

Laboratory 10 “Computational Vision”

Laboratory 10: Video segmentation

In this laboratory, we will practice how we can segment videos and extract background and foreground from it. The main issues are:

- 1) Shots detection
- 2) Descriptors based on image colour and texture.
- 3) Background extraction for videos from static cameras.

To complete the laboratory it is necessary to know and apply the basic concepts on video segmentation. For this, refer to the material of the theoretical slides.

10.1 "Background subtraction" Given the video 'Barcelona.mp4', which contains images acquired with static cameras, detect the shots and remove all movement-related artifacts by extracting the background of the shots.

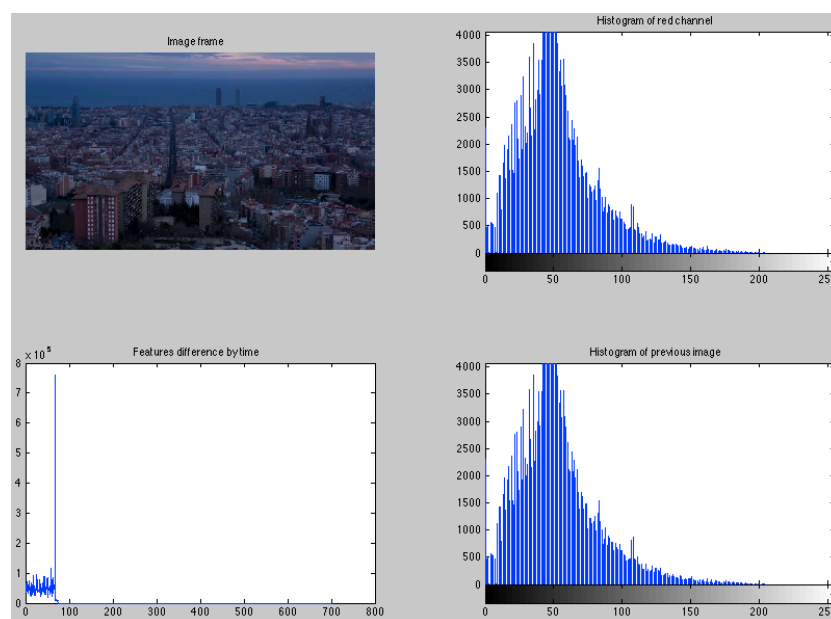


Fig.1 The current image frame (top, left), histogram of red channel (top, right), histogram of red channel of previous frame (bottom, right). Differences between frames features (bottom, left).

Note: One of the applications of this option can be related to the "tourists delete button" invented for the cameras. For example, Adobe uses the "Monument mode" that automatically suppresses people passing in front of the camera.

To perform the algorithm you need to:

- A) Find where one scene ends and another begins (i.e. the video shots) (see Fig.1). What measure of the images can you use to distinguish the scenes?
- B) Apply a background subtraction algorithm (e.g. using the median filter to detect the static scene forming the background).

Display, for each segment delimited by two shots of the video, the extracted static image and sequentially the images (as a movie) from which it was obtained (Figure 2).

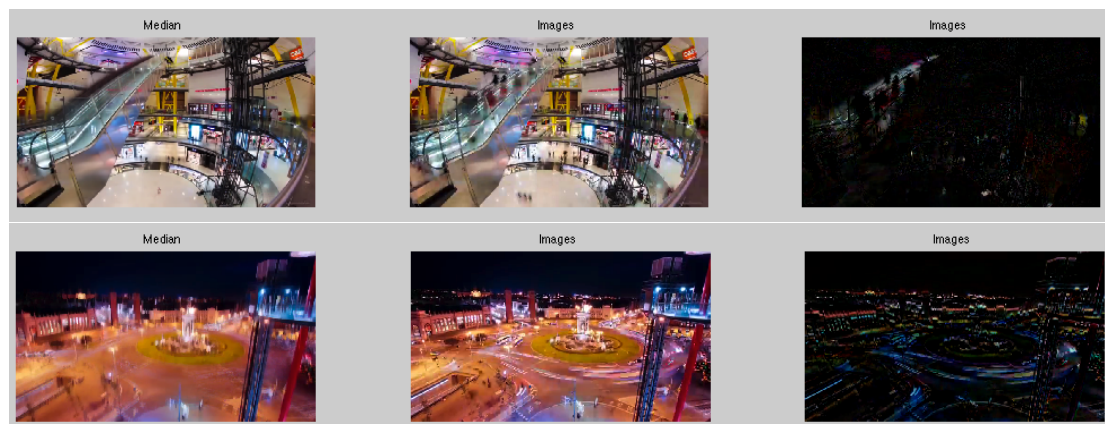


Figure 2. Background extracted (left), a frame from the shot (center) and segmented objects (right)

Discuss your implementation in detail. What were the optimal visual features you found? Comment advantages and limitations of your proposal. Comment on what the static images represent. Do you see other applications of this algorithm?

