta.
- Personas con sombrero.
- Personas con gorra.

El presente TFM plantea el reto de identificar y clasificar sujetos con distintas características físicas.

Los seres humanos utilizamos mucha información visual para realizar esta acción, principalmente de la silueta de los sujetos y ya somos capaces de identificar a los mismos cómo se puede ver en la Figura 0-1.



*Figura 0-1 - Ejemplo de siluetas de sujetos con diferentes complementos*

La mayor dificultad que afronta este TFM es que sólo se va a disponer de la información de la silueta obtenida por una cámara ToF situada en posición cenital (ver Figura 2-3) para identificar sujetos y sus complementos.

# 1 INTRODUCCIÓN

---

El trabajo de detección y clasificación de personas en secuencias de mapas de profundidad cenitales va a desarrollarse siguiendo los pasos que se muestran en la Figura 1-1 y se detallan a continuación:

- Primeramente se pretende realizar un sistema de caracterización de individuos que sea capaz de conseguir clasificar, únicamente mediante los datos obtenidos de una cámara de tiempo de vuelo (ToF), e identificar distintos atributos.
- Una vez obtenido este sistema, realizar una base de datos con distintos sujetos de diferentes características físicas y con diferentes complementos correctamente identificados.
- Posteriormente modificar el sistema para identificar y clasificar mediante una comparación contra la base de datos obtenida al sujeto presente en la escena a examen.

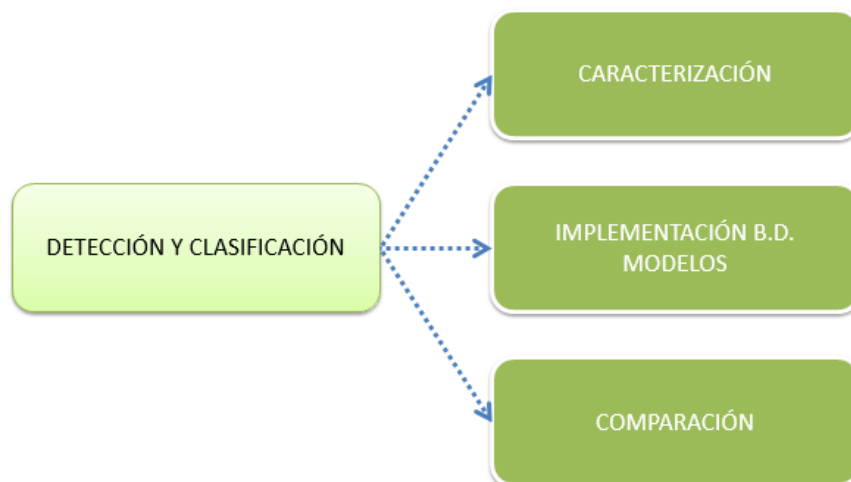


Figura 1-1 - Diagrama en bloques a realizar en el TFM.

## Medios:

En el presente TFM se van a utilizar los medios siguientes:

- Cámara Kinect 2 situada en el Espacio Inteligente, perteneciente al grupo de investigación GEINTRA de la UAH.
- Ordenador personal.
- Programas de ejemplos y bibliotecas de libre acceso o que se pueden utilizar con licencias de prueba.

---

## 2 BASES TEÓRICAS

---

### 2.1 ESTADO DEL ARTE

La clasificación de personas manteniendo la privacidad de los sujetos se encuentra en la actualidad en el punto de mira de investigadores y empresas. Este interés creciente es debido a la gran variedad de aplicaciones posibles para analizar todo tipo de información. Una opción de mantener la privacidad de las personas es el uso de cámaras ToF en posición cenital, donde resulta imposible reconocer a un individuo a partir del mapa de profundidad capturado con dicha cámara.

Los humanos disponemos de la capacidad de reconocer características de distintos sujetos desde diferentes puntos de vista, basándonos en formas, colores y proporciones, el algoritmo que se va a presentar pretende realizar esta clasificación desde un punto de vista cenital, utilizando únicamente la información del mapa de profundidad.

Existe gran cantidad de literatura en la que se puede observar que los autores han trabajado en las distintas tareas que se pretenden unificar en el presente TFM con imágenes a color, como pueden ser trabajos en los que se clasifican personas, se cuentan y se siguen sus trayectorias, pero o no se clasifican ni se sacan características de los sujetos o no se realizan utilizando únicamente cámaras de tiempo de vuelo.

18

La clasificación de personas utilizando imágenes a color según el estudio indicado en [1], donde se propone un método de reconocimiento robusto de características de los individuos, tales como: el género, corte de pelo o tipo de ropa, independientemente de grandes variaciones de puntos de vista o poses típicas de álbumes de fotos sirve de inspiración para este trabajo donde se intenta clasificar estas características partiendo desde un punto de vista superior, con sujetos dinámicos y solamente usando información de profundidad.

Para los autores de [1], la información obtenida se compara con una base de datos que posee una gran variedad de poses factibles para múltiples ocupaciones y utilizando un filtro de partículas se pueden evaluar distintas hipótesis sobre diferentes actividades realizadas y seleccionar la más consistente con la imagen grabada.

El método usado en [1] está evaluado con un conjunto de datos que contiene 10 actividades para 10 personas y consiguen rastrear y clasificar las poses con alta precisión. Este estudio aporta la idea sobre el volumen de individuos necesarios para tratar de implementar una base de datos entrenamiento con individuos perfectamente identificados y con características físicas controladas contra la que se realizará una comparación para realizar la clasificación final.

Esta clasificación final no solamente pretende diferenciar entre sujetos, sino que también se propone un método para el reconocimiento de atributos, estilo de pelo y tipos de ropa, cómo realizan en [2] a través de fotografías a color, para lograr los resultados han entrenado un clasificador para las poses y otro distinto para los atributos.

Primeramente se deberá detectar a los sujetos en la escena y en [2] se puede observar que los autores realizan la detección y el seguimiento de individuos utilizando cámaras en posición cenital.

La metodología utilizada en [2] para el seguimiento simultáneo de personas a través de datos obtenidos mediante cámaras ToF y tiene como objetivo crear un método para el seguimiento y reconocimiento de la actividad del cuerpo humano.

La detección de los individuos es el mayor problema en [3] y los autores tratan de usar los datos obtenidos de una cámara ToF junto con un modelo de movimiento humano para limitar el problema del algoritmo de seguimiento de sujetos.

En este TFM no es necesario el seguimiento simultáneo de personas debido a que se limitan las trayectorias posibles de estas mediante la utilización de barreras físicas, ya que la localización de la cámara se estima en el paso de una puerta, tampoco es necesario realizar el seguimiento individual.

Se va a proponer a continuación un algoritmo para poder clasificar personas y otras características físicas utilizando la información obtenida por una cámara de tiempo de vuelo para identificar los atributos de los sujetos, solamente usando información de profundidad.

## 2.2 DESCRIPCIÓN TEÓRICA DEL TRABAJO

Para llevar a cabo el proceso de clasificación, se van a establecer dos procesos, uno de entrenamiento que denominaremos “proceso off-line” y otro donde se realiza la clasificación en tiempo real que denominaremos “proceso on-line”. En la Figura 2-1 se muestra un esquema general del proceso de clasificación de complementos de personas que se ha seguido para la realización del TFM.

Estos procesos no se realizan en paralelo. El proceso previo para extraer las características de los sujetos objetos de estudio y establecer una base de datos con los modelos es el proceso off-line. En el proceso on-line cada nuevo objeto detectado en la escena, es comparado con cada uno de los modelos obtenidos en el proceso off-line para determinar a cuál de ellos se asemeja más.

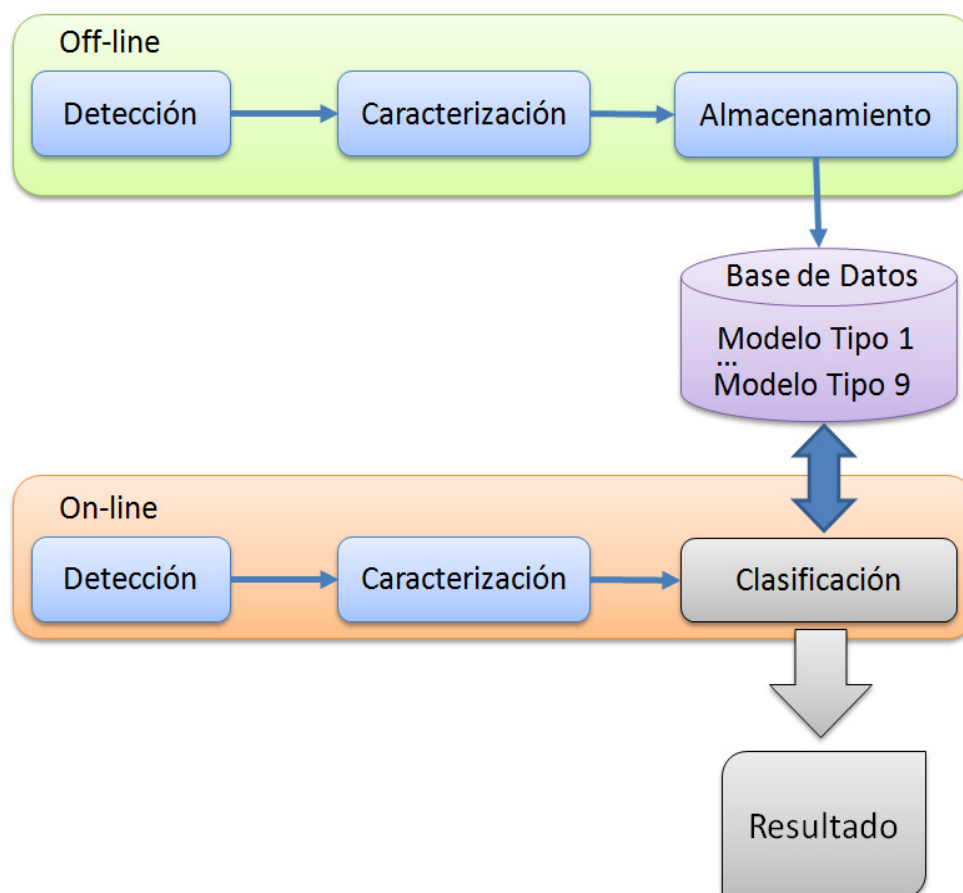


Figura 2-1 - Descripción proceso de clasificación.

Tanto el proceso off-line como el on-line tienen similitudes ya que en ambos se debe detectar la presencia de un sujeto en la escena y se tiene que caracterizar, en el primer caso, proceso offline, para tener una referencia conocida y en el segundo para poder comparar e identificar con los modelos.



### 2.2.1 Proceso Off-line

Un esquema en bloques del proceso off-line se muestra en la Figura 2-2. El mismo está compuesto por las siguientes etapas:

- **Detección** de un posible sujeto en la escena.
- **Caracterización** del sujeto en función de sus características físicas o de sus complementos.
- **Almacenamiento** de los vectores de características de los sujetos para poder contrastar contra ellos las características obtenidas de los futuros sujetos en el proceso on-line.

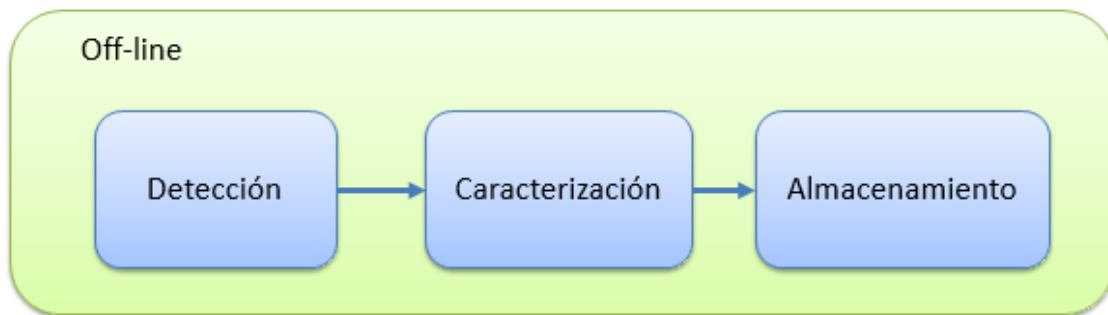


Figura 2-2 - Descripción proceso off-line.

Para desarrollar la detección de los sujetos en la escena, teniendo en cuenta el uso de cámaras ToF, podemos establecer que sólo los puntos de altura máxima detectados, pueden ser candidatos para tratar la información de la existencia de un posible sujeto, cuyo máximo es la cabeza, para a continuación extraer las características dentro de la zona de vecindad de dicho máximo.

A continuación, se va a proceder a describir los apartados de detección y caracterización que son comunes a ambos procesos.

#### 2.2.1.1 Detección

Se trata de aclarar la lógica utilizada para localizar sujetos en una escena utilizando una cámara de tiempo de vuelo colocada en posición cenital y usando la información proveniente de su sensor.

##### 2.2.1.1.1 Percepción de sujetos

Para realizar la detección se deben tener en cuenta diversos factores asociados al tipo de cámara y su ubicación. Las cámaras de tiempo de vuelo se sitúan en posición cenital y obtienen información de la altura de los puntos correspondientes a los objetos que se encuentren en su área de visión. Y el número de puntos obtenidos depende de la resolución del sensor de la cámara.

Teniendo en cuenta la posición de la cámara, en nuestro caso situada perpendicularmente al suelo, los datos se obtienen estableciendo el eje de coordenadas en el centro del objetivo de la cámara y se establecen los ejes X, Y y Z colocados según se indica en la Figura 2-3.

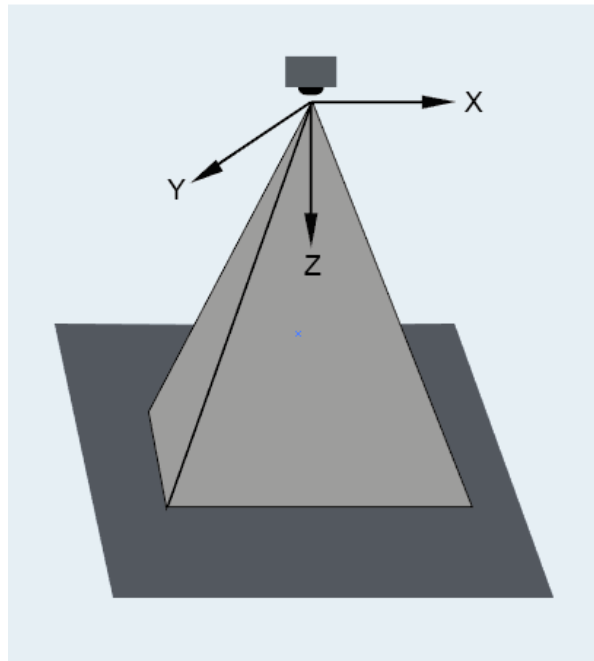


Figura 2-3 - Posición de la cámara.

De tal manera que la distancia de cualquier punto obtenido hasta el origen de coordenadas viene dado por la ecuación (1).

$$d = \sqrt{x^2 + y^2 + z^2} \quad (1)$$

La altura en la que se encuentra un punto respecto al suelo viene dada por la ecuación (2) que sigue de la siguiente forma:

$$h_{punto} = h_{camara} - z_{punto} \quad (2)$$

Para cada punto obtenido por la cámara, se tienen las coordenadas 3D del punto y su distancia al origen de coordenadas.

En este trabajo sólo se va a trabajar con el mapa de profundidad formado por una matriz de medidas de altura, que será el conjunto de medidas en el eje Z por cada pixel obtenido por la cámara, si tenemos una resolución de  $M \times N$  píxeles, se obtendrá una matriz de alturas con las características indicadas por la ecuación (3).

$$\text{Matriz de alturas medias } H = \begin{bmatrix} h_{1,1} & \cdots & h_{1,N} \\ \vdots & \ddots & \vdots \\ h_{M,1} & \cdots & h_{M,N} \end{bmatrix} \quad (3)$$

Dónde cada valor de la matriz, desde  $h_{1,1}$  hasta  $h_{M,N}$  viene de la ecuación (2).

No todos los valores obtenidos dan información relevante para el estudio propuesto en este TFM, por lo que se va a limitar la información mediante una estimación de una región de interés.

Para clasificar sujetos mediante la información obtenida de la cámara ToF se estima suficiente con los datos comprendidos entre los hombros y la cabeza de los sujetos. La cabeza y los hombros se han definido cómo las áreas de vecindad más cercanas a los puntos máximos detectados que representan una superficie de interés lo suficientemente grande para poder ser dichas zonas del cuerpo.

Se establece que la altura a analizar no supere una distancia de  $h_{interés} = 40$  cm para extraer las diferentes características físicas necesarias pertenecientes a cabeza, cuello y hombros, como se indica en la Figura 2-4.

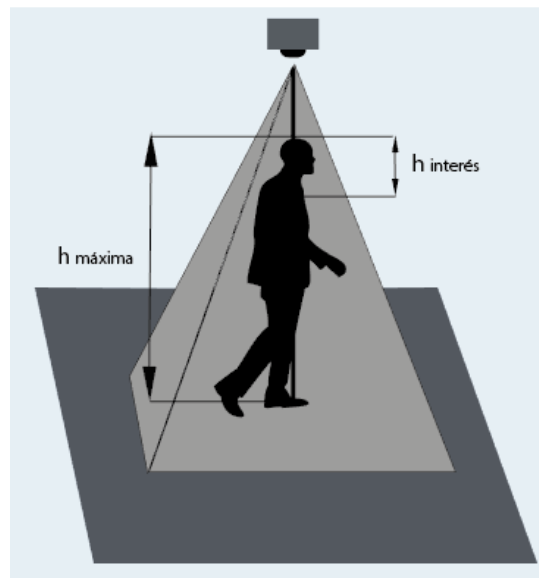


Figura 2-4 - Altura de interés.

