

**UNIVERSIDAD AUTÓNOMA DE MADRID**  
**MASTER'S PROGRAM IN RESEARCH AND INNOVATION IN INFORMATION**  
**AND COMMUNICATION TECHNOLOGIES (I2-ICT)**  
**Video Analysis Techniques for Surveillance (Course 2012/2013)**

## LAB 2: Tracking

Augusto Bourgeat Terán  
José Luis Carrillo Medina

### Objetivo

El objetivo de este trabajo es el desarrollo de un módulo de seguimiento capaz de seguir una región rectangular dado en los marcos de imagen de una secuencia de vídeo de entrada arbitraria. La secuencia de haber sido capturado por una cámara de vídeo estacionaria. El módulo debe ser implementado en lenguaje C utilizando la biblioteca OpenCV público bajo Linux.

### Problema

Implementar en lenguaje C bajo Linux un módulo de seguimiento básico utilizando la biblioteca OpenCV

Características de la Implementación:

El módulo implementa lo siguientes

- Captura una región de video mediante, la selección de un objeto de interés a través de puntos con el mouse.
- Se captura la Región Rectangular del marco de imagen de referencia (Template) que ha de ser rastreado.
- Se realiza el seguimiento.
- Se Implementa Región Rectangular de referencia (Template) para 6 escalas: la primera de acuerdo a la captura de la imagen real, las siguientes con factores de ampliación/reducción 0.5 , 0.75, 1.5 2.0, 2.5 y 3.0
- Se establecen medidas de confianza en la fiabilidad de los resultados (entre 0.0 y 1.0).

### Palabras claves

Template, bounding box, seguimiento, matching, target

## INDICE

<i>Introducción</i> .....	3
<i>1. Análisis de las secuencias de imágenes</i> .....	3
<i>1.1 Identificar los problemas que presenta cada secuencia de imágenes</i> .....	3
<i>2 Seguimiento basado en algoritmo template matching</i> .....	4
<i>2.1 Búsqueda de una región (template) en una imagen</i> .....	4
<i>2.1.1 Implementación: Búsqueda de una región (template) en una imagen</i> .....	5
<i>2.1.2 Interpretación: Búsqueda de una región (template) en una imagen</i> .....	6
<i>3 TRACKING</i> .....	8
<i>3.1. 1Captura: Video blue_object_1</i> .....	9
<i>3. 1. 2 Seguimiento:</i> .....	9
<i>3.1.3Interpretación</i> .....	9
<i>3.2. 1Captura: Video blue_object_2,</i> .....	9
<i>3. 2. 2 Seguimiento:</i> .....	10
<i>3.2.3Interpretación</i> .....	12
<i>3.3. 1Captura: Video cars2,</i> .....	12
<i>3. 3. 2 Seguimiento:</i> .....	12
<i>3.3.3Interpretación</i> .....	13
<i>3.4. 1Captura: Video caviar (I)</i> .....	13
<i>3. 4. 2 Seguimiento:</i> .....	14
<i>3.4.3Interpretación</i> .....	15
<i>3.5. 1Captura: Video caviar (II)</i> .....	16
<i>3. 5. 2 Seguimiento:</i> .....	16
<i>3.5.3Interpretación</i> .....	16
<i>Conclusiones</i> .....	17
<i>Referencias</i> .....	17

## Introducción

El seguimiento de objetos mediante videocámaras es un tema muy explorado en el campo de la visión por computador debido a su gran utilidad. Esta clase de seguimiento es utilizada en sistemas de reconocimiento basados en movimiento, sistemas de videovigilancia, adquisición de metadatos para bases de datos de vídeo, interacción hombre-máquina, etc.

En el tema de la videovigilancia, es un hecho que el análisis de vídeo requerido para un sistema de este tipo se divide en tres principales etapas, las cuales son: Detección de objetos interesantes en escena, seguimiento de éstos o Tracking y posteriormente su reconocimiento.

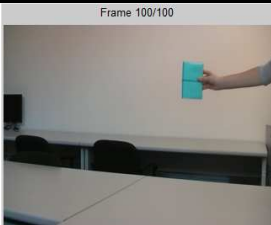

Este trabajo se desarrollo en el lenguaje C utilizando la librería OpenCv 2.4.2 (Librería de manejo de imágenes) bajo Lynus.


## 1. Análisis de las secuencias de imágenes

### 1.1 Identificar los problemas que presenta cada secuencia de imágenes

Existen múltiples factores que afectan a la precisión de un algoritmo de seguimiento de objetos (cambios de escala e iluminación, objetos similares en el fondo, movimiento complejo,...). Con el fin de poder realizar un mejor estudio de los algoritmos se realiza un análisis previo de las secuencias de imágenes a utilizar.

Las secuencia de Video **blue\_object\_1**, corresponde a un objeto de color celeste moviéndose en la escena, fácilmente detectable. **Cars2** que corresponde a una intersección de una avenida. La secuencia **caviar**, se observa a tres personas caminando por un pasillo.

