Sorbonne Université - Master 2 IMA
UE VISION - Practical Work
Object Tracking in Videos

January 2026

The goal of this practical work is to understand the challenges and difficulties of visual tracking, to experiment and develop solutions based on Mean Shift and General Hough Transform algorithms.

To this end, we will use OpenCV library with Python (code tested with python3 and OpenCV 4.1.0). Basic code and test videos are available on the following pages :

— `https://perso.ensta-paris.fr/~manzaner/Cours/IMA/VISION/Tracking_MeanShift.py`
— `https://perso.ensta-paris.fr/~manzaner/Cours/IMA/VISION/Test-Videos.zip`

You are expected to write a report (one report for two students) and upload the PDF file to the Moodle, or send it to antoine.manzanera@ensta.fr.

The purpose of a tracking algorithm is to provide, for each video frame, the position of an object of interest, given here by the coordinates of the bounding rectangle (RoI), whose values are manually initialised on the first image of the video (the rectangle is to be defined using the mouse left button, and validated with the key 'q').

1 Mean Shift

Q1 Experiment the tracking performed by the provided code `Tracking_MeanShift.py` that uses the Mean Shift algorithm, on the marginal density f_H of the hue component H. Briefly recall the principle of this algorithm, discuss and illustrate its advantages and limits by your experiments.

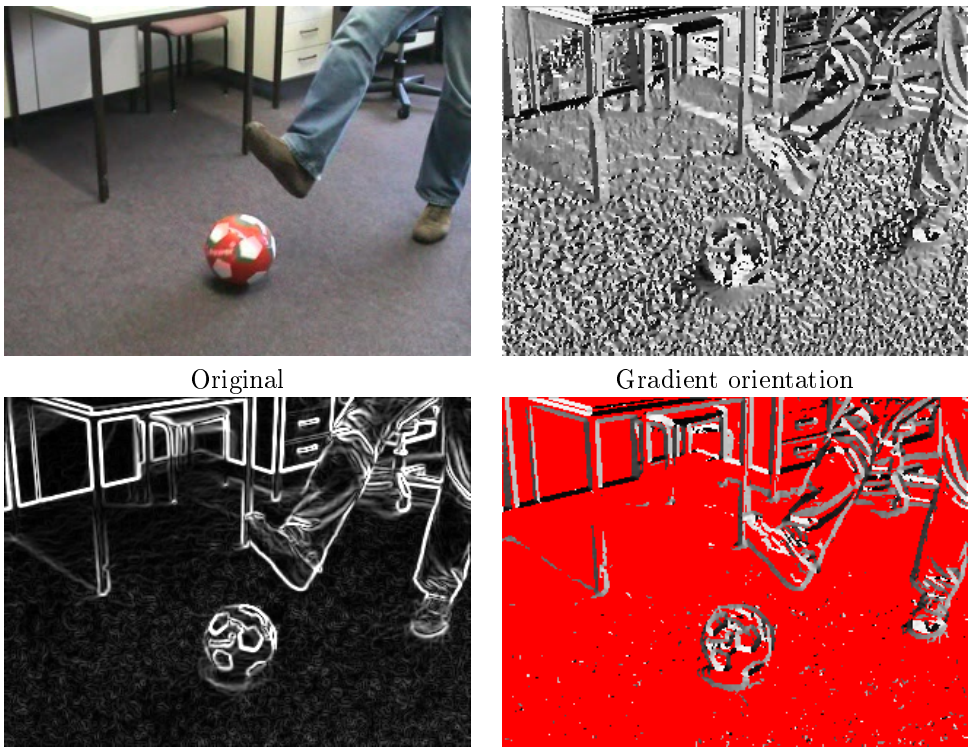
Q2 Analyse more in-depth the result by displaying the sequences of hue images, and also the weight images, calculated from the back-projection R_H of the hue histogram f_H :

$$R_H(x, y) = f_H\left(H(x, y)\right).$$

Try to improve the tracking results, by performing one or several of the following changes :
— changing the computed density
— combining several weight maps
— updating recursively the model histogram

2 Generalised Hough Transform

Q3 Calculate for each frame, the local orientation, i.e. the gradient argument of pixels, and also the gradient magnitude. Use a threshold on the gradient magnitude to mask pixels whose orientation



Original | Gradient orientation

Gradient magnitude | Masked orientations

FIGURE 1 – Computing the index of the R-Table (gradient orientation), with selection of the voting pixels (based on the gradient magnitude).

is not significant. Display the sequence of orientations, where the masked pixels appear in red. An example of expected result is shown on Figure 1.

Q4 BUILD a model of the initial object under the form of an implicit model indexed on the orientation (R-Table), calculated on significant (unmasked) pixels. Then, calculate the associated Hough transform on all the images of the sequence. Calculate the straightforward tracking, corresponding to the maximal value of the Hough transform at each image. Comment and criticise the obtained results. Illustrate your answers by examples of Hough Transforms and the corresponding detections (see an example on Figure 2).

Hough Transform $H(\mathbf{x})$ | Detection by $\arg\max_{\mathbf{x}} H(\mathbf{x})$

FIGURE 2 – Tracking by Hough Transform detection.

Q5 PROPOSE a *Prediction* strategy by considering the whole vote map (and not only its arg max), as a state likelihood to make the tracking more stable.

Q6 PROPOSE an *Update* strategy of the model to make the tracking more robust to aspect changes of the object.

3 Deep Features

Q7 PROPOSE a way to improve the two previous approaches (i.e. histogram and Hough based) by using components from a collection of Feature Channels learned by a Deep Neural Network. Take a pre-trained network of your choice, choose a layer and use the corresponding features within the Mean Shift or Hough algorithm, as you prefer.

— On which task was your network trained and why ?
— How to choose the best layer ?
— How to choose the best channels ?