Una vez detectado un máximo, se incluyen en la ROI todos los puntos que queden comprendidos dentro de una altura determinada correspondiente a la ecuación (4).

$$h_{max} - h_{p,q} \leq h_{interés} = 40 \text{ cm} \quad (4)$$

#### 2.2.1.1.2 Detección de máximos

El algoritmo utilizado para seleccionar los valores máximos que pueden dar la información correspondiente a la existencia de una persona en la escena está basado en el trabajo [4][4][4] y podría ser descrito como sigue:

- I. Creación de subregiones cuadradas de  $D \times D$  píxeles cómo se describe en la ecuación (5).
- II. Si la cámara presenta una resolución espacial de  $M \times N$ , el número de subregiones será:

$$S_X = \frac{M}{D} ; S_Y = \frac{N}{D} \quad (5)$$

Tanto  $M$  como  $N$  son valores de la cámara (resolución de la cámara), pero  $D$  proviene de la ecuación (6) y es un valor que depende de varios factores, tales como:

- La altura de la cámara (que representa la altura a la que se encuentra el eje de coordenadas respecto al suelo).
- La altura mínima de personas que se desea detectar.
- El área mínima que se estima que ocupa la parte superior de una cabeza de persona.

En este caso el tamaño de las subregiones  $D$  viene dado por:

$$D = \frac{d_f}{t_p} \frac{t_{mc}}{(h_{cámara} - h_{mínima})} \quad (6)$$

Siendo:

- $d_f$  = distancia focal de la cámara (dato del fabricante).
- $t_p$  = tamaño del pixel (dato del fabricante).
- $t_{mc}$  = tamaño medio cabeza.
- $h_{cámara}$  = altura del eje de coordenadas respecto del suelo.
- $h_{mínima}$  = altura mínima de los sujetos que se va a tener en cuenta como máximo.

Para la cámara Kinect 2 y el entorno de desarrollo, se obtienen los siguientes valores:

$$\frac{d_f}{t_p} = 365.77 \quad h_{cámara} = 3400 \text{ mm} \quad h_{mínima} = 1500 \text{ mm} \quad t_{mc} = 130 \text{ mm} \quad (7)$$

De tal manera se garantiza un número suficiente de píxeles  $D = 25$  píxeles por cada subregión cómo se puede ver en la Figura 2-5 (Subregiones cuadradas de 25 x 25 píxeles).

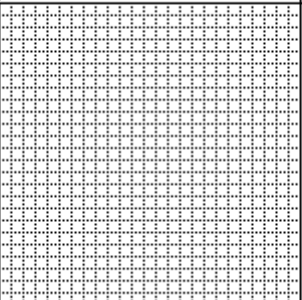
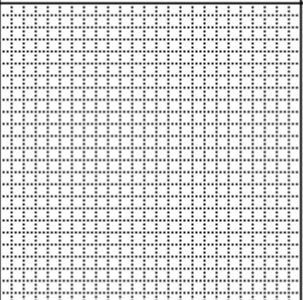
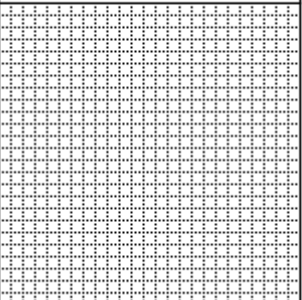
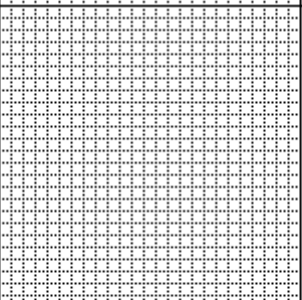
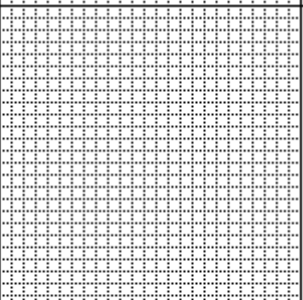
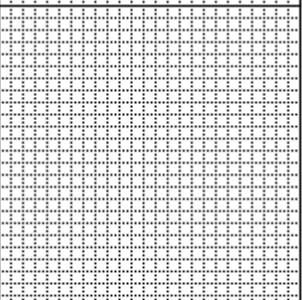
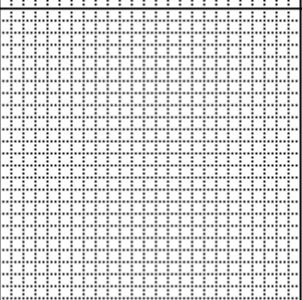
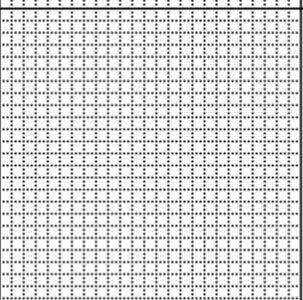
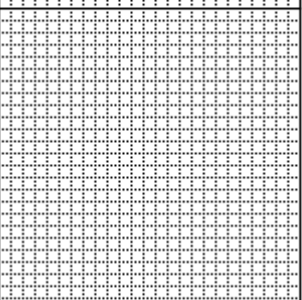
	1 - 25	26 - 50	51 - 75	...
1 - 25				...
26 - 50				...
51 - 75				...
⋮	⋮	⋮	⋮	↘

Figura 2-5 - Píxeles agrupados por subregiones

III. Obtención de máximos por subregiones en forma de matriz de alturas, para ello se debe cumplir:

$$h_{\text{mínima}} \leq h_{\text{max}}^{\text{Subregión}} \geq h_{\text{max}}^{\text{Subregión}} \quad (8)$$

Dónde:

- $h_{\text{mínima}}$  = altura mínima de los sujetos que se va a tener en cuenta como máximo.
- $h_{\text{max}}^{\text{Subregión}}$  = máximo de la subregión evaluado.
- $h_{\text{max}}^{\text{Subregión}}$  = máximo previo de la subregión.

Formando una matriz de máximos por subregiones de la forma:

$$H_{\text{max}}^{\text{Subregión}} = \begin{bmatrix} h_{1,1}^{\text{Subregión}} & \dots & h_{1,S_N}^{\text{Subregión}} \\ \vdots & \ddots & \vdots \\ h_{S_M,1}^{\text{Subregión}} & \dots & h_{S_M,S_N}^{\text{Subregión}} \end{bmatrix} \quad (9)$$

	1 - 25	26 - 50	51 - 75	...
1 - 25	$h_{1,1}$	$h_{1,2}$	$h_{1,3}$	...
26 - 50	$h_{2,1}$	$h_{2,2}$	$h_{2,3}$	...
51 - 75	$h_{3,1}$	$h_{3,2}$	$h_{3,3}$	...
⋮	⋮	⋮	⋮	⋮

Figura 2-6 - Matriz de máximos por subregiones

Dado que se pueden dar subregiones con máximos muy próximas, debido a la fisionomía humana, los complementos de vestimenta, la posición de la cámara y la resolución de la misma, se sustituirán todas las subregiones próximas con un margen de vecindad de orden 2 por el máximo de mayor valor de todas las subregiones incluidas en este área y se le asignan las coordenadas de la subregión que esté más cerca de la centroide de la mayor medida.

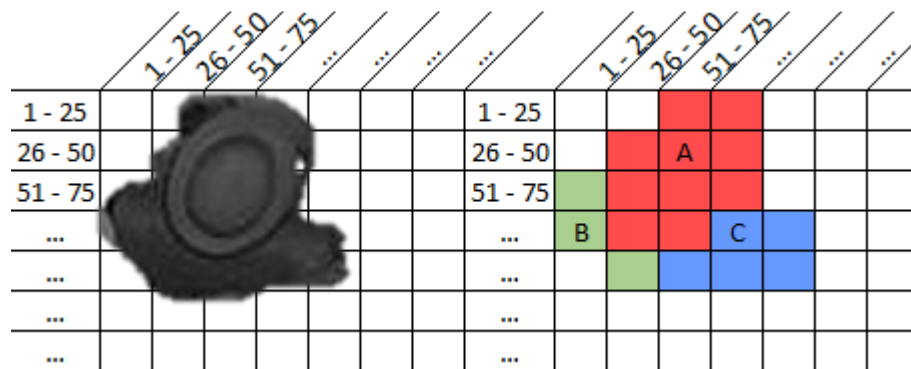


Figura 2-7 - Agrupación de máximos de subregiones colindantes

Cómo se puede ver en la Figura 2-7, el sombrero tiene el máximo valor en la posición dónde se encuentra A y se le asignan las coordenadas de la subregión con máximo más cercana al centroide de todas las subregiones con orden de vecindad 2 al igual que pasa en B y C con los hombros.

2.2.1.1.3 Ajuste del área de interés

Se deben establecer las subregiones que forman parte del área de interés que comprenden cabeza, cuello y hombros. Para ello se deben buscar los límites de cada individuo en distintas zonas de vecindad en direcciones radiales.

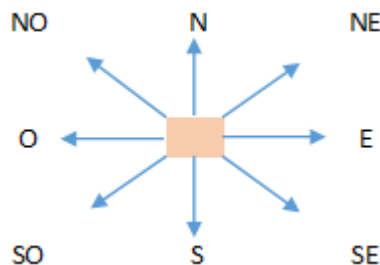


Figura 2-8 - Direcciones radiales de búsqueda

Se van a seguir 8 direcciones radiales, desde el punto de interés, N, S, E, O, NE, NO, SE y SO y los siguientes requisitos para incluirlos en la ROI:

- La subregión en la que se encuentra el punto máximo siempre pertenece al ROI.
- Las subregiones de la zona de vecindad de nivel 1 siempre pertenecen a la ROI, simplemente por la morfología de las personas y la resolución de la cámara.
- La altura máxima de la subregión debe ser mayor q la  $h_{\min}$ .
- Para todas las direcciones N, NE, NO, S, SE, SO, E y O se analizarán, como mínimo, dos zonas de vecindad
- La altura máxima de la subregión tiene que ser inmediatamente inferior a la subregión previa.

### 2.2.1.2 Creación del vector de características

Una vez definidos los datos que se van a extraer de toda la información obtenida del sensor de la cámara, se debe proceder a tratarla para crear una base de datos con vectores de características correspondientes a varios ejemplos de sujetos con distintas características físicas.

#### 2.2.1.2.1 Extracción de características

Se pretende clasificar las características físicas de los sujetos presentes en la escena teniendo en cuenta la morfología presente en la zona comprendida entre la cabeza y los hombros, para ello se va a generar un vector de características que comprenda la densidad de los puntos asociados a cada superficie.

28

Se han definido 6 niveles de altura en la  $h_{interés}$  para la ROI obtenida.

Cinco de estas características se corresponden con las superficies detectadas por la cámara entre la parte superior de la cabeza y los hombros, la sexta es una relación entre el diámetro mayor y menor de la superficie superior detectada.

#### 2.2.1.2.2 Asignación de los píxeles obtenidos

Para los valores obtenidos dentro de la ROI y de la  $h_{interés}$  se cuantifican el número de píxeles obtenidos cada  $X$  cm, obteniendo un número de agrupaciones de:

$$N_a = \frac{h_{interés}}{X} \quad (10)$$

En este estudio, se fija  $X = 2$  cm, teniendo fijada la  $h_{interés} = 40$  cm, se va a obtener un número de  $N_a = 20$  valores.

Se cuenta el número de píxeles existentes según el sensor ToF cada  $X$  cm y se obtiene un valor de densidad de píxeles por cada valor de altura, cómo se puede ver indicado en la columna de la izquierda de la Figura 2-9.



A partir de esta información, se conforma un vector de características de la siguiente forma: En caso de obtener un máximo aislado se desprecia, dado que puede ser un ruido, en caso de obtener una ROI con la suficiente densidad de puntos, se unifican los valores de densidad de píxeles cada X cm en grupos de 3, obteniendo los 3 primeros componentes del vector correspondientes a la cabeza, a partir del 3 componente se descartan los siguientes hasta obtener una mayor densidad de píxeles que corresponda a los hombros.

Para completar el vector de características, hacen falta las dos siguientes componentes darán la información correspondiente a los hombros y se establecerán en la siguiente región con mayor densidad de puntos correspondiente a la zona de los hombros cómo se puede observar en la columna de la derecha de la Figura 2-9 y el sexto valor del vector será una relación geométrica correspondiente a la parte superior de la cabeza.

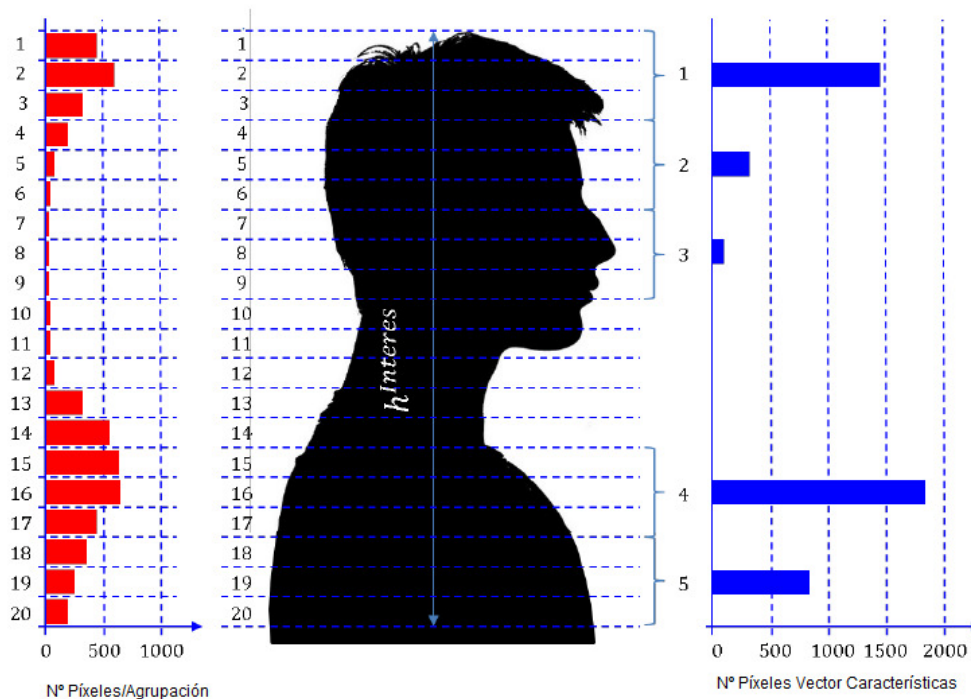


Figura 2-9 - Densidad de pixeles por nivel.

Fuente: Elaboración propia a partir de[4].

### 2.2.1.2.3 Normalización

El número de píxeles asociados a las agrupaciones depende de la altura de los sujetos por lo que es necesario realizar una normalización para uniformar los datos obtenidos.

Se plantea establecer una relación cuadrática entre la altura del sujeto y el número de puntos detectados por la cámara en la agrupación superior.

$$R = a * (h_{max})^2 + b * h_{max} + c \tag{11}$$

Donde **a**, **b** y **c** son los coeficientes que se van a estimar usando un algoritmo de métodos numéricos (Levenberg-Marquardt). El vector normalizado final se obtiene dividiendo cada componente de las 5 anteriormente mencionadas por el valor obtenido en la ecuación (11) que relaciona la densidad de píxeles con la altura del sujeto.

El sexto componente del vector de características es una relación entre el diámetro mayor y el diámetro menor de la superficie que forma la parte superior de la cabeza.

Se localizan los dos ejes, el eje mayor y el eje menor de dicha franja de píxeles y se establece la relación de excentricidad que forma el último componente del vector de características.

### 2.2.1.3 Creación de la base de datos

La creación de una base de datos contra la que comparar los vectores de características que se van a obtener de los sujetos a estudio en el proceso on-line es imprescindible para poder clasificarlos.

La base de datos va a disponer de los vectores identificados procedentes de las secuencias modelo, por lo que se verifica con total seguridad la correspondencia de los vectores con los complementos.

Una vez obtenidos los vectores de características durante el proceso off-line de todos los sujetos que van a servir como modelo, se almacenan en dicha base de datos y se comparan durante el proceso on-line con el vector de características obtenido a partir del sujeto que esté formando parte de la escena en un determinado momento.

Se estima necesaria la creación de una base de datos que contenga información sobre varones y hembras de distintas alturas y características físicas (distintos peinados) y la utilización de distintos complementos que puedan modificar los datos obtenidos por la cámara (gorras, sombreros...).

### 2.2.2 Proceso On-line

En el proceso on-line, están presentes varias etapas indicadas en la Figura 2-10, se puede observar que dos de los bloques son idénticos a los dos del proceso off-line y ya han sido descritos en los puntos anteriores 2.2.1.1 y 2.2.1.2 y dos bloques nuevos que implican:

- **Identificación**, proceso en el que se compara la caracterización del sujeto presente en la escena con las almacenadas como modelo.
- **Resultado** en función de la desviación existente entre los vectores de características obtenidos del sujeto de la escena con los existentes en la base de datos de modelos.