Secuencia	Problemas que podría tener en el seguimiento de objetos	
blue_object_1	Cambio de escala, al alejarse y acercarse  Variación de intensidad luminosa.  Cambios muy rápidos de velocidades del objeto, incluso las trayectorias son cambiadas abruptamente.	
cars2	Muchos Autos y personas en movimiento, a una gran distancia.  Oclusiones  Cambio de escala, al alejarse	

caviar	<p>Cambio de escala, al alejarse</p> <p>Oclusión de objetos, al hacer el seguimiento de las 3 personas que se alejan de la cámara.</p>	
--------	--	---

## 2 Seguimiento basado en algoritmo template matching

El ajuste de plantillas (o *template matching*) es uno de los algoritmos más utilizados para el seguimiento de objetos. Este algoritmo se puede aplicar para seguir objetos en secuencias de imágenes mediante la definición de una región de la imagen que represente al objeto (llamado plantilla, modelo o *template*) y su búsqueda en posteriores imágenes utilizando diversas medidas de similitud.



Figura 1: Ejemplo de funcionamiento de *Template Matching* ([http://en.wikipedia.org/wiki/Template\\_matching](http://en.wikipedia.org/wiki/Template_matching))

La medidas de similitud utilizada corresponde a la suma de las diferencias al cuadrado (*sum of squared differences, SSD*), que se define como:

$$SSD = \sum_{x,y \in W} (R(x,y) - T(x,y))^2$$

donde  $T(x, y)$  es un píxel de la región (template)  $T$  que representa al objeto,  $R(x, y)$  es un píxel de una región  $R$  extraída de la imagen donde queremos localizar el objeto y  $W$  es el tamaño del template.

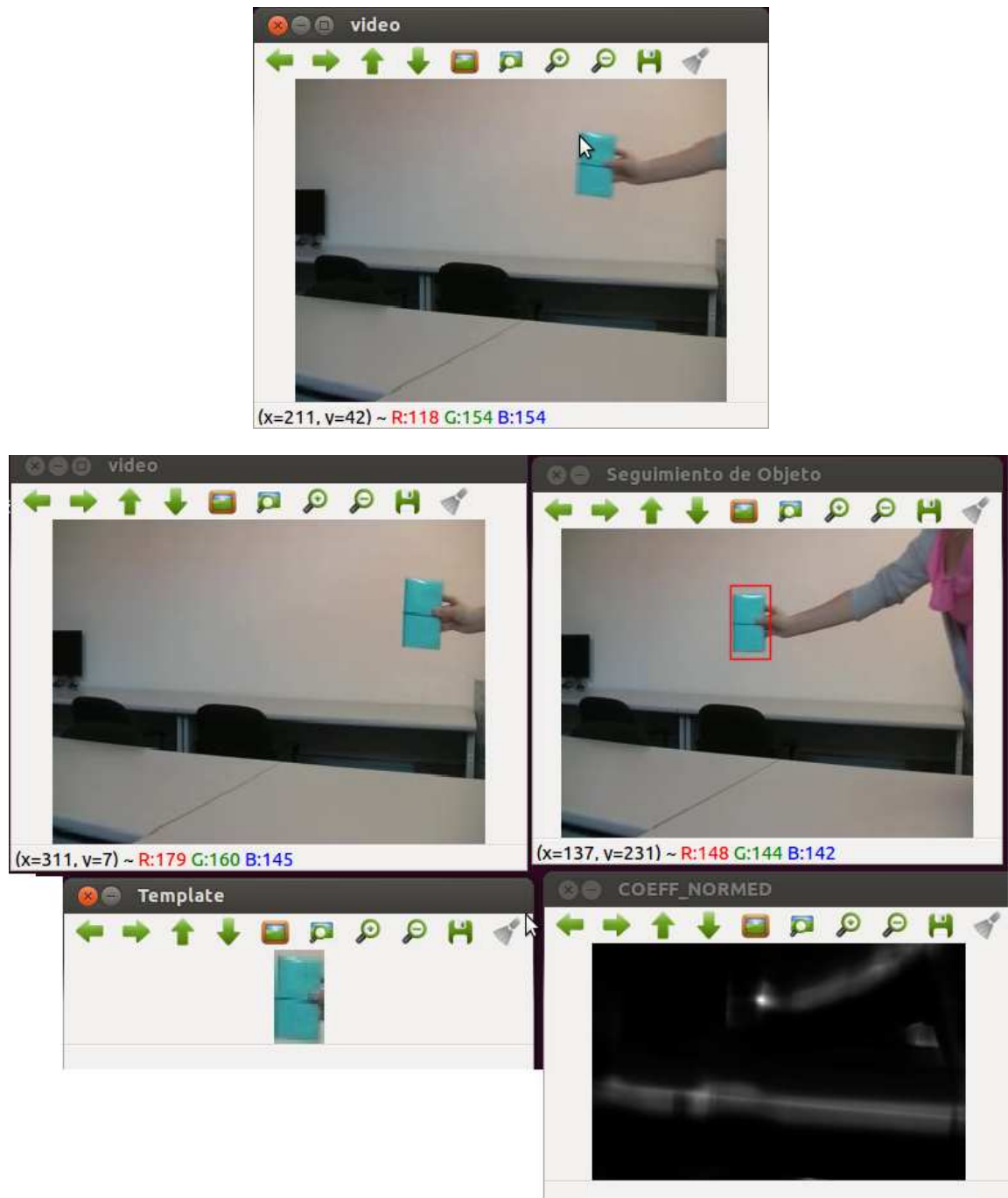
El uso de SSD para buscar una región  $T$  en una imagen  $I$  requiere extraer regiones de tamaño  $W$  para cada píxel  $(x, y)$  de la imagen  $I$ . Finalmente, la localización de la región de interés se obtiene como el valor mínimo de la imagen  $SSD(x, y)$ .

