Video Analytics in Basketball

Computer Vision Course - Second Assignment

Riccardo Scilla - 213723
University of Trento
riccardo.scilla@studenti.unitn.it

1 INTRODUCTION

The aim of this work is to extract some basic analytics starting from a video sequence of a basketball play, using Computer Vision techniques. The analytics extracted are:

- Run a people detector on the whole video.
  Write on the output video the number of people detected at each frame in the scene.
  Qualitatively evaluate the accuracy of the detector when reporting the results.
- Choose one person in the scene (either player or referee) and try to track him.
  Plot the trajectory on the output video.
  Qualitatively evaluate the obtained results.
- Detect how many times the ball possession change.
  Display the result on the output video.

The Workflow of the video analysis system is shown in Figure 1.

2 VIDEO PREPROCESSING

2.1 Initial Scripts

A script (`background.py`) is executed to extract the background of the scene: this is possible because the camera is static and there are no big illumination changes. A sample of frames are taken and the median is computed. The result is the court field image without players, here called *full_background*.

A script (`mask.py`) is executed to extract a mask of the court field. This is applied to each frame in the processing to filter out all the moving objects outside the court field. The script takes the *full_background* an allow the user to select a set of edges of the polygon that will be used as mask.

The scene has also an advertisement display that could be removed by the *full_background*: in order to remove it, another mask with the display shape is created.

2.2 Background Subtraction

A classical background subtraction is executed to find moving objects: frame is converted to gray and a difference is applied with the previous background; then the background is updated according to the rule with alpha=0.3. Thresholding, Closing and Dilation are performed to remove noisy points and enhance moving objects. Then the contours are extracted as bounding boxes and they filtered so that smaller bboxes are removed, and only the good ones are returned for further steps.

A similar procedure is performed also using the *full_background*: in this case it is considered as fixed background for all the frames.

The advantage of using both adaptive and fixed background subtraction is that with the adaptive we can better detect players that are in front of the advertising display. On the other hand with the *full_background*, we can better detect players that are static for many consecutive frames.

The two functions return a set of good Rects of moving objects. In many cases these sets overlap for some objects: to solve this, a *non_max_suppression* operation is performed to merge overlapping Rects.

3 DETECTION

All the Rects found are enlarged by 10 pixels to better capture the players inside them. Then HOG People Detection with pretrained SVM is applied: this process is heavy to be computed and does not perform well when players fall, jump, or crunch to get the ball, and will not be detected by the HOG detector.

To solve this, we rely on the fact that only the players are detected with the background subtraction method, so that they can be considered in the final count of the player detected. The HOG is still used (`HOG.py`) in the case the Rects are bigger than a threshold: in these cases we will have that multiple players overlap and cannot be counted as one. With HOG there is the possibility that players that are too close together or partially obstructed by other players, will be detected singularly. Also in HOG, *non_max_suppression* operation is performed to merge overlapping Rects.

At the end of the Detection step, we will have a frame displaying all the players detected by background subtraction (green) and by HOG (blue), together with the count of the players found.

4 TRACKING

Tracking of a player or referee is performed using CSRT tracker: this is the implementation based on [1] provided by OpenCV.
It uses only 2 standard features: HOGs and Colornames.

To start tracking, the user is asked to select the ROI: this step is fundamental to get good results with tracking. For this reason two bounding boxes regarding a player and a referee are provided. The CSRT tracker is initiated in the first frame using the bbox and it is updated at each frame. The trajectory of the last 100 frame is plotted on the output frame.

5 BALL POSSESSION CHANGES

The ball possession is computed starting from the player detection: by computing the mean of all the centroids of the Rects, if the mean is bigger than the half of court width, then the possession is of the right court, otherwise it is of the left court.

To stabilize this measurement when the mean is close to the half of court width, a counter is used: 30 consecutive frames with the same possession result are required in order to increment the Ball possession changes counter.

6 USAGE

Arg	Values (**default**)	Description
-ts	**player**,referee,None	Select the subject to track
-bgs	True,**False**	Show bg subtraction
-bgsf	True,**False**	Show bg subtraction full
-fhog	True,**False**	Use only HOG Detector

REFERENCES

[1] Alan Lukezic, Tomas Vojir, Luka Cehovin Zajc, Jiri Matas, and Matej Kristan. Discriminative correlation filter with channel and spatial reliability. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.