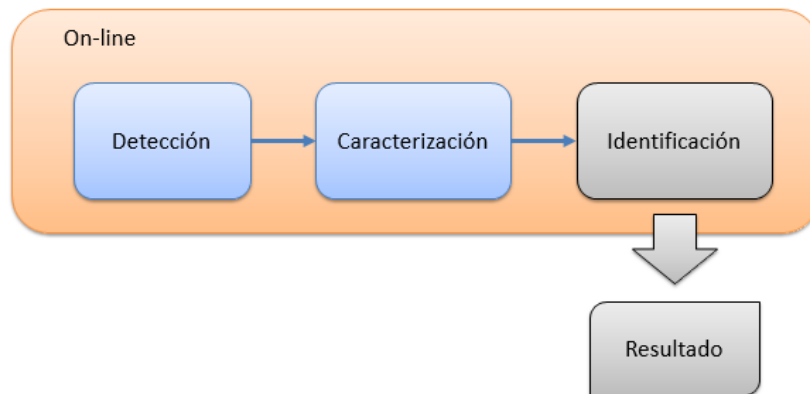


Figura 2-10 - Descripción proceso on-line.

31

En la Figura 2-1 se puede apreciar más destacado el bloque clave que implica cómo se identifican los valores obtenidos de los nuevos sujetos contra los valores previamente almacenados de los sujetos que sirvieron de modelo para cada una de las clases disponibles.

### 2.2.2.1 Identificación

Para poder diferenciar y clasificar sujetos en función de los atributos obtenidos con la cámara, se han definido en la base de datos una serie de clases de sujetos en función de diferentes características físicas.

#### 2.2.2.1.1 Clasificador de sujetos.

##### I. Off-line:

A partir de los ( $N_v$ ) vectores de entrenamiento obtenidos, se generan las matrices de transformación, se calcula la media y la scatter matrix para cada modelo:

$$\bar{\mathbf{W}}\alpha = \frac{1}{N_v} \sum_{i=1}^{N_v} \mathbf{W}\alpha_i \quad (12)$$

$$\mathbf{V}_{T\alpha} = \sum_{l=1}^{N_v} (\mathbf{W}\alpha_l - \bar{\mathbf{W}}\alpha)(\mathbf{W}\alpha_l - \bar{\mathbf{W}}\alpha)^T$$

Cada matriz de cada clase  $\alpha$  están formadas por los autovectores asociados a los autovalores mayores de la  $\mathbf{V}_{T\alpha}$  (Scatter Matrix)

Se han elegido los tres autovectores asociados a los tres autovalores mayores siempre y cuando el error cuadrático medio residual normalizado (RMSE) sea mayor al 90%, de esta manera:

$$RMSE = \frac{\sum_{j=m+1}^6 \gamma_{\alpha j}}{\sum_{j=1}^6 \gamma_{\alpha j}} > 0.9 \quad (13)$$

Con  $m = 3$

## II. On-line:

Para cada clase  $\alpha$ , se proyecta en el espacio transformado la diferencia entre el vector de características  $\mathbf{W}$  y el vector de medias de la clase  $\bar{\mathbf{W}}\alpha$ . como se puede ver indicado en (12).

$$\mathbf{V}\alpha = \mathbf{W} - \bar{\mathbf{W}}\alpha \quad (14)$$

Dónde:

- $\mathbf{V}\alpha$ = Diferencia entre vector obtenido y vector medio de la clase.
- $\mathbf{W}$  = Vector obtenido del sujeto de la escena.
- $\bar{\mathbf{W}}\alpha$ = Vector medio de clase  $\alpha$ .

Este vector  $\mathbf{W}$  se proyecta en el espacio transformado como se indica en (15).

$$\mathbf{X} = \mathbf{U}_\alpha^T (\mathbf{V}\alpha) \quad (15)$$

Posteriormente se recupera:

$$\widehat{\mathbf{V}}\alpha = \mathbf{U}_\alpha(\mathbf{X}) \quad (16)$$

Se define que la distancia Euclidea entre  $\mathbf{V}\alpha$  y  $\widehat{\mathbf{V}}\alpha$  se llama error de reconstrucción  $E_R$  y se obtiene para cada clase modelo.

Se establece un umbral mínimo que en caso de no superarse descarta la correspondencia del vector de características con el modelo de una persona.

Este umbral se establece experimentalmente por:

$$h_{\alpha} = \overline{E_{\alpha}} + 3\delta_{\alpha} \quad (17)$$

Dónde:

- $\overline{E_{\alpha}}$  = Valor medio del error de reconstrucción.
- $\delta_{\alpha}$  = Desviación estándar del error.

El umbral quedaría cómo se indica en (18):

$$\mathbf{W} \text{ se considera si cumple que } E_{\alpha} = [\mathbf{V}\alpha - \widehat{\mathbf{V}\alpha}] \leq h_{\alpha} \quad (18)$$

## 3 DESCRIPCIÓN EXPERIMENTAL

### 3.1 ADQUISICIÓN DE SECUENCIAS DE MAPAS DE PROFUNDIDAD

Para capturar las secuencias de mapas de profundidad cenitales se va a utilizar la cámara ToF del módulo *Kinect 2* de *Microsoft*, colocada en posición cenital. Los principales motivos por los que se usa esta cámara son: la posibilidad de usar herramientas de desarrollo gratuitas, controladores gratuitos de Microsoft y el hecho de ser un hardware de precio asequible.

Para poder utilizar esta cámara para la adquisición de datos, se deben seguir los pasos descritos por la Figura 3-1 instalar en el ordenador los drivers y crear un programa que permita archivar las secuencias de datos deseadas. Para crear un programa que use las librerías de la cámara, al ser de Microsoft, se debe recurrir al programa Microsoft Visual Studio y generar el interfaz de grabación final.

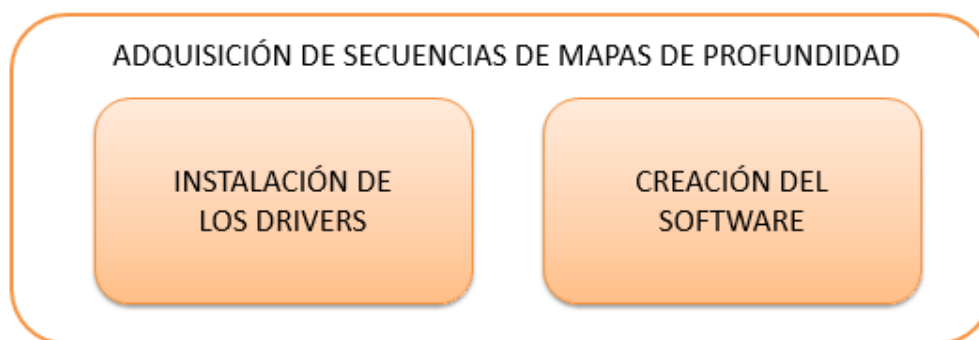


Figura 3-1 - Proceso de obtención de secuencias de mapas de profundidad.

#### 3.1.1 Procedimiento de configuración del entorno de trabajo.

Principalmente se debe trabajar con un ordenador personal que posea las características adecuadas a las especificaciones requeridas por la cámara, como son:

- I. Sistemas operativos soportados:
  - Windows Embedded 8 Standard, Windows 8, Windows 8.1 o versiones superiores.

II. Hardware necesario:

- Procesador de 64-bit (x64).
- 2 Gb de memoria RAM (o más).
- 4 Gb de memoria ROM (o más).
- Procesador de doble núcleo físico de 3.1 GHz (dos núcleos lógicos por cada físico) o superior.
- Controlador USB 3.0 dedicado.
- Adaptador gráfico que soporte DX11.

III. Software necesario:

- Visual Studio 2012 o Visual Studio 2013.
- .NET Framework 4.0.
- Microsoft Kinect for Windows SDK v2.0.

### 3.1.1.1 Instalación de los drivers

El conjunto de librerías y ejemplos que ofrece de manera totalmente libre Microsoft para el desarrollo de software que utilice su cámara se llama “Kinect for Windows SDK v2.0” y es de libre uso y están disponibles en el sitio de descargas oficial de la compañía, el “Microsoft Download Center” [5]. Mediante el uso de estas librerías se consigue tener acceso a la cámara.

35

A parte de los drivers, este software incluye interfaces de programación de aplicaciones (API) y códigos de ejemplo que utilizan diferentes características de la cámara. De entre todos los disponibles cabe destacar el ejemplo “Depth Basics” del que se va a partir para realizar el presente proyecto.

El ejemplo usa la cámara ToF de la Kinect 2. Partiendo de una posición cenital muestra los valores de profundidad obtenidos, consta con un pequeño interfaz gráfico que dispone de:

- Área de visualización dónde se pueden ver en tiempo real las secuencias de profundidad obtenidas por la cámara.
- Botón para realizar capturas de pantalla de las secuencias.
- Cuadro de texto dónde se obtiene la ruta en la que se ha grabado el archivo en el ordenador.

Presenta una amplia zona para representar las imágenes con las dimensiones 512 x 424 y un cuadro de salida de texto acompañado por el botón “Screenshot” que indica la dirección dónde se ha guardado el archivo generado al pulsar el botón.

### 3.1.2 Software de adquisición de datos de profundidad.

Una vez instalados los drivers y habiendo comprobado el correcto funcionamiento de la cámara con los diferentes ejemplos disponibles, se procederá a modificar el ejemplo “Depth Basics”, ya mencionado, para conseguir las funcionalidades deseadas de grabación de los datos.

Tratándose de una cámara de Microsoft y de un ejemplo programado en C++ para poder editarlo Microsoft recomienda instalar el programa Microsoft Visual Studio Community de la página web oficial de la compañía. [6]

Microsoft Visual Studio es un entorno de desarrollo integrado (IDE) que permite realizar las modificaciones requeridas en el ejemplo seleccionado para aportarle las funcionalidades necesarias, este programa también es de libre distribución.

#### 3.1.2.1 Edición del programa

Principalmente se van a realizar dos modificaciones a este ejemplo, la más visual es la modificación de la interfaz gráfica, pero es necesario realizar modificaciones software para acomodar el ejemplo a nuestras necesidades.

Los cambios en la interfaz gráfica radican en la inclusión de botones para grabar secuencia y parar secuencia y la utilización del cuadro de texto para poner nombre a los archivos a almacenar.

---

36

Los botones, que se pueden ver en la Figura 3-2, funcionarán de la siguiente manera:

Una vez indicado un nombre de archivo en el cuadro de texto de entrada, se pulsa el botón “SaveVideo”, se crea un archivo con el nombre indicado en el cuadro de texto y se comienza a grabar la secuencia de datos de profundidad obtenidos por la cámara, en el fichero con el nombre indicado en el cuadro de texto, hasta recibir una pulsación en el botón “StopVideo” con la que se cierra el archivo y se deja al programa en espera de la próxima pulsación.

Con la modificación sugerida del cuadro de texto se pretende conseguir la posibilidad de asignar nombres personalizados para cada archivo generado de tal manera que permita al usuario identificar las características, trayectoria seguida y prendas usadas por el sujeto grabado directamente con el nombre del archivo.



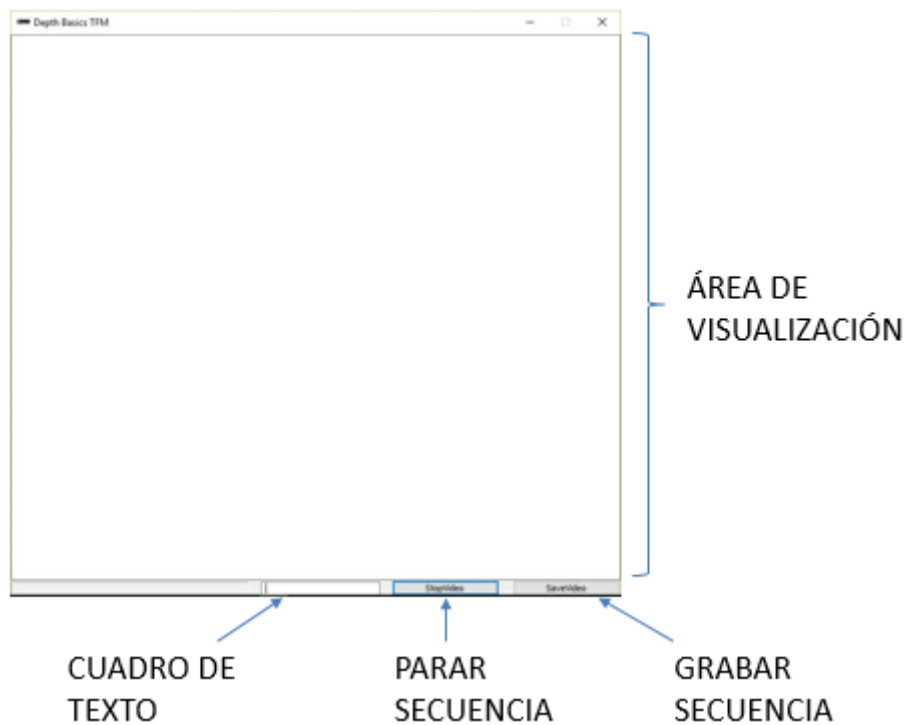


Figura 3-2- Interfaz gráfica modificada.

### 3.2 CREACIÓN DE LA BASE DE DATOS

Una vez creado el programa que gestiona de la manera deseada la información proporcionada por la cámara, se debe proceder a la creación de las bases de datos de secuencias de medidas de profundidad.

La propuesta final de este proyecto es el análisis en tiempo real de los individuos que atraviesan una escena concreta, pero debido a la imposibilidad de recibir secuencias de mapas de profundidad válidos en los momentos de desarrollo del trabajo, se genera una base de datos de secuencias de medidas de profundidad para poder recurrir a ella en cualquier momento.

Para trabajar con los mapas de profundidad, se tiene que gestionar la información obtenida por el sensor de la cámara, para ello se crea un nuevo tipo de archivo que tendrá los datos de profundidad obtenidos.

La extensión del archivo se va a definir como **“\*.z16”** debido al tipo de datos a archivar, medidas en el eje *Z* y de 16 bits, que es el tipo de datos procedentes de la cámara (valores de la coordenada *Z* en *mm*, representados en enteros de 16 bits).

En la creación de la base de datos es necesario tener en cuenta una serie de factores que influyen tanto en la validez de los resultados que se obtengan utilizando la misma, como en la facilidad del manejo de los datos almacenados. Los factores que se han tenido en cuenta son el entorno y la población a estudiar, la codificación del nombre de las secuencias y la diversidad de las secuencias, a continuación se describen en detalle cada uno de ellos.

### 3.2.1 Entorno y población a estudio

La población que se va a utilizar es de 9 sujetos, cada uno con características físicas distintas que según el estudio realizado en [7] garantiza una población suficiente para comprobar el funcionamiento del trabajo, estos 9 sujetos usarán distintos tipos de complementos.

La localización de la cámara, como ya se ha mencionado en la Introducción, obliga a simular la existencia de una puerta mediante la incursión de obstáculos en la escena para delimitar el área de estudio a las características de un teórico pasillo.

### 3.2.2 Codificación del nombre de las secuencias

La codificación de los archivos corresponde al siguiente esquema:

**“SUJ\_(Número de sujeto) + \_SEC\_(Número de identificación secuencia)  
+ \_COMP\_(Número de identificación del complemento).z16”**

38

Sujeto	SUJ_
Persona 1	1
Persona 2	2
Persona 3	3
Persona 4	4
Persona 5	5
Persona 6	6
Persona 7	7
Persona 8	8
Persona 9	9

*Tabla 1 - Codificación de sujetos*

<b>Secuencia</b>	<b>_SEC_</b>
Izquierda a derecha	1
Derecha a izquierda	2

*Tabla 2 - Codificación de secuencia*

<b>Complemento</b>	<b>_COMP_</b>	<b>Correspondencia</b>
Sombrero tipo 1	1	Barriero
Sombrero tipo 2	2	Chulapo
Pelo tipo 1	3	Coleta
Pelo tipo 2	4	Corto
Sombrero tipo 5	5	Gorra
Pelo tipo 3	6	Liso
Sombrero tipo 7	7	Pamela
Sombrero tipo 8	8	Pirata
Sombrero tipo 9	9	Vaquero

*Tabla 3 - Codificación de complementos*

39

**Ejemplo de codificación:**

Sujeto	Persona 5	5
Secuencia	Izquierda a derecha	1
Complemento	Sombrero tipo 9	9

*Tabla 4 - Ejemplo de codificación.*

Quedando el archivo cómo: **SUJ\_5\_SEC\_1\_COMP\_9.z16**

Se han obtenido un total de **106** archivos de vídeo distintos, con un tamaño total de 5,78 Gb.

### 3.2.3 Diversidad de secuencias

Se van a realizar distintas grabaciones con 9 sujetos con diferentes características físicas, de distintos sexos y con distintos cortes de pelo.

A su vez, se han elegido 6 complementos diferentes que sumados a la variedad de los sujetos grabados garantizan una diversidad suficiente.