### 2.1 Búsqueda de una región (template) en una imagen

Se genera un programa que cargue dos imágenes, selecciona un objeto en la primera de ellas (descrito mediante su rectángulo englobante o **bounding box**), posteriormente se realiza su búsqueda en la segunda imagen mediante SSD y finalmente muestre por pantalla los resultados de la nueva localización del objeto (centro y *bounding box*) e imagen SSD. Se hace uso de las funciones `cvMatchTemplate`, `cvNormalize`, `cvRectangle` y `cvMinMaxloc`

Para realizar el seguimiento de un objeto de interés primero se detiene la imagen de video, con la tecla ESC, luego mediante un click izquierdo del mouse, se selecciona la primera coordenada del bounding box (esquina superior-izquierda), y con un click derecho del mouse se selecciona la segunda coordenada (esquina inferior-derecha).

### 2.1.1 Implementación: Búsqueda de una región (template) en una imagen



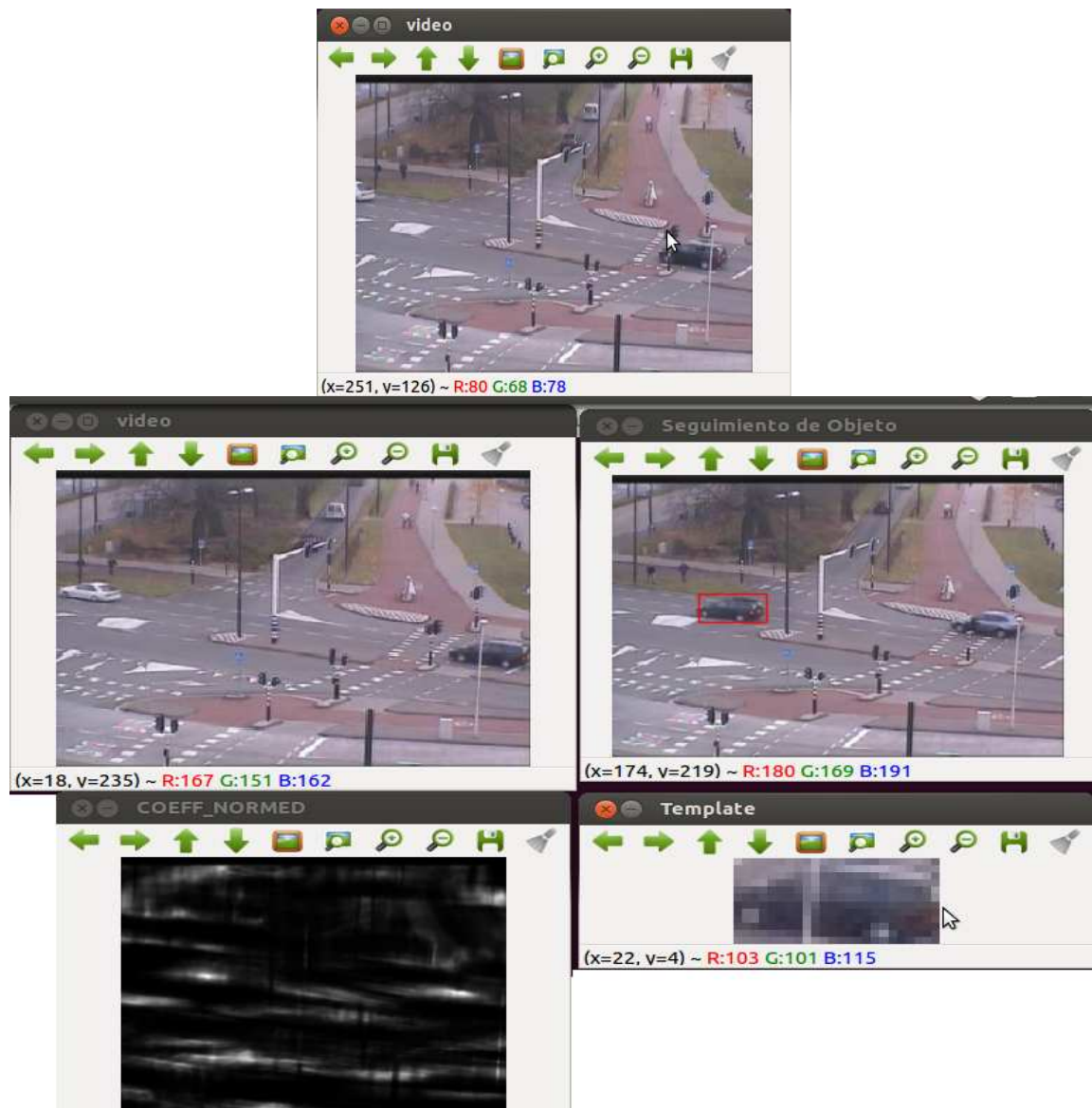
### 2.1.2 Interpretación: Búsqueda de una región (template) en una imagen