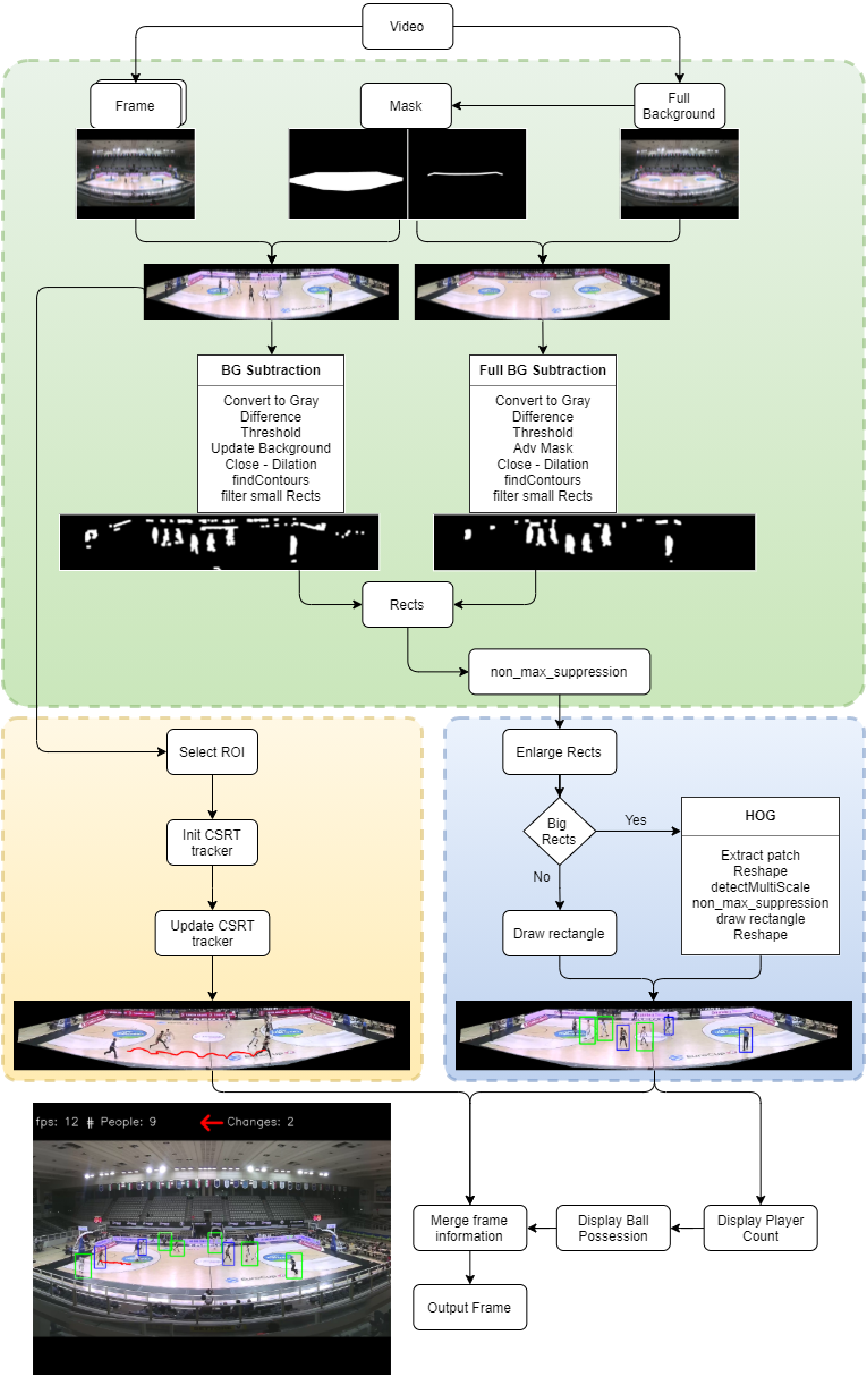

Figure 1: Workflow of the proposed video analysis system