Los complementos utilizados han sido:



Figura 3-3 - Sombrero tipo 1

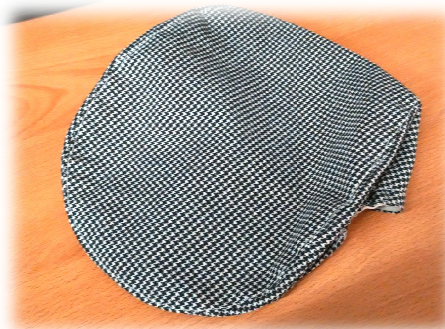


Figura 3-4 - Sombrero tipo 2



Figura 3-5 - Sombrero tipo 5



Figura 3-6 - Sombrero tipo 7



Figura 3-7 - Sombrero tipo 8



Figura 3-8 - Sombrero tipo 9

### 3.2.4 Base de datos de secuencias

Según el principio de funcionamiento indicado en la Descripción teórica del trabajo se van a establecer dentro de una base de datos dos tipos de secuencias distintas: secuencias de entrenamiento y secuencias de pruebas.

En el Proceso Off-line, se van a establecer una serie de secuencias de mapas de profundidad de distintos sujetos con diferentes complementos como secuencias de entrenamiento que se van a almacenar en una base de datos y contra las que se van a comparar las secuencias de prueba de sujetos posteriores.

Para seleccionar las secuencias de entrenamiento se intenta garantizar la mayor pluralidad posible, teniendo en cuenta los complementos y las características físicas de los sujetos.

#### I. Secuencias de entrenamiento.

Se seleccionan las siguientes secuencias indicadas en la Tabla 6 como secuencias de entrenamiento:

Nombre de la secuencia:	Tipo de sujeto
SUJ_1_SEC_1_COMP_6.z16	Mujer pelo liso.
SUJ_2_SEC_1_COMP_4.z16	Hombre pelo corto.
SUJ_7_SEC_1_COMP_2.z16	Hombre sombrero chulapo.
SUJ_1_SEC_2_COMP_3.z16	Mujer pelo en coleta.
SUJ_5_SEC_2_COMP_8.z16	Hombre sombrero pirata.
SUJ_6_SEC_1_COMP_7.z16	Hombre sombrero pamea.
SUJ_7_SEC_2_COMP_5.z16	Hombre gorra.
SUJ_8_SEC_2_COMP_9.z16	Hombre sombrero vaquero.
SUJ_9_SEC_1_COMP_1.z16	Mujer sombrero barriero.

*Tabla 5 - Secuencias seleccionadas como archivos modelo*

#### II. Secuencias de pruebas.

Se van a grabar secuencias para pruebas con diferentes sujetos y numerosos complementos.

Dejando en la Tabla 6 como secuencias de pruebas las siguientes:

SUJ_1_SEC_1_COMP_1	SUJ_6_SEC_2_COMP_2	SUJ_7_SEC_1_COMP_7
SUJ_2_SEC_2_COMP_1	SUJ_1_SEC_1_COMP_3	SUJ_7_SEC_2_COMP_7
SUJ_2_SEC_1_COMP_1	SUJ_3_SEC_2_COMP_3	SUJ_5_SEC_2_COMP_8
SUJ_1_SEC_2_COMP_1	SUJ_3_SEC_1_COMP_3	SUJ_1_SEC_2_COMP_8
SUJ_3_SEC_2_COMP_1	SUJ_3_SEC_2_COMP_3	SUJ_1_SEC_1_COMP_8
SUJ_3_SEC_1_COMP_1	SUJ_3_SEC_1_COMP_3	SUJ_2_SEC_2_COMP_8
SUJ_9_SEC_2_COMP_1	SUJ_7_SEC_1_COMP_4	SUJ_2_SEC_2_COMP_8
SUJ_5_SEC_2_COMP_1	SUJ_6_SEC_2_COMP_4	SUJ_2_SEC_1_COMP_8
SUJ_5_SEC_1_COMP_1	SUJ_2_SEC_2_COMP_4	SUJ_3_SEC_2_COMP_8
SUJ_7_SEC_2_COMP_1	SUJ_8_SEC_2_COMP_4	SUJ_3_SEC_1_COMP_8
SUJ_6_SEC_1_COMP_1	SUJ_8_SEC_2_COMP_4	SUJ_8_SEC_2_COMP_8
SUJ_6_SEC_2_COMP_1	SUJ_6_SEC_1_COMP_4	SUJ_7_SEC_1_COMP_8
SUJ_1_SEC_2_COMP_2	SUJ_8_SEC_1_COMP_4	SUJ_7_SEC_2_COMP_8
SUJ_1_SEC_1_COMP_2	SUJ_5_SEC_2_COMP_4	SUJ_6_SEC_1_COMP_8
SUJ_2_SEC_2_COMP_2	SUJ_5_SEC_1_COMP_4	SUJ_6_SEC_2_COMP_8
SUJ_2_SEC_1_COMP_2	SUJ_7_SEC_2_COMP_4	SUJ_6_SEC_2_COMP_9
SUJ_3_SEC_2_COMP_2	SUJ_1_SEC_1_COMP_5	SUJ_7_SEC_1_COMP_9
SUJ_3_SEC_1_COMP_2	SUJ_8_SEC_1_COMP_5	SUJ_2_SEC_2_COMP_9
SUJ_5_SEC_2_COMP_2	SUJ_8_SEC_2_COMP_5	SUJ_2_SEC_1_COMP_9
SUJ_5_SEC_1_COMP_2	SUJ_6_SEC_2_COMP_5	SUJ_1_SEC_2_COMP_9
SUJ_6_SEC_1_COMP_2	SUJ_5_SEC_1_COMP_5	SUJ_1_SEC_1_COMP_9
SUJ_5_SEC_2_COMP_7	SUJ_1_SEC_2_COMP_5	SUJ_3_SEC_2_COMP_9
SUJ_7_SEC_2_COMP_2	SUJ_3_SEC_2_COMP_5	SUJ_3_SEC_1_COMP_9
SUJ_7_SEC_1_COMP_1	SUJ_3_SEC_1_COMP_5	SUJ_8_SEC_2_COMP_9
SUJ_5_SEC_1_COMP_7	SUJ_5_SEC_2_COMP_5	SUJ_9_SEC_1_COMP_9
SUJ_8_SEC_2_COMP_7	SUJ_7_SEC_1_COMP_5	SUJ_9_SEC_2_COMP_9
SUJ_8_SEC_1_COMP_7	SUJ_6_SEC_1_COMP_5	SUJ_5_SEC_2_COMP_9
SUJ_9_SEC_1_COMP_7	SUJ_1_SEC_1_COMP_6	SUJ_5_SEC_1_COMP_9
SUJ_9_SEC_2_COMP_7	SUJ_3_SEC_2_COMP_6	SUJ_7_SEC_2_COMP_9
SUJ_1_SEC_2_COMP_7	SUJ_3_SEC_1_COMP_6	SUJ_6_SEC_1_COMP_9
SUJ_3_SEC_2_COMP_7	SUJ_4_SEC_2_COMP_6	SUJ_9_SEC_1_COMP_6
SUJ_3_SEC_1_COMP_7	SUJ_4_SEC_1_COMP_6	SUJ_9_SEC_2_COMP_6
SUJ_6_SEC_2_COMP_7	SUJ_8_SEC_1_COMP_7	

Tabla 6 - Secuencias de pruebas.

### III. Ejemplo de visualización de datos de las secuencias.

Los datos obtenidos en las secuencias de mapas de profundidad del entorno de trabajo, se visualizan de la forma que se puede observar en la Figura 3-9:

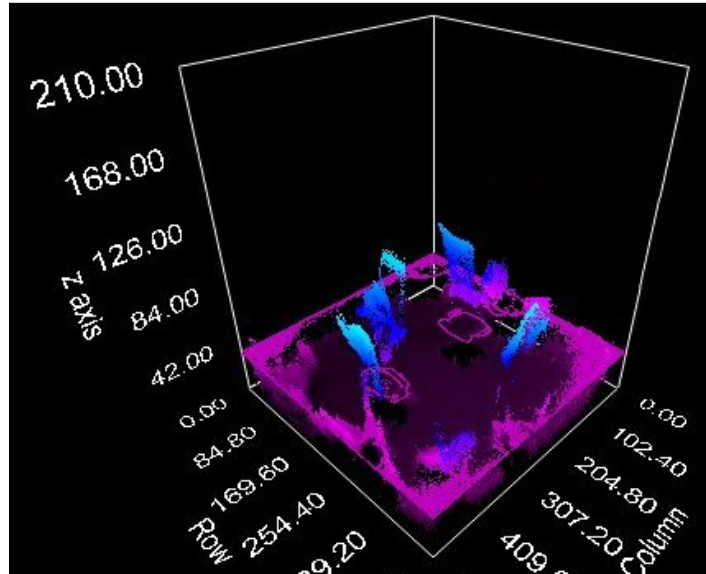


Figura 3-9 - Entorno de trabajo.

43

Se puede observar que los datos obtenidos son muy amplios, para delimitar la zona de interés, se colocan dos obstáculos, dos sillas, que harán la función de "puerta".

A continuación se muestran en detalle en forma de imagen las principales de cada tipo de complemento y característica física.

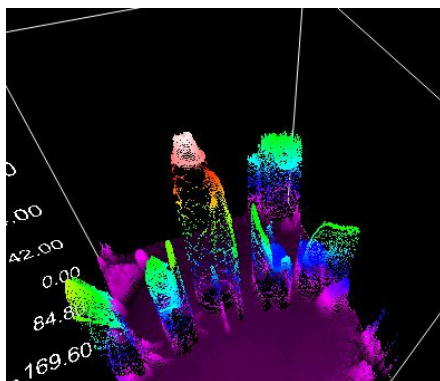


Figura 3-10 - Complemento tipo 1.

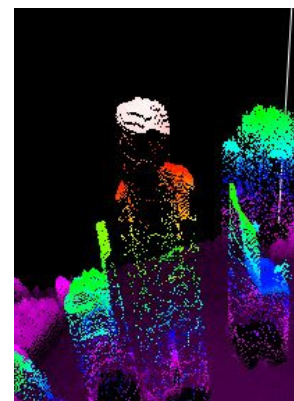


Figura 3-11 - Complemento tipo 2.



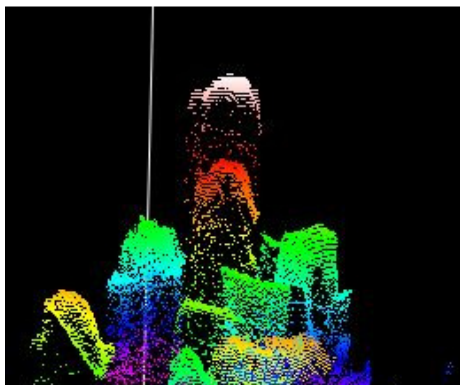


Figura 3-12 - Complemento tipo 3.

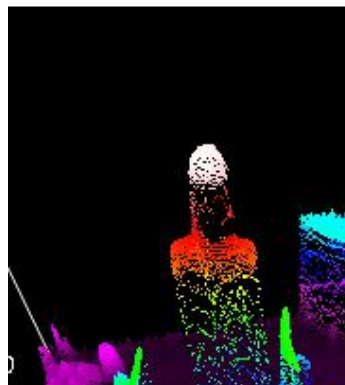


Figura 3-13 - Complemento tipo 4.

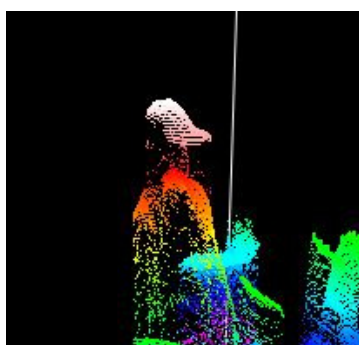


Figura 3-14 - Complemento tipo 5.

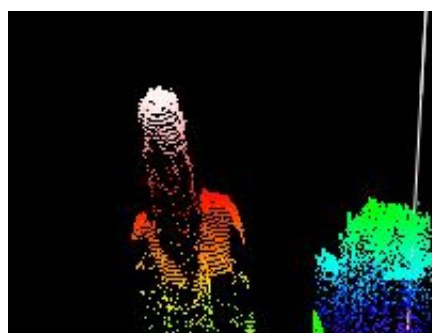


Figura 3-15 - Complemento tipo 6.

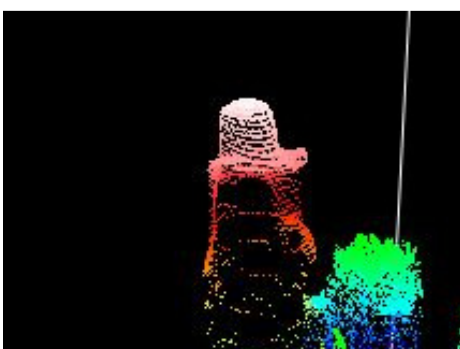


Figura 3-16 - Complemento tipo 7.

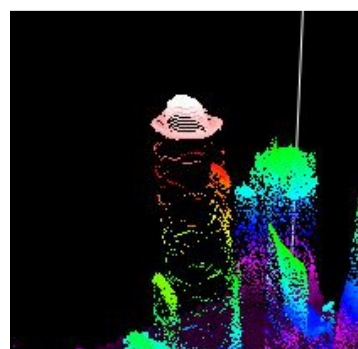


Figura 3-17 - Complemento tipo 8.

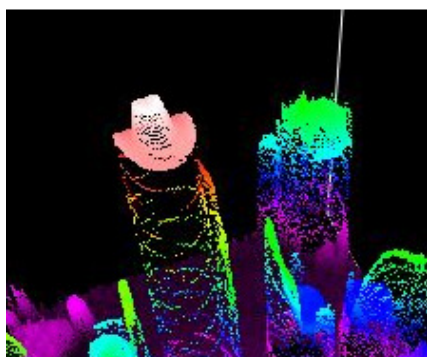


Figura 3-18 - Complemento tipo 9.



### 3.3 CLASIFICACIÓN

Se procede a intentar clasificar automáticamente las secuencias de mapas de profundidad obtenidas.

Para ello se deben establecer unos vectores modelo, a partir de las secuencias de entrenamiento con características conocidas, para comparar las futuras secuencias de mapas de profundidad correspondientes a las secuencias de pruebas.

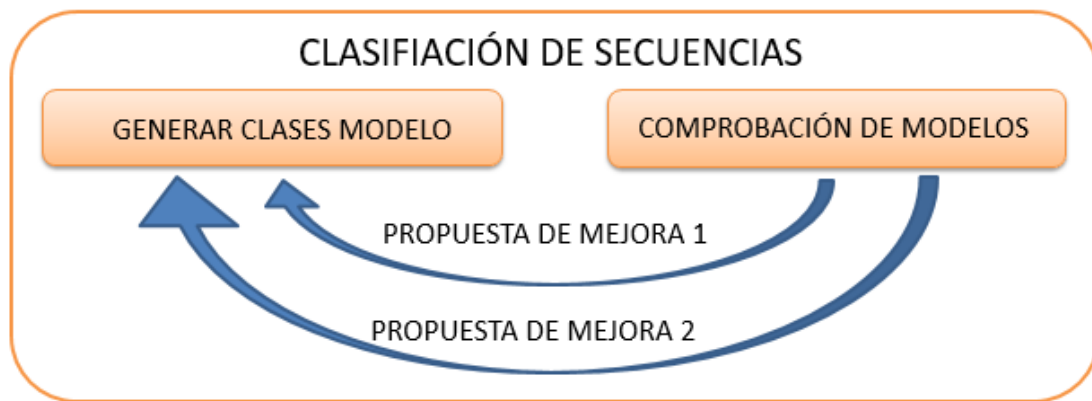


Figura 3-19 - Clasificación de secuencias

#### 3.3.1 Generación de clases modelo

Para identificar correctamente los diferentes tipos de complementos posibles se van a analizar varias secuencias de diferentes sujetos en diferentes trayectorias y con distintos complementos, utilizando para este método los archivos que ya se establecieron como **secuencias de entrenamiento** en 3.2.4 - I..

Posteriormente y una vez obtenidos los vectores de características modelo de las secuencias de entrenamiento, se van a analizar el resto de las secuencias, llamadas **secuencias de prueba** en 3.2.4 II., con los que se pretende comprobar el correcto funcionamiento de la clasificación.

### 3.3.2 Comprobación de clases modelo

Una vez agregados los vectores característicos de los archivos modelo al programa, se procede a analizar los vídeos de entrenamiento para comparar los vectores característicos de los mismos con los modelos ya almacenados.

Tras el primer análisis de resultados, se observa una gran cantidad de ruidos debidos a los obstáculos introducidos como sillas y otros objetos que se encuentran dentro del rango de visión de la cámara.

Teniendo en cuenta que el entorno para el que va a estar desarrollada esta aplicación es en el acceso de una puerta, se eliminan todos los máximos obtenidos de objetos menores de 1,6 m. De esta manera se consigue que todos los máximos a evaluar estén situados en la trayectoria teórica que seguirían los sujetos al atravesar una puerta.

