INF6803

VIDEO PROCESSING AND APPLICATIONS

H2018 – Practical Assignment #3

Detection and tracking of an object of interest

Objectives :

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

Submission:

- All your source code (if we need to re-run your tests, we should be able to!)
- A report (in **.pdf format**), between 8 and 15 pages (roughly)
- Submit before April 17th, 3:00 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

Other directives :

- The assignments can be done alone or in teams of two; however, submit only one version of your work!

Presentation

In this assignment, you will have to propose a method capable of detecting and tracking an 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.

We provide on Moodle a video sequence for which only the first bounding box is available. Your final results will be evaluated on this sequence after the submission deadline using our "secret" groundtruth, and a ranking of the methods submitted by all teams will then be published. Note that your rank does not influence your grade in this assignment --- like previously, the presentation of the method, its strengths and weaknesses, and your experiments are the most important parts of this work (see the marking scheme at the end).

The evaluation for the final ranking will be based on the Intersection-over-Union (IoU) metric obtained for the produced bounding boxes with respect to the groundtruth of the sequence provided on Moodle. In order to validate the performance of your method during its development, we expect you to evaluate it using another public dataset (e.g. VOT2013).

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 choice of the solution must reflect the challenges present in the video sequence provided on Moodle.

In your report, along with an in-depth description of your solution, you will have to:

1. Identify the challenges specific to the video sequence provided on Moodle for the detection and tracking of the object of interest. In other words, what makes this sequence hard to process?

2. Justify the choice of the proposed solution with respect to the challenges identified above. Which problems would your solution solve easily? Which might still cause some issues, and why?

3. Describe the implementation of the proposed solution. 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 from to write it? In all cases, what are the primary parameters of your methods? How did you set their values?

4. Provide the evaluation results from your validation experiments on the other public dataset (e.g. VOT2013). Use a proper format for their presentation --- we expect to see tables, figures, ...

5. Discuss the results of the fourth point in relation with the challenges identified in the first point. Which challenges seem to be resolved by your solution? Which are not? Can we expect a similar performance for the sequence on Moodle?

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

Marking scheme

Report:

- Presentation of the solution (get into the details! 1-2 pages) = 4 pts
- Identification of the challenges for the Moodle sequence = 3 pts
- Justification of the solution w.r.t. identified challenges = 2 pts
- Description of the implementation used = 2 pts
- Experimentation results for validation tests = 3 pts
- Discussion of results = 3 pts
- Readability, property, and completeness = 3 pts

Total on 20 pts.

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. For the length, we expect something between 8 and 15 pages, but if you produce something longer, it does not matter too much (just please don't hand in 50 pages). The order of presentation for the topics listed above does not matter either, as long as they are all present.

References

- Matlab Cheat Sheet:

  http://web.mit.edu/18.06/www/Spring09/matlab-cheatsheet.pdf

- Full Matlab Guide:

  http://www.mathworks.com/help/pdf_doc/matlab/getstart.pdf

- C++ : Open Source Computer Vision Library (OpenCV)

  http://opencv.org/

  http://docs.opencv.org/doc/tutorials/tutorials.html