De las figuras se observa:

En la primera imagen corresponde al instante en que se captura la región de interés (template). La segunda es el seguimiento de la imagen capturada en el video. La tercera es la región de interés. La cuarta imagen corresponde al punto que más se parece a la región de interés dentro de la imagen.

Como podemos observar, el punto blanco más brillante es la zona en la que los errores cuadráticos son muy cercanos a cero; que corresponde a la zona donde la similitud al objeto buscado es la más aceptada, para ello se utiliza la función cvMatchTemplate.

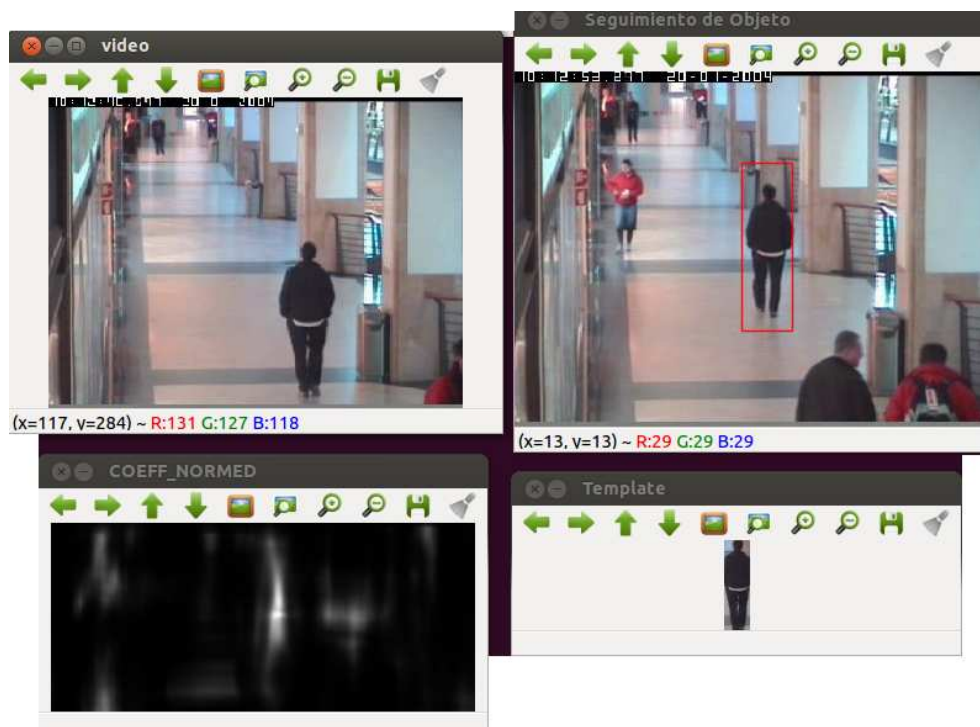
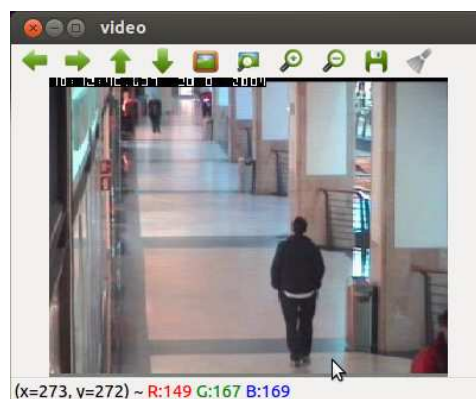
Se repite este experimento con la secuencia de video de una carretera con varios automóviles y se quiere realizar el seguimiento a un automóvil.



De las figuras se observa:

En la primera imagen corresponde al instante en que se captura la región de interés (automovil), la segunda es el seguimiento de la imagen capturada, la tercera es imagen corresponde al punto que más se parece a la región de interés dentro de la imagen y la tercera es la región de interés, pudiéndose observar que el seguimiento se realiza en la mayoría de las secuencias de video denotándose que existen pérdidas debido especialmente cuando aparece otro automóvil que se supone tiene las mismas características, pero el seguimiento se recupera. En las últimas imágenes no se puede hacer seguimiento por cuanto el objeto ha cambiado su escala, y el target se ha alejado de la cámara.