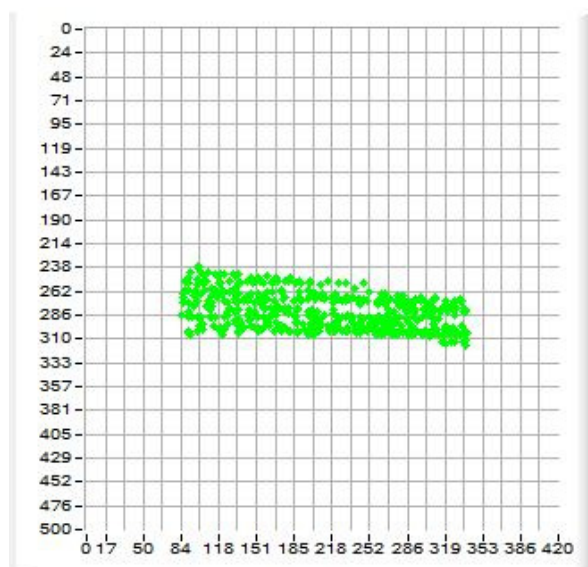


Figura 3-20 - Representación de trayectorias seguidas por los sujetos.

Se puede observar que se consigue el efecto deseado ya que todas las trayectorias están comprendidas en un ancho menor al del marco de una puerta.

Se vuelven a procesar una gran cantidad de secuencias de pruebas y a la vista de los obteniendo indicados en el - Anexo 1 – Resultados sin filtrar página 70 se puede llegar a las siguientes conclusiones:

- Se han obtenido solamente 6 aciertos con gran porcentaje de reconocimiento vs 5 errores con porcentaje de reconocimiento nulo.
- Los porcentajes de reconocimiento de los distintos modelos, por lo general, son muy bajos y están muy repartidos en todos los modelos.
- La tasa de ruido de las secuencias obtenidas es muy elevada.

Estos resultados demuestran la necesidad de buscar estrategias que permitan mejorar las tasas de acierto para todos los modelos. Por ello se decide realizar un filtrado de medidas de profundidad.

### 3.4 FILTRADO DE LAS MEDIDAS DE PROFUNDIDAD

Ante los resultados obtenidos y los problemas observados se decide realizar un filtrado de las secuencias de mapas de profundidad para reducir el ruido y conseguir una mayor precisión con el objetivo de mejorar la tasa de acierto en el reconocimiento de los complementos.

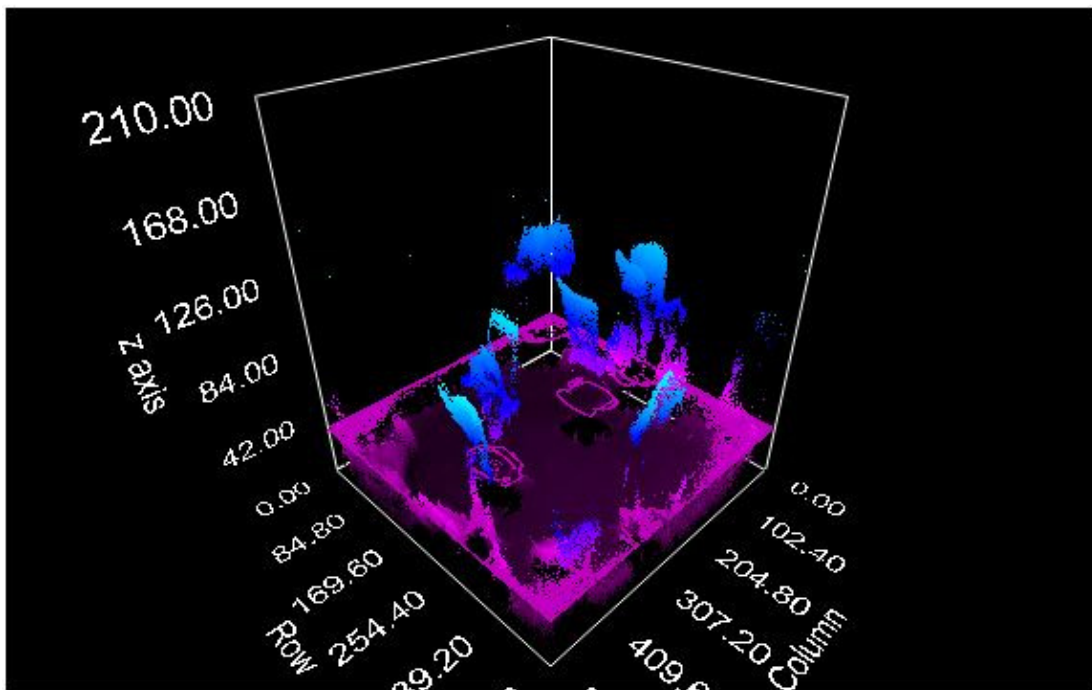


Figura 3-21 - Secuencia cenital sin filtrar.

Se puede observar la existencia de puntos flotando que introducen errores.

Para descartar todos los puntos máximos no correspondientes a sujetos, se propone la realización de un filtrado por superficies, por el cual se descartan como máximos seleccionables como posible sujeto aquellos que no tengan otro punto máximo en la zona de vecindad más próxima.

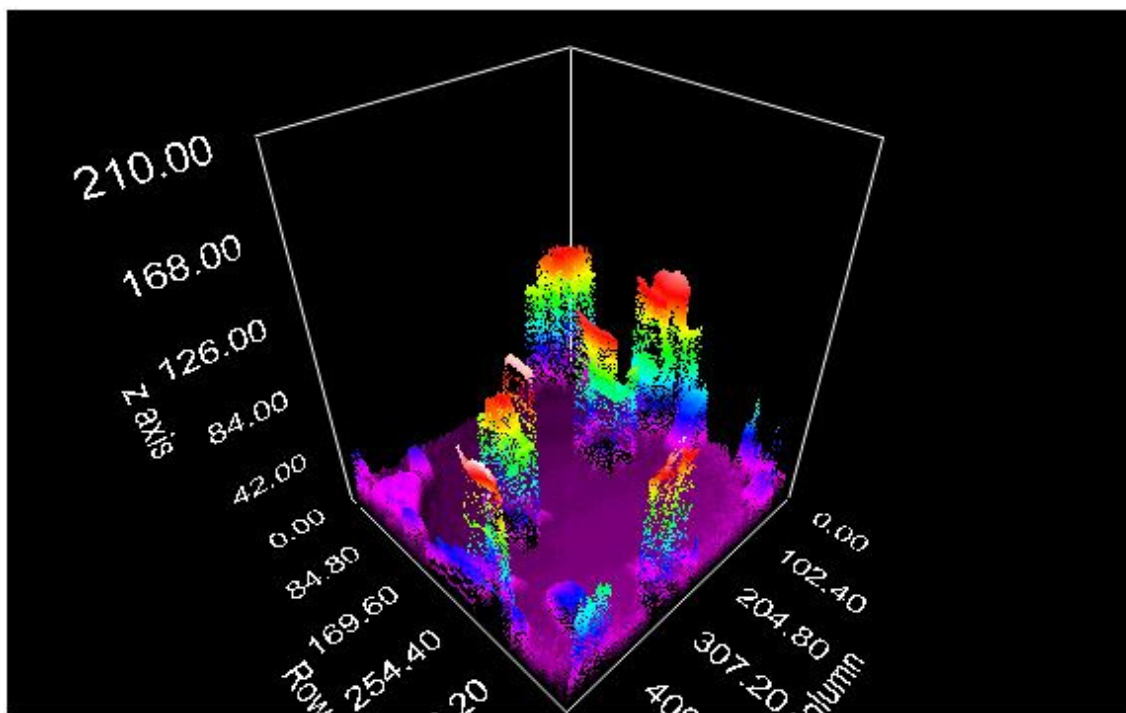
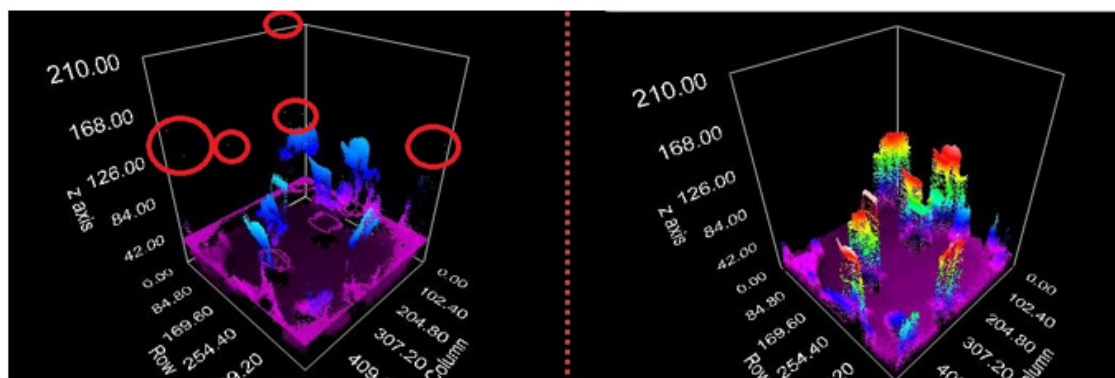


Figura 3-22 - Secuencia cenital filtrada

Como se puede observar, se consigue una mejor visualización de la escena y la eliminación de falsos positivos gracias a la reducción de ruido.



a) Mapa sin filtrar, círculos rodeando ruidos.

b) Mapa filtrado, sin ruidos

Figura 3-23 - Comparativa filtrado

Una vez realizado el mismo, se sigue el procedimiento anterior para analizar las secuencias de datos de profundidad obtenidas:

### 3.4.1.1 Generación de clases modelo

Al haberse realizado modificaciones en el programa se deben obtener de nuevo los vectores de características de las denominadas **secuencias de entrenamiento** en 3.2.4 - I.

Siguiendo el mismo procedimiento anterior se van a analizar las llamadas **secuencias de prueba** en 3.2.4 II.

De los resultados obtenidos en la primera modificación se observa que ni la cámara ni el programa tienen resolución suficiente para detectar un complemento del tipo de las “gafas de sol” por lo que se decide quitarlo de las secuencias de entrenamiento.

### 3.4.1.2 Comprobación de clases modelo

Una vez agregados los nuevos vectores característicos de los archivos modelo, se procede a analizar los vídeos de entrenamiento para comparar los vectores.

Se vuelven a procesar todos los datos obteniendo los resultados que se pueden observar en el - Anexo 2 – Resultados filtrados página 73.

49

	C0	C1	C2	C3	C4	C5	C6	C7	C8
SUJ_1_SEC_2_COMP_8	2%	0%	0%	2%	2%	7%	0%	87%	2%
SUJ_2_SEC_1_COMP_8	4%	1%	4%	0%	2%	1%	0%	88%	0%
SUJ_2_SEC_2_COMP_8	2%	0%	4%	10%	0%	0%	0%	80%	10%
SUJ_3_SEC_2_COMP_8	0%	0%	0%	5%	0%	8%	0%	87%	0%
SUJ_3_SEC_1_COMP_8	2%	3%	2%	0%	3%	2%	0%	85%	5%
SUJ_1_SEC_1_COMP_5	8%	0%	0%	7%	22%	5%	2%	0%	57%
SUJ_1_SEC_2_COMP_5	7%	0%	0%	5%	49%	9%	0%	0%	29%
SUJ_2_SEC_1_COMP_5	10%	0%	15%	68%	0%	0%	8%	0%	0%
SUJ_2_SEC_2_COMP_5	14%	0%	13%	61%	13%	0%	0%	0%	0%
SUJ_3_SEC_1_COMP_5	9%	0%	0%	9%	60%	0%	1%	0%	21%
SUJ_3_SEC_2_COMP_5	9%	0%	0%	21%	62%	0%	0%	0%	9%
SUJ_5_SEC_1_COMP_5	10%	0%	0%	13%	37%	10%	0%	0%	31%
SUJ_5_SEC_2_COMP_5	5%	0%	0%	7%	74%	0%	2%	0%	12%
SUJ_5_SEC_2_COMP_5	25%	0%	16%	25%	32%	1%	0%	0%	0%
SUJ_1_SEC_1_COMP_6	15%	0%	32%	13%	22%	1%	1%	0%	15%
SUJ_1_SEC_2_COMP_6	0%	0%	0%	0%	94%	6%	0%	0%	0%
SUJ_2_SEC_1_COMP_6	5%	2%	3%	0%	82%	0%	0%	0%	8%
SUJ_2_SEC_2_COMP_6	6%	0%	0%	2%	10%	0%	0%	82%	0%
SUJ_3_SEC_1_COMP_6	8%	0%	0%	11%	11%	70%	1%	0%	0%
SUJ_3_SEC_2_COMP_6	5%	0%	0%	6%	85%	0%	0%	0%	5%
SUJ_5_SEC_1_COMP_6	6%	0%	0%	2%	48%	0%	0%	44%	0%
SUJ_5_SEC_2_COMP_6	5%	2%	0%	3%	8%	0%	0%	81%	0%
SUJ_4_SEC_1_COMP_7	0%	0%	2%	0%	16%	10%	2%	0%	71%

SUJ_4_SEC_2_COMP_7	0%	0%	0%	4%	9%	86%	0%	0%	2%
SUJ_3_SEC_1_COMP_7	0%	0%	1%	0%	4%	33%	0%	0%	61%
SUJ_3_SEC_2_COMP_7	2%	0%	0%	2%	27%	70%	0%	0%	0%
SUJ_1_SEC_1_COMP_4	4%	0%	0%	1%	63%	29%	0%	0%	3%
SUJ_1_SEC_2_COMP_4	7%	0%	0%	2%	54%	37%	0%	0%	0%
SUJ_2_SEC_1_COMP_4	8%	5%	19%	56%	0%	0%	13%	0%	0%
SUJ_2_SEC_2_COMP_4	7%	15%	2%	63%	0%	0%	13%	0%	0%
SUJ_3_SEC_1_COMP_4	13%	0%	0%	87%	0%	0%	0%	0%	0%
SUJ_3_SEC_2_COMP_4	11%	0%	0%	83%	4%	0%	2%	0%	0%
SUJ_1_SEC_1_COMP_3	7%	0%	11%	75%	3%	3%	0%	0%	0%
SUJ_1_SEC_2_COMP_3	15%	12%	52%	15%	0%	0%	5%	0%	0%
SUJ_2_SEC_1_COMP_3	9%	0%	29%	50%	0%	0%	13%	0%	0%
SUJ_2_SEC_2_COMP_3	14%	0%	47%	35%	0%	0%	4%	0%	0%
SUJ_3_SEC_1_COMP_3	3%	0%	79%	5%	0%	0%	13%	0%	0%
SUJ_3_SEC_2_COMP_3	6%	0%	57%	4%	10%	22%	0%	0%	0%
SUJ_1_SEC_1_COMP_2	8%	4%	75%	10%	0%	0%	3%	0%	0%
SUJ_1_SEC_2_COMP_2	10%	0%	87%	3%	0%	0%	0%	0%	0%
SUJ_2_SEC_1_COMP_2	7%	3%	36%	0%	0%	0%	54%	0%	0%
SUJ_2_SEC_2_COMP_2	7%	2%	70%	0%	0%	0%	21%	0%	0%
SUJ_3_SEC_1_COMP_2	9%	0%	85%	0%	0%	3%	1%	0%	1%
SUJ_3_SEC_2_COMP_2	7%	7%	29%	0%	0%	0%	57%	0%	0%
SUJ_1_SEC_1_COMP_1	13%	64%	0%	23%	0%	0%	0%	0%	0%
SUJ_1_SEC_2_COMP_1	34%	16%	10%	37%	0%	1%	0%	0%	1%
SUJ_2_SEC_1_COMP_1	47%	4%	3%	45%	0%	0%	1%	0%	0%
SUJ_2_SEC_2_COMP_1	0%	0%	12%	4%	11%	68%	0%	0%	5%
SUJ_3_SEC_1_COMP_1	7%	0%	11%	27%	13%	38%	0%	0%	4%
SUJ_3_SEC_2_COMP_1	82%	15%	0%	3%	0%	0%	0%	0%	0%
SUJ_3_SEC_2_COMP_1	10%	0%	17%	54%	17%	2%	0%	0%	0%
SUJ_1_SEC_2_COMP_8	0%	0%	0%	63%	30%	7%	0%	0%	0%
SUJ_2_SEC_1_COMP_8	16%	5%	0%	79%	0%	0%	0%	0%	0%

Tabla 7 - Tabla de presentación de resultados modificación 2

En esta nueva presentación de resultados, se marcan en verde y en función de su peso específico por orden de intensidad los mayores resultados.

A partir de estos resultados se pueden sacar las siguientes conclusiones:

La resolución de la cámara, en posición cenital, no nos permite identificar correctamente una diferencia clara entre personas con pelo largo suelto liso o personas con pelo corto.

Esto se debe a la forma de analizar las secuencias, se está evaluando un área y esta diferencia fisonómica no produce ningún tipo de modificación entre las áreas evaluadas.

### 3.5 REDUCCIÓN DE LA ZONA DE TRABAJO

Ante los resultados obtenidos en ambas modificaciones que se pueden observar en ambos anexos 1 (Pág.70) y 2 (Pág.73), se observan patrones de errores que no han sufrido mejoras con el filtrado.

Desde la modificación 1 los archivos generados presentan muchos errores de clasificación en las primeras y últimas posiciones de los sujetos. Se propone eliminar estas zonas del reconocimiento ya que producen “sombras” en la cámara.



Figura 3-24 - Trayectoria de sujeto bajo cámara ToF.

*Fuente: Elaboración propia desde [8] y [9].*

Estos errores se achacan a la posición cenital de la cámara, produciendo sombras que eliminan características importantes de los sujetos en el inicio y final de las secuencias.



