

# VISIÓ PER COMPUTADOR

## *Short Project*

**Facultat d'Informàtica de Barcelona**

**Manel Frigola  
Joan Climent**

**Barcelona, Abril de 2023**

## 1. Objectius del projecte

L'objectiu del projecte és implementar un sistema de seguiment d'objectes, animals o persones en moviment en seqüències de vídeos. Proposem treballar amb un conjunt de dades de vídeos amb anotacions anomenat TLP (de Track Long and Prosper, <https://amoudgl.github.io/tlp/>). Aquesta base de dades pretén ser un punt de referència per al seguiment d'objectes visuals per a comparar algoritmes de seguiment en imatges de vídeo. El conjunt de dades consta de 50 vídeos d'escenaris del món real, que abasten una durada de més de 400 minuts, cosa que el fa el més gran en durada mitjana per seqüència en comparació amb altres conjunts de dades genèrics existents per al seguiment visual.

Les imatges amb les que es treballaran seran principalment amb una part reduïda del TLP anomenada TinyTLP. El conjunt de dades TLP consta de 50 seqüències d'imatges ( $1280 \times 720$  pixels) amb anotacions del quadre delimitador de l'objecte a seguir. Cada seqüència consta d'un únic objecte a seguir, indicat al primer fotograma en el fitxer d'anotacions. Les posicions reals d'on es troba l'objecte a seguir (*bbox*) les trobareu en un fitxer que acompanya al vídeo anomenat `groundtruth_rect.txt`.

Les anotacions del quadre delimitador de l'objecte en cada seqüència d'imatges tenen el format següent:

[ID del frame, xmin, ymin, amplada, alçada, isLost] on cada ítem indica:

frameID: numero de frame.

xmin - Coordenada x superior esquerra del quadre delimitador.

ymin - Coordenada y superior esquerra del quadre delimitador.

amplada - Amplada del quadre.

altura - Alçada del quadre.

isLost - Si és 1, l'objecte objectiu no és visible en absolut; altrament 0.

## 2. Treball a realitzar

Les tasques del projecte es dividirien en dues parts principals. La primera seria fer un algoritme de seguiment de l'objecte/animal o persona mitjançant tècniques de *Tracking*. La segona part seria fer la detecció de l'objecte per reconeixement, trobant un bon descriptor d'imatges que funcioni correctament per a detectar la seva posició.

- Construir un sistema de *Tracking* de codi propi (o d'altri per si voleu comparar) que a partir de la posició del quadre delimitador enèsim detecti on està l'objecte en la imatge  $n+1$ . Evidentment, en el codi no es podrà utilitzar la informació de la finestra correcta (`groundtruth_rect.txt`) a excepció per calcular l'encert de la detecció. Opcionalment, es podrà fer servir la predicció de la posició actual (veure filtre de Kalman en l'apartat *Indicacions*) a partir de la posició anterior.
- Un sistema de detecció de codi propi que analitzi tota o part de la seqüència d'imatges per calcular unes característiques que permeti trobar l'objecte mitjançant tècniques de

reconeixement. Per analitzar la seqüència d'imatges es podrà utilitzar la informació de la finestra correcta (*groundtruth\_rect.txt*) però aquesta mateixa informació no es podrà utilitzar en l'algoritme de reconeixement.

### 3. Es demana

**Un informe didàctic i complet.** Cal que lliureu un informe comentant de cadascuna de les etapes amb una descripció del treball realitzat, resultats obtinguts i destacant les parts més rellevants del vostre programa, quins problemes us heu trobat i com els heu resolt (o no).

En especial l'informe ha de contenir:

- i) Una explicació detallada del mètode, amb els resultats obtinguts. Per mesurar els resultats obtinguts experimentalment, podeu fer servir els valors correctes de referència (*groundtruth\_rect.txt*), indicant el nº de *frames* amb prou encert, primer *frame* en que la pèrdua és total. En l'apartat de reconeixement indiqueu la taula de confusió de l'entrenament del vostre classificador.
- ii) Enumeració de les funcions utilitzades, deixant ben clar quines són les implementades per vosaltres. Tot el software provinent d'altres fonts, ha **d'estar correctament referenciat, en cas contrari es considerarà plagi.**

**iii) L'última secció de l'informe serà un annex amb tot el codi.**

### Presentació del funcionament del programa

Haureu d'executar el programa que implementi el sistema de reconeixement demanat en aquesta pràctica. Tingueu en compte que el dia de la presentació es provarà el vostre programa de reconeixement amb noves imatges, no incloses en el *dataset*.

### 4. Indicacions

Les següents indicacions són de caràcter orientatiu.

#### a) Esquelet inicial del programa

Per construir el vostre sistema de *Tracking* us proposem utilitzar el següent programa inicial. Aquest programa es comentarà a classe durant les hores assignades al Short Project

```
% myTracker, codi inicial del short project

close all

% Llegim el fitxer d'anotacions groundtruth_rect.txt: frame, bounding boxes, is_lost
BB = importdata('./Aquarium2/groundtruth_rect.txt');
ldir = dir('./Aquarium2/img/*.jpg');

figure
hold on % mantenim sempre oberta la mateixa figura

% filtre de Kalman
initialStateP = [BB(1,2) 0 BB(1,3) 0]; % posicio inicial i velocitat
initialStateW = [BB(1,4) 0 BB(1,5) 0]; % mida inicial i velocitat
KFP = trackingKF('MotionModel','2D Constant Velocity','State',initialStateP);
KFW = trackingKF('MotionModel','2D Constant Velocity','State',initialStateW);
```

```

nf = size(Iidir); % nombre total de fitxers imatges
for i = 1:nf
    filename = horzcat(Iidir(i).folder,'/',Iidir(i).name);
    I = imread(filename);

    predP = predict(KFP); % predictor de la posicio de la finestra per Kalman

    correct(KFP,BB(i,2:3)); % indiquem la nova posicio detectada

    predW = predict(KFW); % predictor de l'amplada/alcada de la finestra per Kalman

    correct(KFW,BB(i,4:5));

    %calculem el -overlap, intersection over union-- de les finestres
    overlapRatio = bboxOverlapRatio([predP(1),predP(3),predW(1),predW(3)],BB(i,2:5))

    % mostrem els resultats sobre la imatge
    imshow(I); % mostra el frame
    % mostrem els rectangles: correcte (en groc) i el predit (en vermell)
    rectangle('Position',[predP(1),predP(3),predW(1),predW(3)],'EdgeColor','red');
    rectangle('Position',BB(i,2:5),'EdgeColor','yellow');
    drawnow
end

```

Fig 3. Codi d'exemple inicial.

## 5. Valoració del projecte

Els factors que més es valoraran del vostre projecte són:

- Un informe del projecte ben estructurat i complet.
- Comparació del resultat usant diferents tècniques i descriptors, avaluant les seves prestacions, etc.
- Utilització de tècniques que permetin una reducció del cost d'execució a l'hora que es manté un rati elevat d'encerts en la classificació/detecció.
- Avaluar el funcionament del vostre programa amb imatges de test prou variades.