Utilizamos la secuencia track3 correspondiente a personas que caminan en un pasillo y se realiza el seguimiento a una persona.





De las figuras se observa:

En la primera imagen corresponde al instante en que se captura la región de interés (persona), la segunda es el seguimiento de la imagen capturada, la tercera es imagen corresponde al punto que más se parece a la región de interés dentro de la imagen y la tercera es la región de interés, pudiéndose observar que el seguimiento se realiza en la mayoría de las secuencias de video denotándose que existen pérdidas debido especialmente cuando se va alejando de la cámara. En las últimas imágenes no se puede hacer seguimiento por cuanto el objeto ha cambiado su escala, y el target se ha alejado de la cámara.

### 3 TRACKING

Se ha implementado un módulo de seguimiento capaz de seguir una región rectangular dado en los marcos de imagen de una secuencia de vídeo de entrada arbitraria. El módulo esta implementado en lenguaje C utilizando la biblioteca OpenCV bajo Linux.

Primeramente Captura una región de video mediante, la selección de un objeto de interés a través de puntos con el mouse.

Luego la región rectangular de referencia (Template) se escala para 7 tamaños: la primera de acuerdo a la captura de la imagen real, las siguientes con factores de ampliación/ reducción 0.5 , 0.75, 1.5 2.0, 2.5 y 3.0, creación plantillas para posibles reducciones o aumentos de los objetos



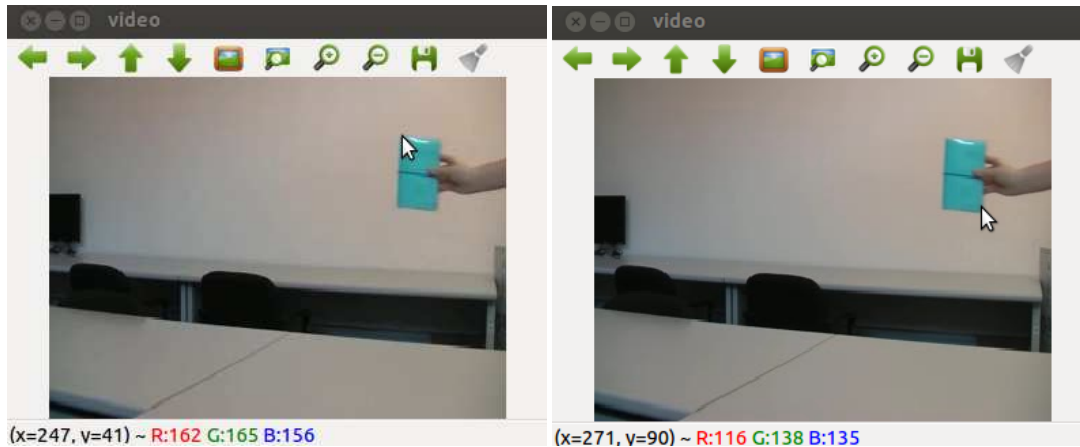
El programa cargue las 7 imágenes que se van a utilizar como referencia (Template), realiza una por una la búsqueda dinámica en la imagen actual mediante el uso de la función `cvMatchTemplate`, encontrando 7 valores para cada referencia (Template) de los sectores que más se parecen a las imágenes.

A continuación se selecciona el Template que tiene mayor valor de similitud en la imagen, este representa el que mejor se ha adaptado a la escalas, y con los datos descrito mediante su rectángulo englobante, se realiza el seguimiento en la Imagen actual.



Ejemplos:

### 3.1. 1Captura: Video blue\_object\_1



### 3. 1. 2 Seguimiento:

Las primera secuencia corresponde al Video **blue\_object\_1**, corresponde a un objeto de color celeste moviéndose en la escena, fácilmente detectable, y su tamaño no cambia apreciablemente.



### 3.1.3Interpretación

En el seguimiento de esta secuencia de video no se observa ningún problema en la detección, sin ninguna pérdida de trayectoria.