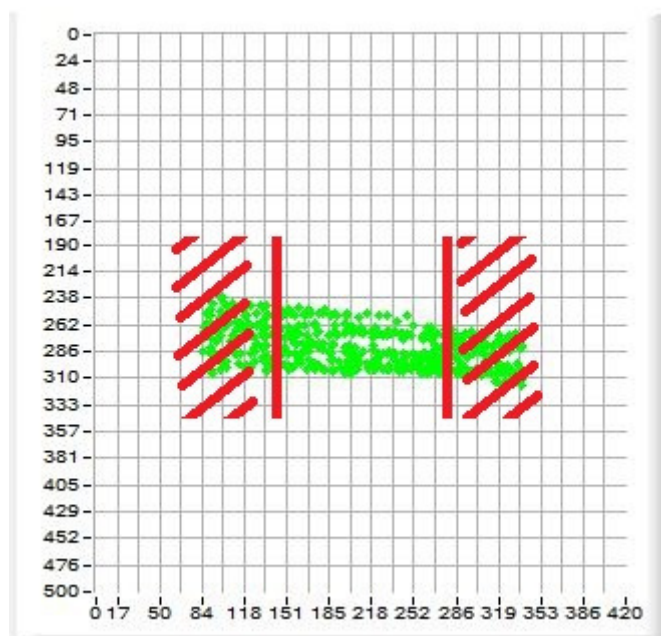


Figura 3-25 - Acotación de datos.

De esta manera, los datos analizados corresponden al sujeto completamente.



Figura 3-26 - Sujeto centrado en la secuencia.

. Fuente: Elaboración propia desde [8] y [9].

Debido a las características físicas de los complementos empleados y al tipo de análisis que se está empleando utilizando cómo clasificador las superficies obtenidas, los complementos que no presentan una gran diferencia de superficie no se pueden diferenciar.



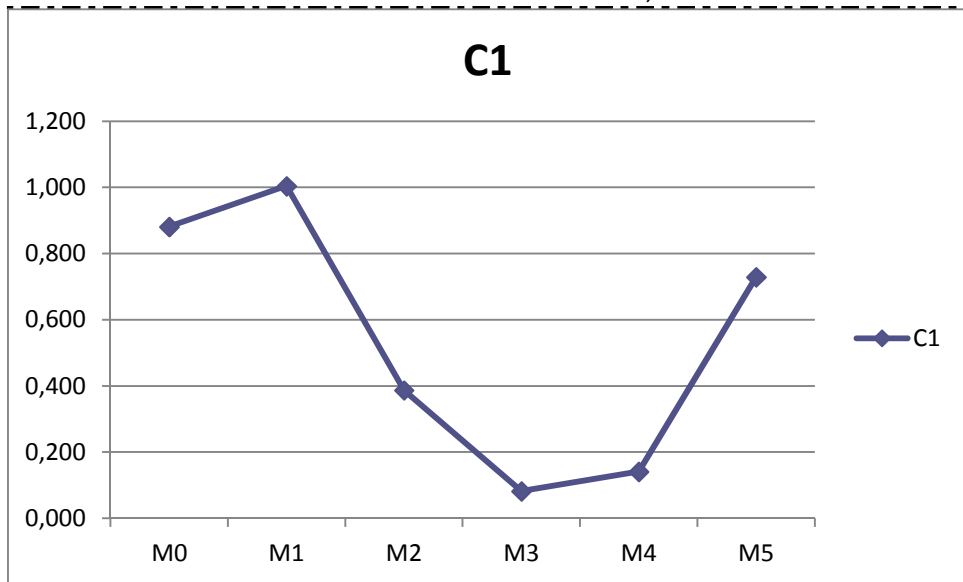
### 3.5.1 Acotación de la región de interés de análisis

Una vez realizadas las modificaciones sugeridas en el apartado anterior, se procede a generar los vectores modelo y a realizar el análisis final y completo de los vídeos de entrenamiento.

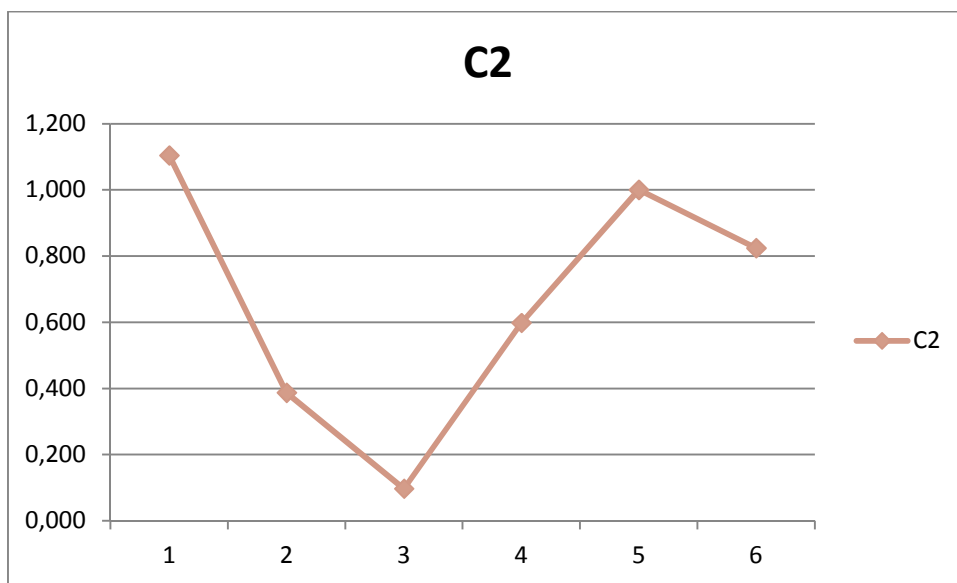
#### 3.5.1.1 Vectores de características de las clases modelo

Se procede a analizar gráficamente las secuencias de entrenamiento obteniendo los siguientes vectores de características:

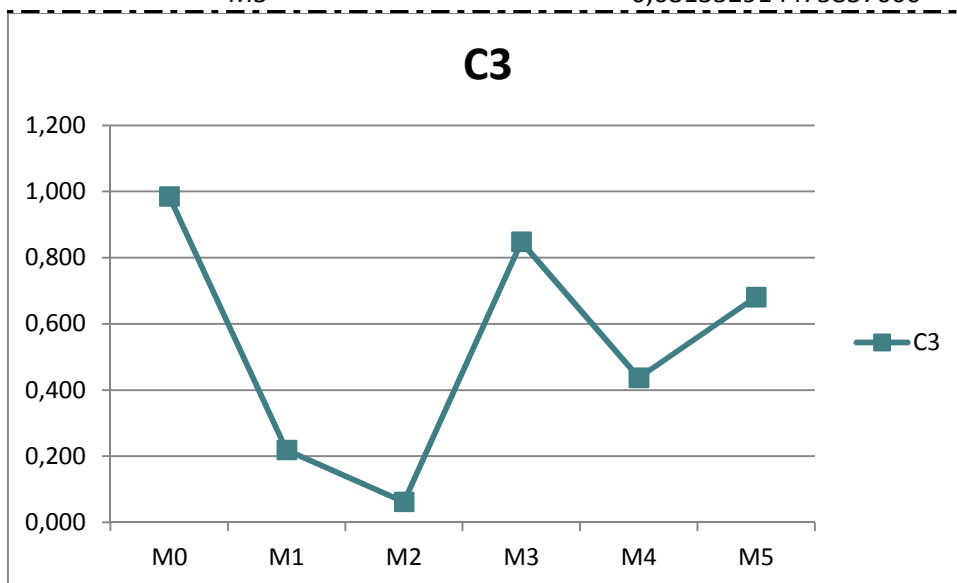
C 1	M0	0,880938536493373000
	M1	1,003627149657000000
	M2	0,386813067345758000
	M3	0,082052454836938200
	M4	0,141367312844979000
	M5	0,728712264212981000



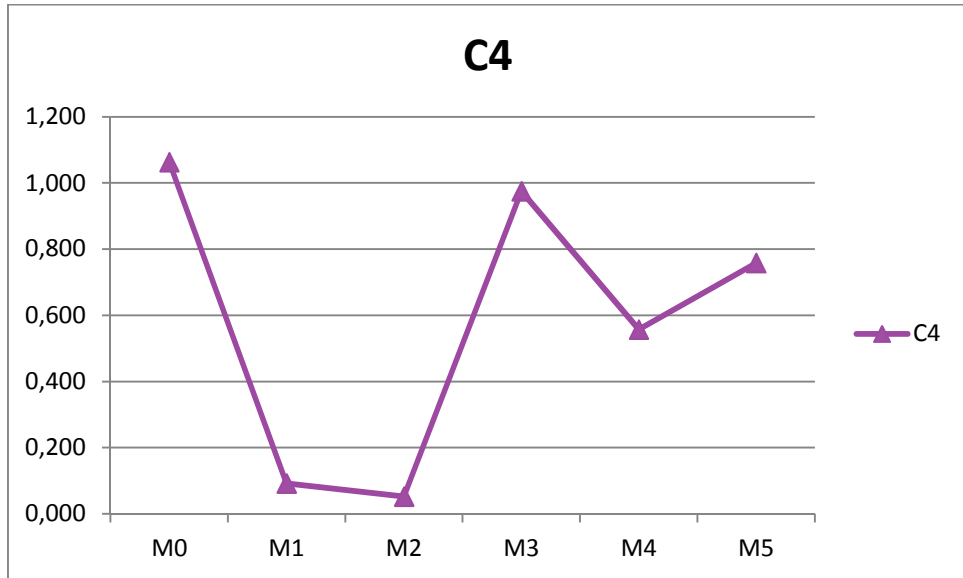
C 2	M0	1,104418524889710000
	M1	0,386611990864318000
	M2	0,096759797967962900
	M3	0,598415841871641000
	M4	1,000036428870040000
	M5	0,823923121010432000



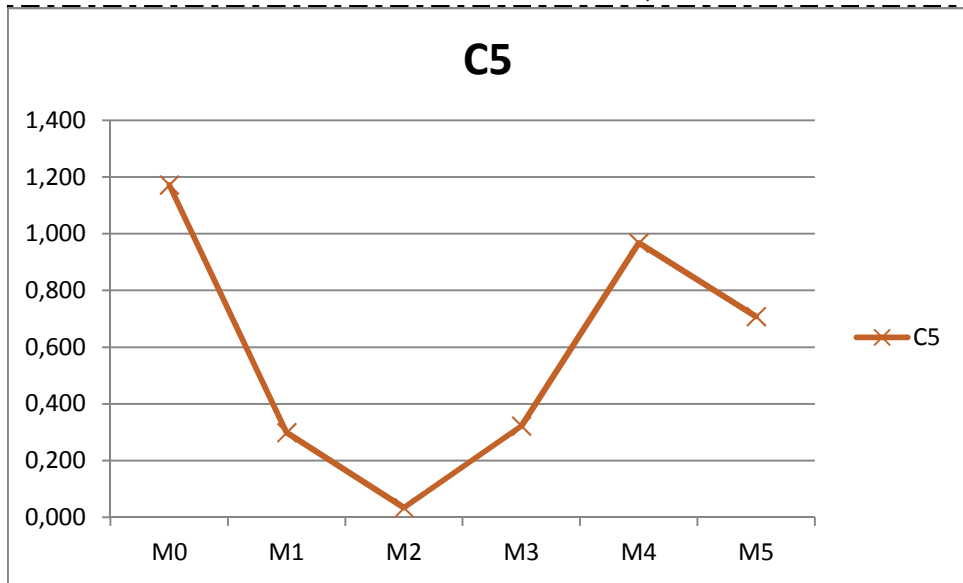
C 3		
	M0	0,984772163967778000
	M1	0,219466163782999000
	M2	0,061628612204210800
	M3	0,847915580701247000
	M4	0,436992427995736000
	M5	0,681352914475837000



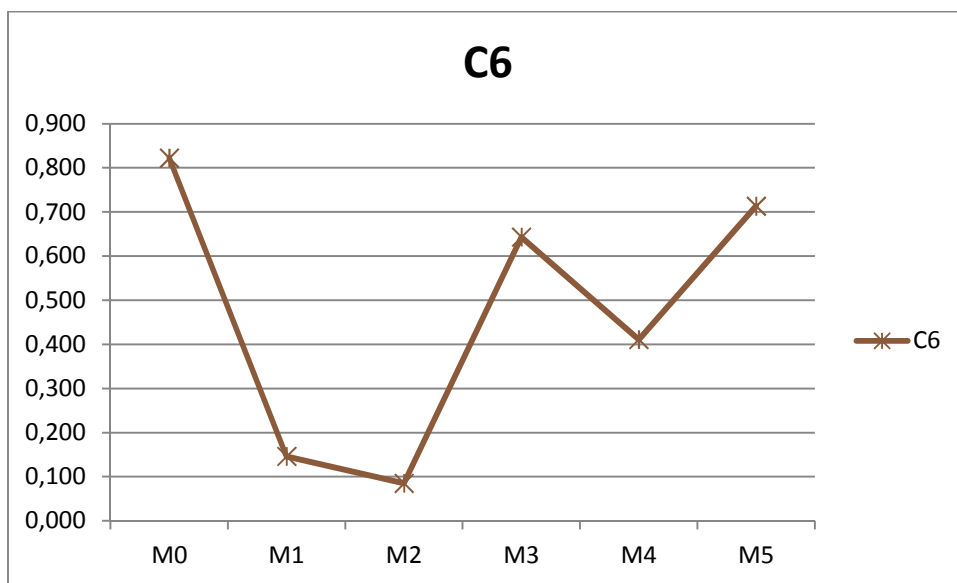
C 4		
	M0	1,063421047953750000
	M1	0,092258445531389800
	M2	0,052022075283302900
	M3	0,974911245105281000
	M4	0,556905972493071000
	M5	0,757675839504331000



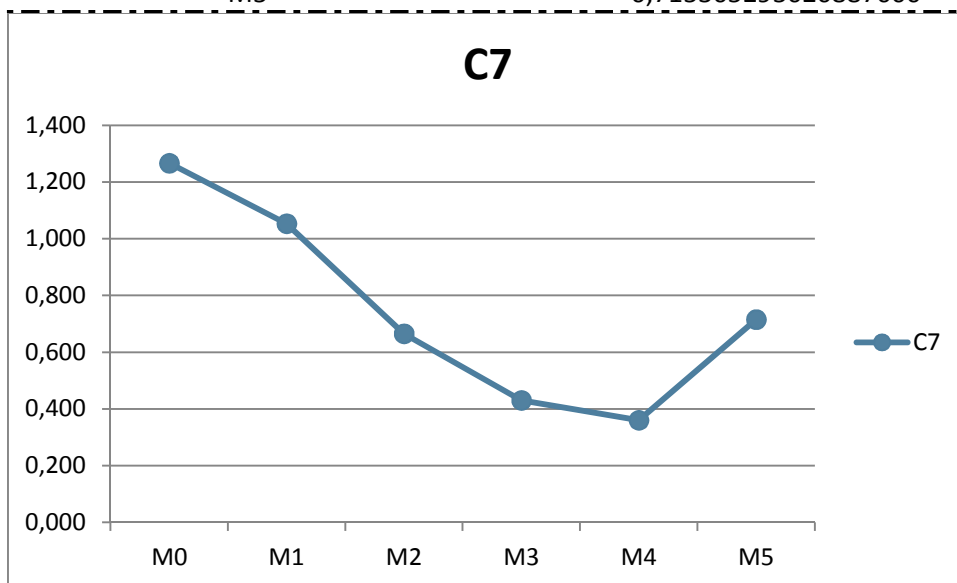
C 5	M0	1,171982079978190000
	M1	0,299325726169239000
	M2	0,034357766203585200
	M3	0,321834206267006000
	M4	0,966596537756147000
	M5	0,707583495504991000



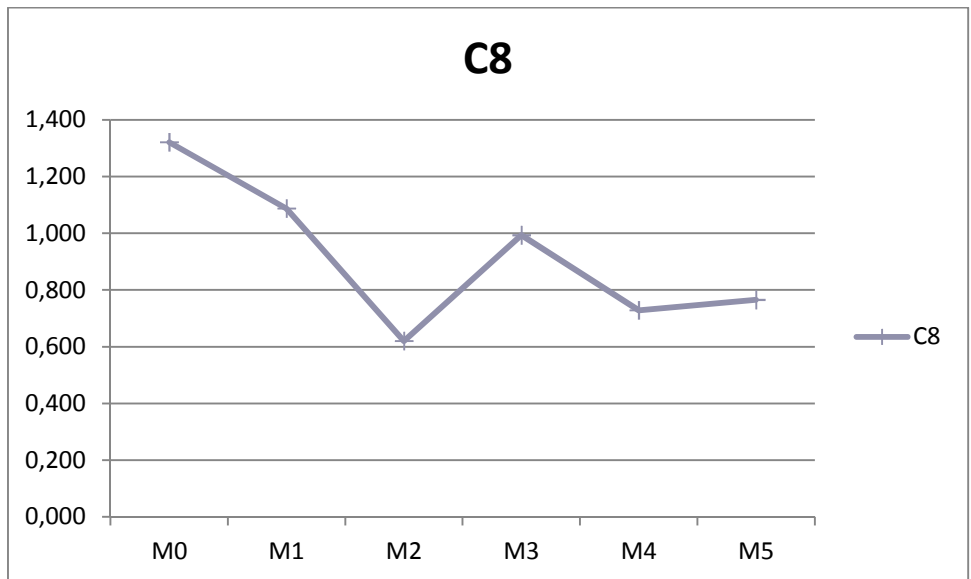
C 6	M0	0,822122548055113000
	M1	0,145837393184973000
	M2	0,085410525610395800
	M3	0,643171967672122000
	M4	0,411348137100430000
	M5	0,713061950264263000



