INF6804 Computer Vision
H2020 – Practical Assignment 3
Detection and visual object tracking

Objectives:

- Allow the student to learn about object detection and/or tracking methods for image sequences.

Submission:

- *All your source code* (we should be able to run your tests)
- A report (*.pdf format* of 8 to 15 pages with font size of 10)
- Submit before April 16th, 5:00 PM, on Moodle –*late submissions will not be accepted*
- *You must also submit your report on TurnItIn*
  - Register at www.turnitin.com using the info available on Moodle!

References:

- See course notes on Moodle (Chapter 5)

Other directives:

- The assignments must be made in teams of two, submit only one version of your work!

Presentation

In this assignment, you will have to propose a method capable of detecting and tracking a single object of interest through a video sequence. The object in question is identified via a bounding box in the first frame of the video; your goal is to relocate it similarly in all other frames of the sequence.

Your task is to find a method that works well, under various conditions (e.g. occlusions, scale changing, etc.). To do this, you will test your method on **three** sequences from the VOT2013 benchmark, each sequence must represent specific conditions.

For the method itself, you can once again implement your own solution, or use/modify an existing one. Note that two approaches can be considered here: you can use a solution based on tracking by detection (which implies using an object detector trained a priori), or use a classic visual object tracking solution (i.e. based on the modelling of the object of interest using only the first frame, and searching for it in the subsequent ones).

The final evaluation of your method will be done based on the robustness and the precision, depending on the calculation of the intersection over union (IoU) between the bounding boxes produced by your method and the ground truth. We also want you to try different values of **IoU {30% ; 50% ; 70%}**.

However to help you, a part of the code is provided in the file '***main.py***', in which you only have to implement your tracking method *track()*. Like before, the presentation of the method, its strengths and weaknesses, and your experiments are the most important parts of this work.

In your report, you have to include the following elements (marked on 20 pts):

1. Detailed presentation of the method (4 pts) :

   In your own words, give the description and principles of your method. What are its strengths/weaknesses?

2. Description of experiments, datasets and evaluation criteria (3 pts) :

   Describe in detail the realized experiments. What are the difficulties of the chosen sequences? Explain the evaluation criteria at your disposal?

3. Description of the implementation used (3 pts) :

   Describe the implementation of the proposed method. If you did not write all the code yourself, where does it come from? Did it require modifications? Otherwise, from which papers or websites did you inspire yourself to write it? In all cases, what are the primary parameters of your methods? How did you set their values?

4. Experimentation results for validation tests (3 pts) :

Provide the evaluation results from your experiments. Use a proper format for their presentation — tables, figures, ...

5. Discussion of results (4 pts) :

   Discuss the results in relation with the challenges identified on the chosen sequences. Which challenges seem to be resolved by your method? Which are not? How could you improve these tests?

6. Readability and completeness (3 pts) :

   In addition to the content, the format must be clean and complete.

For the dataset, we ask that you use the *VOT2013* benchmark mentioned in the resources section.

During the lab periods, do not hesitate to ask questions to the TAs by chat/video-conference— they can help you with any technical issue if you are working on Windows/Linux, or if you are coding in C/C++, Python or Matlab.

You will be penalized by 50% of the total grade if you do not hand in your code. Also, if your report is not submitted to TurnItIn, it will not be graded. The order of presentation for the topics listed above does not matter, as long as they are all present.

Resources

Datasets:

- VOT 2013 Benchmark (https://www.votchallenge.net/vot2013/)

Vision libraries:

- OpenCV (https://docs.opencv.org/4.0.0/d9/df8/tutorial_root.html)
- scikit-image (https://scikit-image.org/docs/stable/auto_examples/index.html)

Deep learning frameworks:

- PyTorch (https://pytorch.org/tutorials/)
- Tensorflow (https://www.tensorflow.org/tutorials)

Python:

- Guide (https://wiki.python.org/moin/BeginnersGuide/Programmers)
- NumPy (https://docs.scipy.org/doc/numpy/user/quickstart.html)

Matlab:

- Guide (http://www.mathworks.com/help/pdf_doc/matlab/getstart.pdf)
- Cheatsheet (http://web.mit.edu/18.06/www/Spring09/matlab-cheatsheet.pdf)