### 3.2. 1Captura: Video blue\_object\_2,

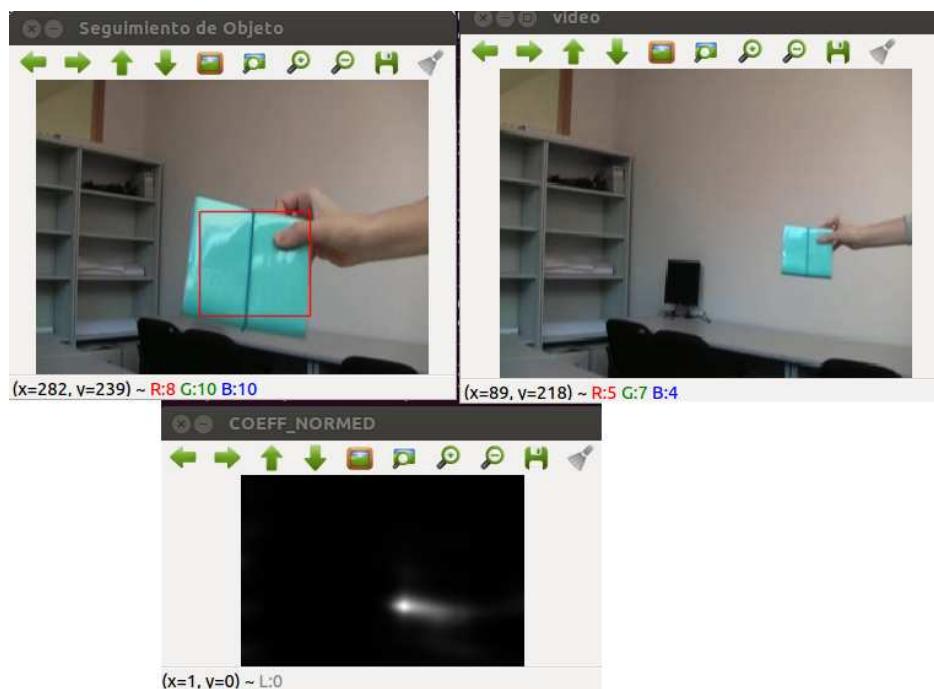
Las segunda secuencia corresponde al **Video blue\_object\_2**, corresponde a un objeto de color celeste moviéndose en la escena, fácilmente detectable, y su tamaño cambia considerablemente.



### 3. 2. 2 Seguimiento:

El Objeto se acerca y se aleja de la cámara, existiendo un cambio de escala.