C 7	Metric	Value
	M0	1,265519900598660000
	M1	1,053271771473020000
	M2	0,665262575573177000
	M3	0,429758303018028000
	M4	0,360444909379059000
	M5	0,715305295020887000



C 8	Metric	Value
	M0	1,320819612871030000
	M1	1,087150533310020000
	M2	0,620163547871988000
	M3	0,993283932368036000
	M4	0,727778549461306000
	M5	0,765478989238674000



C 9	M0	1,046209416728470000
	M1	2,117153228479690000
	M2	1,343225897549300000
	M3	0,049910872845429600
	M4	0,207890339265869000
	M5	0,755396804255149000

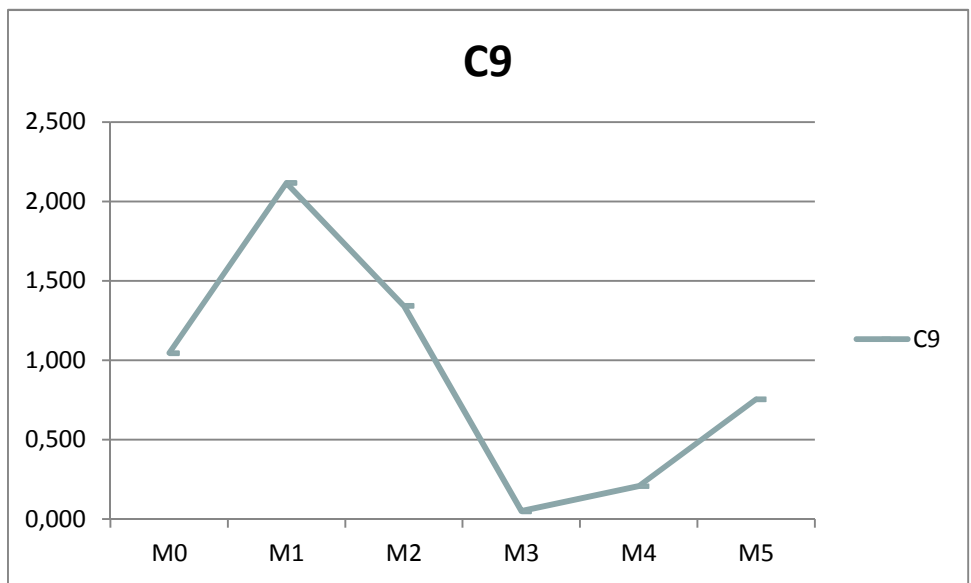


Tabla 8 - Vectores características modelos

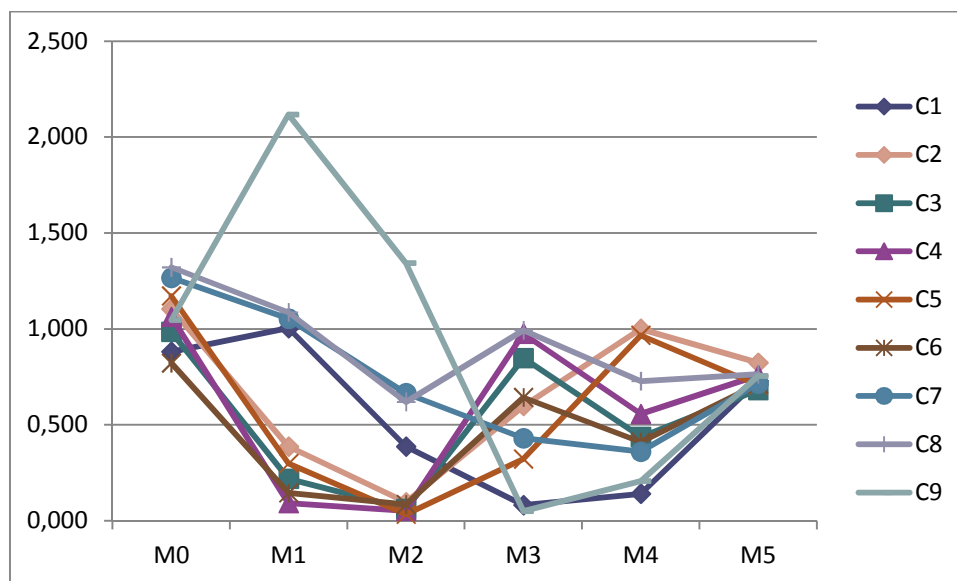


Tabla 9 - Comparativa de vectores

### 3.5.1.2 Comprobación de clases modelo

Utilizando el coeficiente de correlación de Pearson a los vectores característicos de los modelos, se puede medir el grado de relación entre 2 clases. A diferencia de utilizar la covarianza, este es independiente de las unidades de medida de las variables usadas.

### 3.5.1.3 Coeficiente de correlación de Pearson

	C1	C2	C3	C4	C5	C6	C7	C8	C9
C1	1								
C2	0,046552537	1							
C3	0,004282033	0,755115159	1						
C4	-0,140428986	0,77330841	0,986193908	1					
C5	0,168239068	0,97066267	0,647325743	0,656535646	1				
C6	0,031201925	0,801527015	0,970718471	0,970732496	0,704689973	1			
C7	0,910468975	0,111900294	0,171356883	0,029396277	0,251450484	0,138532667	1		
C8	0,535692908	0,409476678	0,607572628	0,486155503	0,419188026	0,469016898	0,734140818	1	
C9	0,795797925	-0,494886545	-0,540912979	-0,666256203	-0,33696279	-0,559740141	0,708868345	0,22164878	1

Tabla 10 - Coeficiente de correlación de Pearson

Los valores de cualquier coeficiente de correlación deben encontrarse entre -1 y +1.

El coeficiente de correlación de Pearson nos indica: “Si se demuestra que dos variables pueden variar conjuntamente los valores obtenidos tienden a -1 y +1, si no están relacionados, tienden a 0”.

Observando los complementos de la Tabla 3, vemos que la diferenciación entre los complementos 1, 7 y 9, correspondientes a sombreros con ala, es muy pequeña y puede dar pie a errores, al igual que entre 2, 3, 4 y 5, correspondientes a pelo corto, coleta, chulapo y gorra que presentan muy poca superficie.

A continuación se procede a analizar la efectividad de este método de identificación analizando los resultados.

## 4 ANÁLISIS DE RESULTADOS

Ante los resultados arrojados por el coeficiente de correlación de Pearson se puede observar claramente que existe la posibilidad de clasificar las diferentes secuencias obtenidas.

Por lo tanto se encuentra acertado el objeto de estudio del presente trabajo.

Se han obtenido 3483 muestras mediante la adquisición de secuencias de mapas de profundidad, analizando las diferentes formas posibles de clasificación, se pueden agrupar de las siguientes maneras:

### 4.1 DIFERENCIACIÓN DOBLE

En este apartado se pretende clasificar el resultado en dos grandes grupos, sujetos con complemento y sujetos sin complemento, obteniéndose los siguientes resultados sobre el total de 3483 muestras:

	Con complemento	Sin complemento
Con complemento	2287	394
Sin complemento	147	655

61

*Tabla 11 - Comparativa resultados con y sin complemento*

La tasa de acierto obtenida es de:

	Con complemento	Sin complemento
Con complemento	85%	15%
Sin complemento	18%	82%

*Tabla 12 - Tasa de acierto diferenciación doble.*

Se observa una tasa de acierto muy similar, los fallos obtenidos son debidos a la utilización de complementos que presentan una superficie mínima, sombreros con poco vuelo.



## 4.2 DIFERENCIACIÓN MÚLTIPLE

Para esta sección de resultados se van a agrupar los resultados según las características físicas de los complementos:

	Sin Complemento	Complemento Pequeño	Complemento Grande
Sin Complemento	655	83	64
Complemento Pequeño	387	600	251
Complemento Grande	7	313	1123

Tabla 13 - Comparativa resultados agrupados por tamaños.

La tasa de acierto obtenida es de:

	Sin Complemento	Complemento Pequeño	Complemento Grande
Sin Complemento	82%	10%	8%
Complemento Pequeño	48%	48%	20%
Complemento Grande	1%	22%	78%

Tabla 14 - Tasa de acierto diferenciación múltiple.

Se puede apreciar la buena diferenciación existente entre la existencia de un complemento de gran tamaño y los casos sin complemento. El problema aparece con los complementos de pequeño tamaño ya que estos se pueden confundir.

## 4.3 DIFERENCIACIÓN INDIVIDUAL

En vista de los resultados anteriores, se puede predecir que en la diferenciación por complementos, se van a obtener algunos resultados malos para algunos complementos pequeños que provocan errores al sistema.

	C1	C2	Sin Complemento	C5	C7	C8	C9
C1	175	35	28	1	6	238	0
C2	0	234	60	54	7	0	0
Sin Complemento	5	69	655	9	51	13	0
C5	0	67	299	34	0	0	0
C7	128	0	0	0	277	119	0
C8	21	22	7	14	49	179	68
C9	127	0	0	1	23	0	408

Tabla 15 - Comparativa resultados diferenciación individual.

	C1	C2	Sin Complemento	C5	C7	C8	C9
C1	36%	7%	6%	0%	1%	49%	0%
C2	0%	66%	17%	15%	2%	0%	0%
Sin Complemento	1%	9%	82%	1%	6%	2%	0%
C5	0%	17%	75%	9%	0%	0%	0%
C7	24%	0%	0%	0%	53%	23%	0%
C8	6%	6%	2%	4%	14%	50%	19%
C9	23%	0%	0%	0%	4%	0%	73%

Tabla 16 - Tasa de acierto diferenciación individual

Cómo puede observarse la diagonal principal debería encontrarse en verde completamente, pero se observan fallos con los complemento C1 y C5 debidos a sus características físicas.

	C1	C2	Sin Complemento	C7	C8	C9
C1 + C5	210	102	327	6	238	0
C2	54	234	60	7	0	0
Sin Complemento	14	69	655	51	13	0
C7	128	0	0	277	119	0
C8	35	22	7	49	179	68
C9	128	0	0	23	0	408

Tabla 17 - Comparativa resultados agrupados.

Por lo que se agrupan en una misma categoría:

	C1 + C5	C2	Sin Complemento	C7	C8	C9
C1 + C5	43%	12%	37%	1%	27%	0%
C2	15%	66%	17%	2%	0%	0%
Sin Complemento	2%	9%	82%	6%	2%	0%
C7	24%	0%	0%	53%	23%	0%
C8	10%	6%	2%	14%	50%	19%
C9	23%	0%	0%	4%	0%	73%

Tabla 18 - Tasa de acierto agrupada.

Con este tipo de clasificación y tomando este número de muestras disponibles se consiguen unos resultados satisfactorios.

## 5 CONCLUSIONES

---

A continuación, se indican las conclusiones y logros obtenidos en el presente trabajo.

Este trabajo se ha centrado en el análisis de diferentes personas con diferentes atributos físicos y complementos mediante el uso del sensor ToF de una cámara Kinect 2.

En primer lugar se ha desarrollado un software para capturar y almacenar las secuencias de datos, posteriormente se han analizado y se han podido clasificar satisfactoriamente.

Dependiendo de la aplicación para la que se vaya a usar, el sistema puede ofrecer clasificación básica con una alta tasa de acierto, entre el 82% y el 85% para la clasificación entre sujetos con complementos y sin ellos, para una clasificación que discrimina entre sin complementos, complementos de pequeño tamaño (gorras, boinas...) y complementos de gran tamaño se observa que los porcentajes son similares a los anteriores con los complementos de mayor tamaño respecto a la ausencia de complementos (82% sin complemento, 48% complemento de pequeño tamaño y 75% complementos de gran tamaño), hasta una clasificación particularizada para distintos tipos de complementos agrupando los que presentan similitudes físicas, con una tasa de acierto menor indicada que mantiene un 82% para sin complemento, tasas elevadas 73% y 66% entre complementos de gran tamaño y unos valores aceptables para los de pequeño tamaño 53%, 50% y 43%.

El principal logro obtenido es conseguir clasificar con diferentes niveles de fiabilidad, distintos tipos de personas mediante el uso de una sola cámara de tiempo de vuelo, utilizando únicamente la información obtenida de las superficies más elevadas de su cuerpo respetando completamente la privacidad de los individuos.

## 6 LÍNEAS DE TRABAJO FUTURAS

---

Teniendo en cuenta las características del trabajo expuesto anteriormente las líneas futuras de investigación y mejora podrían ser las siguientes:

Del presente trabajo:

- Aumento de la biblioteca de secuencias de las secuencias de entrenamiento para mejorar la tasa de acierto de las secuencias de entrenamiento.
- Aumentar el número de secuencias de entrenamiento para poder evaluar con mayor fiabilidad los resultados.
- Mejora en la posible detección múltiple de personas y objetos.
- Aumento de la resolución mediante el uso de mejores cámaras.

Para futuros trabajos:

- Mejora de la tasa de acierto utilizando varias cámaras en diferentes posiciones.
- Desarrollo de un posible software adaptable a distintas alturas, auto calibración de altura mediante la medida enviada por la cámara desde la posición cenital.
- Estudio del comportamiento del software desarrollado en diferentes entornos o con diferentes secuencias obtenidas en diferentes entornos no controlados, a distintas alturas, etc...
- Actualización del software para la detección en tiempo real y no el uso de secuencias grabadas.
- Mejora del algoritmo de clasificación para mejorar la precisión en complementos de pequeño tamaño.

## 7 POSIBLES APLICACIONES

---

Principalmente el presente trabajo se orienta a la clasificación de personas pero realizando las modificaciones oportunas se podría aplicar a la detección y clasificación de objetos.

Aplicaciones para clasificación de personas:

- Seguridad.
- Marketing.
- Fronteras.

Aplicaciones para objetos:

- Identificación de vehículos.
- Tasación de daños.

## 8 PLIEGO DE CONDICIONES

---

En el siguiente listado se mencionan los requisitos necesarios para la correcta utilización del sistema desarrollado.

Se debe disponer como mínimo del siguiente hardware y software:

I. Sistemas operativos utilizado:

- Windows 10.

II. Hardware necesario:

- Cámara Microsoft Kinect 2.
- Procesador de 64-bit (x64).
- 6 Gb de memoria RAM.
- 500 Gb de memoria ROM.
- Procesador de doble núcleo físico de 2.7 GHz (cuatro núcleos lógicos por cada físico).
- Gráfica NVIDIA GeForce GT 720 M (2 Gb VRAM)
- Controlador USB 3.0 dedicado.
- Adaptador gráfico que soporte DX11.

III. Software necesario:

- LabWindows CVI.
- Visual Studio 2013.
- .NET Framework 4.0.
- Microsoft Kinect for Windows SDK v2.0.

## 9 PRESUPUESTO

A continuación, se detallan, aproximadamente, los costes estimados para implementar el sistema descrito en el presente trabajo:

- Equipamiento Hardware:

	<b>Unidades</b>	<b>Precio unitario</b>	<b>Subtotal</b>
PC	1	700,00 €	700,00 €
Kinect 2	1	200,00 €	200,00 €
		<b>Total</b>	<b>900,00 €</b>

Tabla 19 - Precio Hardware

- Recursos Software:

	<b>Unidades</b>	<b>Precio unitario</b>	<b>Subtotal</b>
Windows 10	1	135,00 €	135,00 €
Kinect for Windows SDK	1	0,00 €	0,00 €
Microsoft Visual Studio	1	574,00 €	574,00 €
LabWindows CVI	1	1.597,00 €	1.597,00 €
		<b>Total</b>	<b>2.306,00 €</b>

68

Tabla 20 - Precio Software

- Costes de mano de obra:

	<b>Unidades</b>	<b>Precio unitario</b>	<b>Subtotal</b>
Mano de obra	250	70,00 €	17.500,00 €
Redacción	50	25,00 €	1.250,00 €
		<b>Total</b>	<b>18.750,00 €</b>

Tabla 21 - Precio Mano de Obra

- Presupuesto Total:

	<b>Subtotal</b>
Equipamiento Hardware	900,00 €
Recursos Software	2.306,00 €
Costes de mano de obra	18.750,00 €
<b>Total</b>	<b>21.956,00 €</b>

Tabla 22 - Precio Total

## 10 BIBLIOGRAFÍA

---

- [1] A. Bevilacqua, L. Di Stefano y P. Azzari, «People tracking using a Time-of-Flight depth sensor,» de *Multimedia Communications, Services and Security. MCSS 2013*, Sydney, Australia, 2013.
- [2] D. Ramanan, D. A. Forsyth y A. Zisserman, «Tracking People by Learning Their Appearance,» *IEEE Transactions on pattern analysis and machine intelligence*, vol. 29, pp. 65-81, January 2007.
- [3] G. Rodríguez-Canosa, J. Del Cerro y A. Barrientos, «Detection and tracking of dynamic objects by using a multirobot system: Application to critical Infrastructures Surveillance,» *Sensors*, nº 14, pp. 2911-2943, 2013.
- [4] C. Luna Vazquez, C. Losada Gutierrez, D. Fuentes, A. Fernandez, M. Mazo y J. Macias Guarasa, «Robust people counting using a ceil mounted ToF camera,» *Expert Systems with Applications*, vol. 71, nº C, pp. 240-256, Abril 2017.
- [5] Microsoft, «Microsoft Download Center,» Microsoft, 15 Junio 2016. [En línea]. Available: <https://www.microsoft.com/en-us/download/default.aspx>. [Último acceso: Microsoft Junio 2016].
- [6] Visual Studio Community, «Visual Studio,» 22 Junio 2016. [En línea]. Available: <https://www.visualstudio.com/post-download-vs?sku=community&clid=0x409&downloadrename=true>. [Último acceso: 22 Junio 2016].
- [7] L. Bourdev, S. Maji y J. Malik, «Describing People: A Poselet-Based Approach to Attribute Classification,» de *IEEE International Conference on Computer Vision, ICCV 2011*, Barcelona, Spain, November 6-13, 2011.
- [8] P. Exilii, «Passus Exilii,» [En línea]. Available: <http://www.passusexilii.com/>. [Último acceso: 2 Marzo 2017].
- [9] 33723407, «Pixattitude,» [En línea]. Available: <http://www.dreamstime.com>. [Último acceso: 2 Marzo 2017].