[OPTIONAL] Apply the algorithm on another video you can find in internet.

10.2 Video segmentation for event extraction Let us imagine that you have been given a wearable camera, you went abroad, you were wearing the camera all days so you have around 2000 images per day, and you want to separate, measure and store the different events (travelling by plane, taxi, sightseeing, eating, shopping, walking, etc.). Develop an algorithm to separate the outdoor from indoor events of your day (Fig. 3).

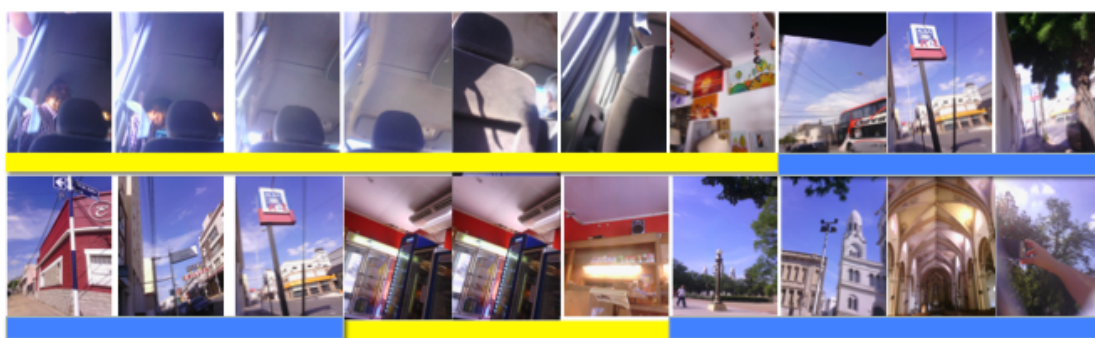


Fig.3 Indoor (yellow) and outdoor (blue) image segment

Explore if the texture descriptors from Laboratory 9 improve the results. The filters to generate the texture descriptors can be downloaded from <http://www.robots.ox.ac.uk/~vgg/research/texclass/code/makeLMfilters.m>.

What feature descriptors did you find optimal in order to separate the event shots?

Visualize the function of descriptors differences between frames and define a criterion to decide a threshold for the shot cut (Fig. 4).

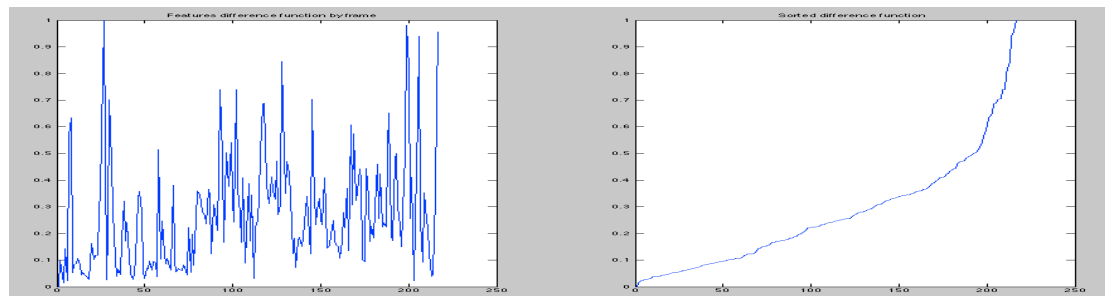


Fig. 4. Features difference function by frame (left) and sorted difference values (right).

Comment the advantages and disadvantages and possible ways to improve the method.

Note: You can reduce the images (imresize) in order to faster the computation time.

Delivery of the laboratory Deadline: 10 January at 23:50h.

This part is to be delivered together with the next laboratory. It should be stored in a compressed file named "names_surnames_P9-10-11.zip" containing: - .m files with the created functions and a pdf file explaining the implementation and the performance of the algorithms and the answers raised in the exercises.