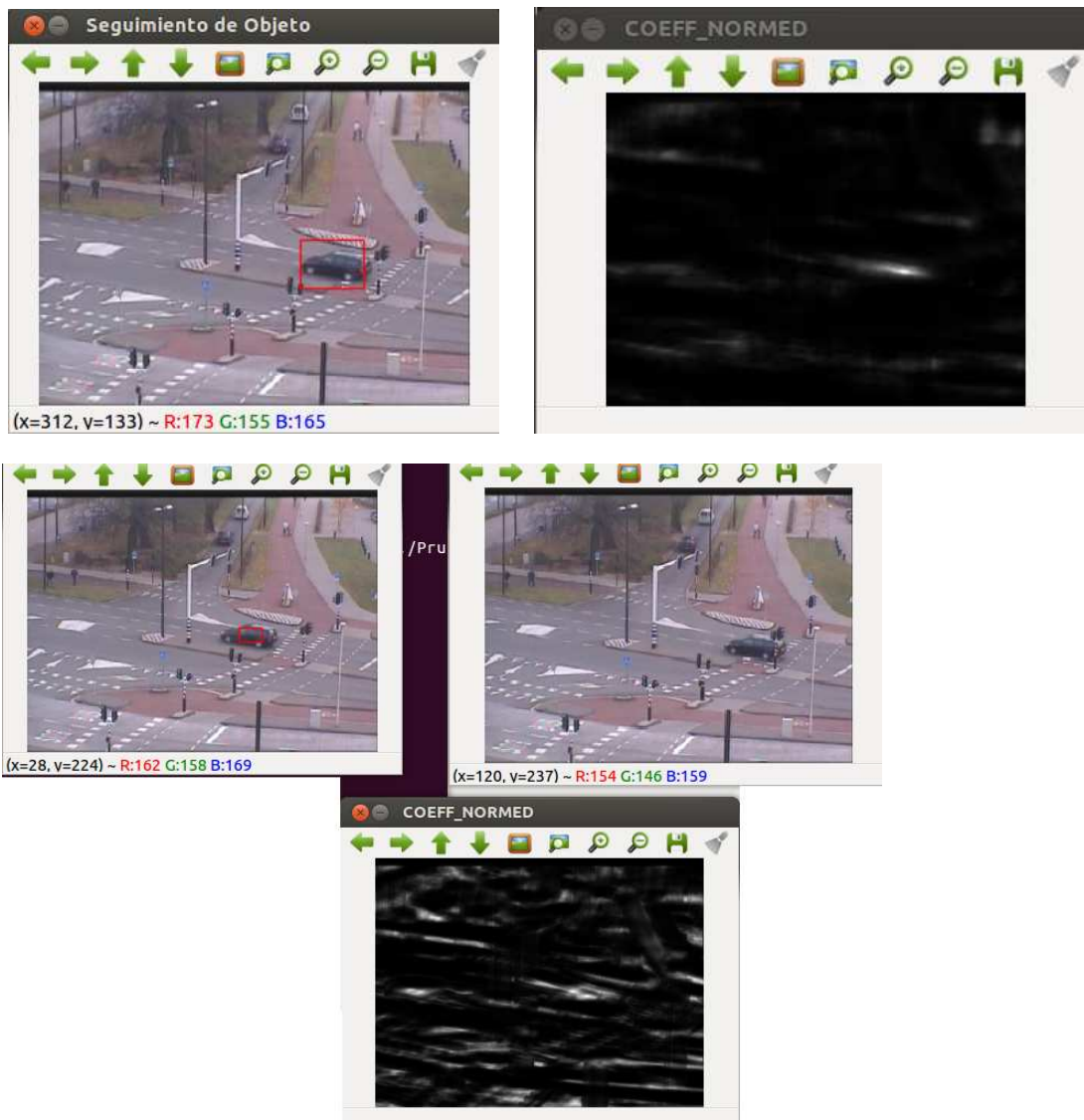
### 3.2.3 Interpretación

La primera imagen corresponde al seguimiento de la imagen o Tracking, la segunda imagen corresponde al instante en que se captura la región de interés y la tercera es imagen corresponde al punto que más se parece a la región de interés dentro de la imagen; se puede observar que el seguimiento se realiza en la secuencia de video notándose que existen seguimiento y la adaptación a diferentes escalas.

#### 3.3. 1 Captura: Video cars2,

#### 3.3. 2 Seguimiento:

Una tercera secuencia correspondiente al **Video cars2**, corresponde una intersección de una avenida, y el movimiento de vehículos fácilmente detectable, y su tamaño cambia a medida que se aleja de la cámara.





### 3.3.3 Interpretación

La primera imagen corresponde al seguimiento de la imagen o Tracking, la segunda imagen corresponde al instante en que se captura la región de interés y la tercera es imagen corresponde al punto que más se parece a la región de interés dentro de la imagen; se puede observar que el seguimiento se realiza en la secuencia de video notándose que existen seguimiento y la adaptación a diferentes escalas.

En esta secuencia, inicialmente el seguimiento al vehículo lo hacía correctamente, en algunas imágenes el seguimiento tiene problemas en el escalamiento posiblemente por la cantidad de zonas donde el target que se compara es muy probable.

### 3.4. 1 Captura: Video caviar (I)

Una cuarta secuencia correspondiente al **Video caviar**, se observa a personas caminando por un pasillo, el movimiento de personas fácilmente detectable, y su tamaño cambia a medida que se aleja o se acercan a la cámara.

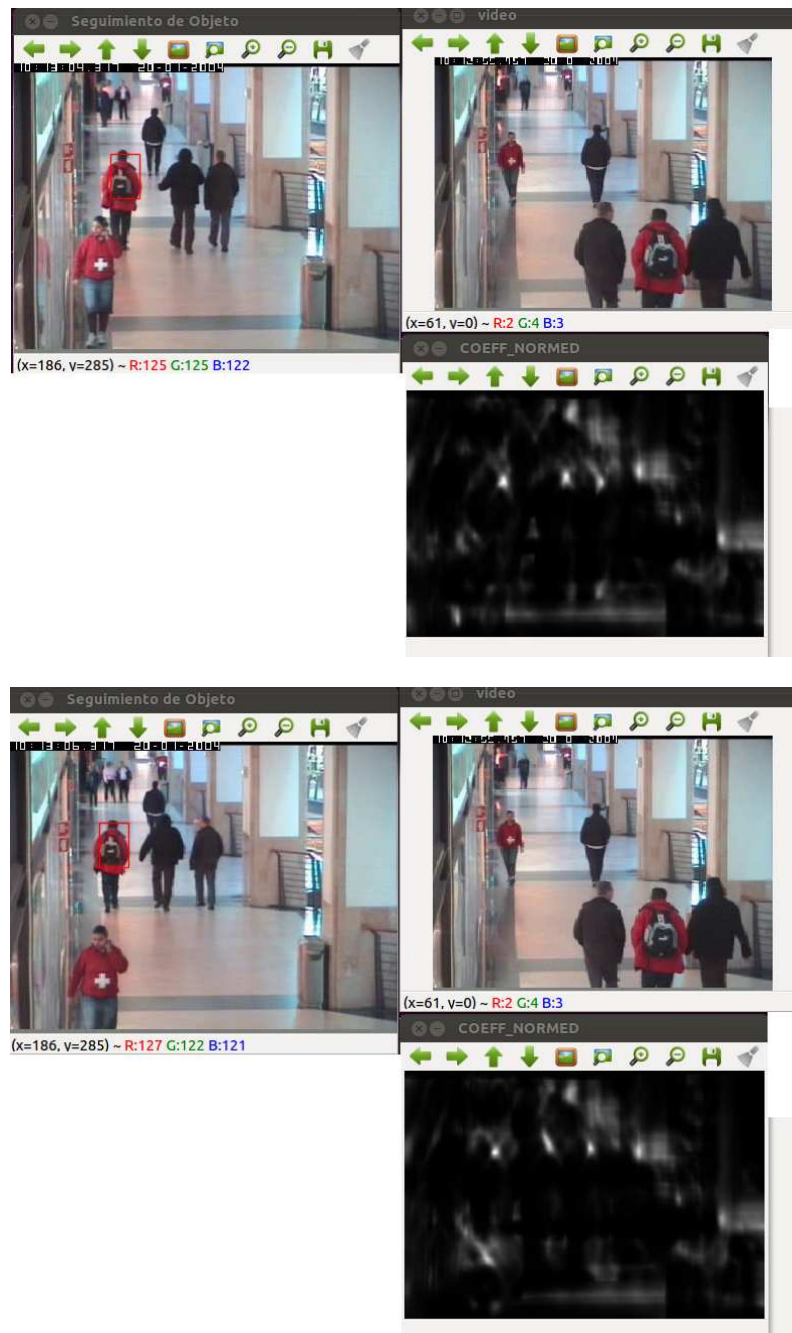
la mochila del estudiante, como objeto que queremos hacer el seguimiento.