## 11- ANEXO 1 – RESULTADOS SIN FILTRAR

A continuación se recoge un ejemplo de los datos obtenidos después de la primera ejecución completa del programa.

Estos datos son presentados directamente de la salida de los archivos generados por el programa, a continuación se detalla el formato de salida de un archivo de entrenamiento una vez analizado y comparado con los vectores de los modelos ya obtenidos y cargados en el programa:

Identificando las letras **N** con el número de frame, **T** con el tipo de modelo contra el que se compara y es reconocido, **X** la posición en el eje X, **Y** la posición en el eje Y, **Z** la posición en el eje Z del máximo (se puede interpretar cómo la altura del sujeto) y **E** que indica el error obtenido entre el vector medido y el vector comparado en ese punto.

A partir de ahí los valores obtenidos son del tratamiento de datos posterior, **C-X** significa un positivo en un tipo determinado de complemento, cada columna se refiere a un tipo diferente, **Total** indica un positivo en reconocimiento de persona, indiferente del complemento.

N	T	X	Y	Z	E	C-X	C-X	C-X	C-X	Total
86	6	70	270	174	0.188	1				1
87	9	70	270	174	0.107		1			1
88	9	90	270	175	0.053		1			1
89	9	90	270	175	0.107		1			1
90	9	90	270	174	0.066		1			1
91	6	90	270	175	0.145	1				1
92	3	110	270	175	0.048			1		1
93	3	110	270	174	0.076			1		1
94	3	130	270	174	0.084			1		1
95	3	130	270	173	0.094			1		1
96	3	130	270	172	0.079			1		1
97	3	130	270	172	0.070			1		1
98	3	130	270	172	0.040			1		1
99	3	150	270	170	0.074			1		1
100	3	150	290	170	0.202			1		1
101	3	150	290	172	0.109			1		1
102	3	170	290	173	0.097			1		1
103	3	170	290	174	0.066			1		1
104	3	190	290	174	0.118			1		1
105	9	190	290	174	0.113		1			1
106	9	190	290	174	0.069		1			1
107	3	210	290	175	0.105			1		1
108	3	210	290	175	0.139			1		1
109	3	210	290	175	0.129			1		1
110	9	210	290	174	0.164		1			1
111	2	230	290	175	0.096				1	1

112	3 230 290 175 0.107					1			1
113	3 230 290 173 0.102					1			1
114	2 250 290 173 0.100						1		1
115	2 250 290 173 0.092						1		1
116	2 250 290 173 0.131						1		1
117	3 250 290 173 0.094					1			1
118	6 270 290 173 0.174	1							1
119	3 270 290 174 0.149					1			1
120	9 290 310 174 0.145		1						1
121	6 290 290 174 0.187	1							1
122	3 290 290 174 0.144					1			1
123	2 310 310 175 0.049						1		1
124	6 310 310 175 0.121	1							1
125	2 310 310 175 0.175						1		1
126	3 310 310 176 0.126					1			1
127	2 330 310 176 0.045						1		1
128	3 330 310 176 0.024					1			1
129	3 330 310 176 0.081					1			1
130	6 330 310 176 0.159	1							1
131	3 330 310 176 0.128					1			1
132	6 350 310 176 0.126	1							1
133	3 350 310 176 0.042					1			1
134	3 350 310 175 0.130					1			1
135	3 350 310 175 0.160					1			1
136	2 370 310 175 0.145						1		1
137	3 370 330 176 0.069					1			1
138	3 370 330 176 0.114					1			1
139	3 370 330 175 0.135					1			1
140	3 370 330 175 0.107					1			1
		7	8	32		8		55	
		13%	15%	58%		15%		1	
		<b>6</b>	<b>9</b>	<b>3</b>		<b>2</b>			

Tabla 23 - Ejemplo de resultados

Una vez analizadas un número importante de secuencias, se procede a comparar los resultados obtenidos con los esperados y se obtiene la siguiente tabla de valores (Teniendo en cuenta que cada columna representa un tipo de clase de complemento y cada fila un video de pruebas):

C0	C1	C2	C3	C4	C5	C6	C7	C8	C9
0%	0%	0%	58%	0%	0%	13%	0%	15%	15%
0%	0%	4%	43%	0%	0%	17%	0%	0%	35%
0%	0%	18%	12%	0%	20%	12%	0%	39%	0%
0%	0%	6%	0%	0%	6%	2%	0%	86%	0%
0%	0%	38%	2%	0%	0%	0%	0%	0%	60%
0%	0%	2%	76%	0%	0%	11%	0%	0%	11%
0%	0%	0%	0%	0%	0%	0%	0%	100%	0%
0%	0%	28%	45%	0%	0%	2%	0%	0%	26%

0%	0%	40%	0%	0%	0%	0%	0%	0%	60%
0%	0%	25%	0%	0%	0%	0%	0%	0%	75%
0%	0%	7%	2%	0%	31%	0%	0%	34%	27%
0%	0%	0%	2%	0%	0%	89%	0%	9%	0%
0%	0%	14%	8%	0%	0%	78%	0%	0%	0%
0%	0%	17%	2%	0%	0%	8%	0%	73%	0%
0%	0%	10%	33%	0%	0%	13%	0%	0%	45%
0%	0%	8%	0%	0%	0%	0%	0%	92%	0%
0%	0%	4%	0%	0%	7%	0%	0%	89%	0%
0%	0%	16%	0%	0%	44%	0%	0%	41%	0%
0%	49%	0%	9%	0%	0%	14%	0%	0%	28%
0%	0%	55%	4%	0%	0%	6%	0%	37%	0%
0%	0%	4%	2%	0%	92%	0%	0%	2%	0%
0%	0%	0%	0%	0%	71%	0%	0%	29%	0%
0%	0%	8%	0%	0%	2%	8%	0%	83%	2%
0%	0%	48%	34%	0%	0%	18%	0%	0%	0%
0%	0%	53%	2%	0%	0%	47%	0%	0%	0%
0%	0%	0%	5%	0%	0%	83%	0%	17%	0%
0%	0%	31%	22%	0%	0%	6%	0%	41%	3%
0%	0%	65%	21%	0%	0%	6%	0%	3%	6%
0%	0%	0%	0%	0%	74%	0%	0%	26%	0%
0%	0%	11%	3%	0%	0%	86%	0%	0%	0%
0%	0%	94%	0%	0%	0%	6%	0%	0%	0%
0%	0%	0%	0%	0%	0%	0%	0%	0%	100%
0%	0%	0%	0%	0%	0%	0%	0%	0%	100%
0%	0%	73%	16%	0%	0%	11%	0%	0%	0%
0%	0%	83%	6%	0%	3%	11%	0%	3%	3%
0%	0%	0%	0%	0%	100%	0%	0%	0%	0%
0%	0%	63%	24%	0%	0%	10%	0%	0%	0%
0%	0%	40%	49%	0%	0%	11%	0%	0%	0%
0%	0%	2%	0%	0%	0%	0%	0%	98%	0%
0%	0%	2%	0%	0%	0%	0%	0%	68%	26%
0%	0%	6%	0%	0%	86%	0%	0%	3%	6%
0%	0%	0%	0%	0%	35%	2%	0%	63%	0%
0%	0%	20%	0%	0%	0%	0%	0%	65%	14%
0%	0%	93%	0%	0%	0%	7%	0%	0%	0%
0%	0%	57%	2%	0%	13%	0%	0%	11%	17%
0%	0%	24%	2%	0%	2%	10%	0%	61%	0%
0%	0%	0%	0%	0%	0%	0%	0%	69%	29%
0%	0%	5%	30%	0%	20%	15%	0%	30%	0%
0%	0%	0%	0%	0%	65%	0%	0%	35%	0%

Tabla 24 - Tabla de presentación de resultados modificación 1

Se pueden observar en rojo los errores de clasificación más significativos y en verde todos los verdaderos positivos obtenidos.

## 12- ANEXO 2 – RESULTADOS FILTRADOS

Después de haber realizado la modificación indicada en 3.5.1 se vuelve a realizar el procesado de las secuencias de entrenamiento.

El análisis de los datos se realiza de una forma más automatizada, evolucionando el sistema mediante las siguientes identificaciones:

Identificando las letras **N** con el número de frame, **T** con el tipo de modelo contra el que se compara y es reconocido, **X** la posición en el eje X, **Y** la posición en el eje Y, **Z** la posición en el eje Z del máximo (se puede interpretar cómo la altura del sujeto) y **E** que indica el error obtenido entre el vector medido y el vector comparado en ese punto.

Hasta aquí los indicadores no han sufrido modificación, pero ahora se incluyen en la comparativa todos los complementos indicados en los archivos modelo, no solamente los complementos en los que se ha dado positivo. Se indican con los números correspondientes de **0** a **8** y finalmente el **Total**.

N	T	X	Y	Z	E	0	1	2	3	4	5	6	7	8	Total
72	0	266	17	167	0.573	1	0	0	0	0	0	0	0	0	1
73	0	268	17	171	0.522	1	0	0	0	0	0	0	0	0	1
74	0	264	17	174	0.507	1	0	0	0	0	0	0	0	0	1
75	0	262	20	176	0.474	1	0	0	0	0	0	0	0	0	1
76	0	274	18	178	0.336	1	0	0	0	0	0	0	0	0	1
77	0	259	23	176	0.258	1	0	0	0	0	0	0	0	0	1
78	0	265	20	176	0.173	1	0	0	0	0	0	0	0	0	1
79	0	262	28	175	0.239	1	0	0	0	0	0	0	0	0	1
80	0	266	32	175	0.163	1	0	0	0	0	0	0	0	0	1
81	0	264	39	175	0.146	1	0	0	0	0	0	0	0	0	1
82	0	265	41	176	0.106	1	0	0	0	0	0	0	0	0	1
83	0	260	56	176	0.074	1	0	0	0	0	0	0	0	0	1
84	0	258	57	177	0.104	1	0	0	0	0	0	0	0	0	1
85	0	269	57	177	0.069	1	0	0	0	0	0	0	0	0	1
86	0	264	70	176	0.056	1	0	0	0	0	0	0	0	0	1
87	0	265	67	178	0.147	1	0	0	0	0	0	0	0	0	1
88	0	265	86	177	0.036	1	0	0	0	0	0	0	0	0	1
89	0	261	88	177	0.055	1	0	0	0	0	0	0	0	0	1
90	0	268	86	177	0.064	1	0	0	0	0	0	0	0	0	1
91	0	262	97	178	0.044	1	0	0	0	0	0	0	0	0	1
92	0	271	98	177	0.069	1	0	0	0	0	0	0	0	0	1
93	0	271	108	176	0.047	1	0	0	0	0	0	0	0	0	1
94	0	264	121	177	0.023	1	0	0	0	0	0	0	0	0	1
95	0	274	134	176	0.021	1	0	0	0	0	0	0	0	0	1
96	0	272	140	176	0.015	1	0	0	0	0	0	0	0	0	1
97	0	270	132	175	0.043	1	0	0	0	0	0	0	0	0	1

98	0 269 153 175 0.026	1	0	0	0	0	0	0	0	0	1
99	0 278 157 176 0.110	1	0	0	0	0	0	0	0	0	1
100	0 283 147 176 0.064	1	0	0	0	0	0	0	0	0	1
101	0 278 164 176 0.034	1	0	0	0	0	0	0	0	0	1
102	0 281 165 176 0.110	1	0	0	0	0	0	0	0	0	1
103	0 278 176 177 0.041	1	0	0	0	0	0	0	0	0	1
104	0 283 177 176 0.022	1	0	0	0	0	0	0	0	0	1
105	0 278 188 177 0.018	1	0	0	0	0	0	0	0	0	1
106	0 281 200 178 0.128	1	0	0	0	0	0	0	0	0	1
107	0 282 197 177 0.029	1	0	0	0	0	0	0	0	0	1
108	0 285 212 178 0.129	1	0	0	0	0	0	0	0	0	1
109	0 289 203 177 0.053	1	0	0	0	0	0	0	0	0	1
110	0 284 220 178 0.070	1	0	0	0	0	0	0	0	0	1
111	0 280 219 177 0.043	1	0	0	0	0	0	0	0	0	1
112	0 286 232 177 0.030	1	0	0	0	0	0	0	0	0	1
113	0 288 237 176 0.046	1	0	0	0	0	0	0	0	0	1
114	0 287 243 176 0.051	1	0	0	0	0	0	0	0	0	1
115	0 290 235 175 0.050	1	0	0	0	0	0	0	0	0	1
116	0 290 254 176 0.061	1	0	0	0	0	0	0	0	0	1
117	1 287 261 177 0.055	0	1	0	0	0	0	0	0	0	1
118	0 293 253 176 0.031	1	0	0	0	0	0	0	0	0	1
119	0 293 271 177 0.044	1	0	0	0	0	0	0	0	0	1
120	0 294 266 177 0.008	1	0	0	0	0	0	0	0	0	1
121	1 293 277 177 0.053	0	1	0	0	0	0	0	0	0	1
122	1 300 271 177 0.023	0	1	0	0	0	0	0	0	0	1
123	0 292 287 178 0.039	1	0	0	0	0	0	0	0	0	1
124	1 297 290 178 0.075	0	1	0	0	0	0	0	0	0	1
125	0 294 308 179 0.106	1	0	0	0	0	0	0	0	0	1
126	0 299 310 179 0.068	1	0	0	0	0	0	0	0	0	1
127	3 300 318 179 0.089	0	0	0	1	0	0	0	0	0	1
128	1 301 322 179 0.060	0	1	0	0	0	0	0	0	0	1
129	1 301 330 179 0.063	0	1	0	0	0	0	0	0	0	1
130	1 296 320 178 0.047	0	1	0	0	0	0	0	0	0	1
131	1 304 339 178 0.057	0	1	0	0	0	0	0	0	0	1
132	1 310 337 178 0.050	0	1	0	0	0	0	0	0	0	1
133	3 310 351 178 0.075	0	0	0	1	0	0	0	0	0	1
134	1 305 335 177 0.035	0	1	0	0	0	0	0	0	0	1
135	0 311 368 178 0.053	1	0	0	0	0	0	0	0	0	1
136	0 320 344 178 0.045	1	0	0	0	0	0	0	0	0	1
137	1 319 364 178 0.082	0	1	0	0	0	0	0	0	0	1
138	0 330 377 179 0.088	1	0	0	0	0	0	0	0	0	1
139	0 319 385 180 0.066	1	0	0	0	0	0	0	0	0	1
140	0 319 387 179 0.083	1	0	0	0	0	0	0	0	0	1
141	0 320 388 179 0.129	1	0	0	0	0	0	0	0	0	1
142		0	0	0	0	0	0	0	0	0	0

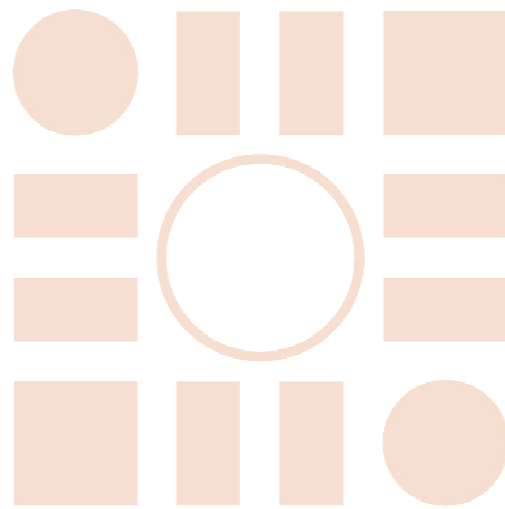
143	0	0	0	0	0	0	0	0	0	0	0
144	0	0	0	0	0	0	0	0	0	0	0
145	0	0	0	0	0	0	0	0	0	0	0
146 0 339 389 180 0.527	1	0	0	0	0	0	0	0	0	0	1
	58	11	0	2	0	0	0	0	0	0	71
	82%	15%	0%	3%	0%	0%	0%	0%	0%	0%	1
	0	1	2	3	4	5	6	7	8		

Tabla 25 - Ejemplo de resultados

Se puede observar, comparando con el ejemplo anterior, basado en la misma secuencia, que los errores han disminuido significativamente.



Universidad  
de Alcalá



ESCUELA POLITECNICA  
SUPERIOR