### 3. 4. 2 Seguimiento:





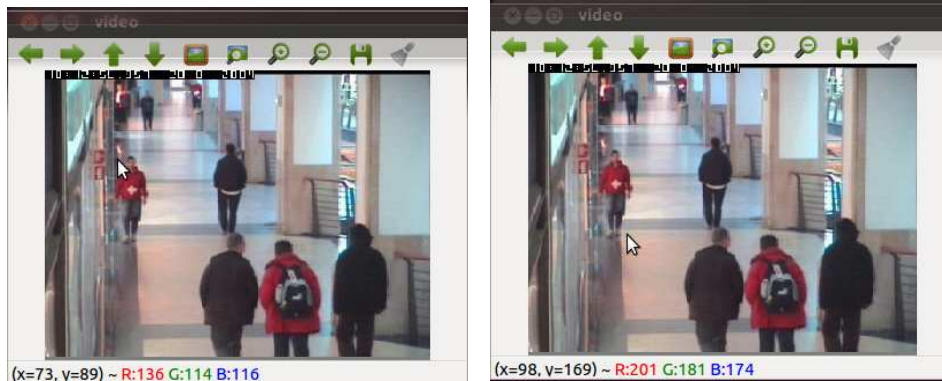
### 3.4.3 Interpretación

En esta secuencia el seguimiento lo hacía correctamente, se puede hacer seguimiento por cuanto el factor de escala a cambiado, debido a que las imágenes y el target se va adaptando a medida que la persona se aleja de la cámara y el tamaño del cajón de búsqueda se acomoda. En algunas imágenes el seguimiento se pierde, especialmente cuando el objeto cambia o rota.



### 3.5. 1Captura: Video caviar (II)

Captura: La persona que tiene un sweater con una cruz



### 3. 5. 2 Seguimiento:



### 3.5.3 Interpretación

Se observa, el acercamiento de la persona a la cámara y el cambio de escala del objeto en el seguimiento en el tiempo, con respecto al bounding box.

## Conclusiones

En esta práctica se han desarrollado con OpenCv diversas técnicas para realizar el seguimiento de objetos (target) en secuencias de imágenes.

El método utilizado en la búsqueda de patrones es el Template Matching, que es una comparación de plantillas.

El seguimiento de objetos de interés en vídeo es la base de muchas aplicaciones que van desde la producción de vídeo hasta la vigilancia remota, y desde la robótica hasta los juegos interactivos.

Los algoritmos implementados dan un buen seguimiento a los cambios de velocidad, cambios de escala, oclusión, y cambios muy bruscos.

En la implementación de los algoritmos el sistema de seguimiento ha podido detectar y seguir el objeto o elementos de una escena, solventando problemas que se pueden encontrar como la oclusión, los cambios de escala, cambios de velocidad o los cambios en la iluminación que impidan una correcta detección y el posterior seguimiento.

El costo computacional en los algoritmos de tratamiento de imágenes es grande, pero se puede mejorar programando con OpenCv.

## References

- [1] Apuntes de Clase, Video Analysis Techniques for Surveillance Unit 2: Video Analysis for Surveillance (Part I), Miguel Ángel García miguelangel.garcia@uam.es
- [2] Introduction to the Analysis of Video, Sequences Introduction to Scene Analysis: Object tracking, Luis Salgado Álvarez de Sotomayor, Luis.Salgadoa@uam.es
- [3] HERRAMIENTA PARA EL ETIQUETAJE DE SECUENCIAS DE IMÁGENES, [http://www.recercat.net/bitstream/handle/2072/43693/PFC\\_CesarCuadrosCortina.pdf?sequence=1](http://www.recercat.net/bitstream/handle/2072/43693/PFC_CesarCuadrosCortina.pdf?sequence=1)

**Realizado por:**

**Augusto Bourgeat**  
**José Luis Carrillo**

**Program:** Master's program in Research and Innovation in Information and Communication Technologies (I2-CIT)

**Center:** Escuela Politécnica Superior

**University:** Universidad Autónoma de Madrid  
Madrid 2013