Tracking of Tennis Ball in Tennis Serving Videos Using Particle Filtering and Segmentation

Yikai Lan: 4474783, Y.Lan@student.tudelft.nl
Xuehan Jiang: 4578791, X.Jiang@student.tudelft.nl

July 2017

1 Problem description

1.1 Introduction

Tracking aims to following objects or objects configuration in a video sequence, which is a significant element of many computer vision systems. Tons of applications such as visual surveillance, video compression, and visual content analysis. Among all these applications, we focus on building a system to track the ball in some sports videos. which can be applied in computer-assisted sports study and automatic annotation of sports video.

1.2 Challenge

Tracking has been an well-studied computer vision topic with various kinds of algorithms during the past decade. However, It is still non-trivial to realize the tracking of the balls because of the following challenges: 1) balls, as Small objectives, usually have less informative features to be used for detecting. 2) the speed of the balls are sometimes so high that they are blurred into backgrounds. 3) In some sports like tennis, the tracking of the ball is stuck with occlusion, out-of-view, shape distortion, false detection and dramatic change of motion direction.

2 Methodology

Inspired by [1][2], we implemented a ball detecting and tracking method based on particle filters.

2.1 Segmentation based detection

Various segmentation-based detection algorithms have been proposed, such as background subtraction, motion estimation and color estimation. In our case, color estimation was chose due to the fact that balls usually have a simple but distinct color distribution compared with background. Meanshift algorithm based on color estimation was implemented to segment the balls from the background.

2.2 Particle filter based tracking

Particle filtering methodology uses a genetic type mutation-selection sampling approach, with a set of particles (also called individuals, or samples) to represent the posterior distribution of some stochastic process given some noisy and/or partial observations. The efficiency and accuracy of the particle filter depends on two key factors: the motion model and the appearance model. Motion model decides how a movement of particles are generated by a proposal distribution. Appearance model decide how these particles are weighted to estimate posterior state distribution. Different from the appearance model of the particle filter we implemented in assignment 7, we developed new appearance model based on the likelihood map of each image.

2.3 Combining mask and particle filter

The idea of our project is to combine mask and particle filter. The mask can give a rather satisfied result when the background color is distinguish with the color of the tennis ball. But it may produce false positive, which can be reduced by applying particle filtering if we assign proper variance to limit the search radius.

And when mask is not working, either the color of the background is similar to tennis ball color or that the tennis ball in occluded, the particle filtering method can still give a predicted result. And when the tennis ball shows up again in the frame, we can use mask to find a more accurate position.

3 Implementation

We implemented the majority of the codes by ourselves using openCV on Python3. Steps of implementation are listed below.

It should be noted that the initial spot and the initial color of the tennis ball is extracted manually in the first chosen frame from the video.

3.1 Color Segmentation

Since tennis balls are generally green, a mask can be used to find the circles in the image.

This can be achieved by first extracting the color of the center of tennis ball and generate a HSV color range for the mask. Then using openCV package, the minimum enclosing circles could be found within the image.

As shown in Figure 1, two tennis balls can be detected accurately without confusion with the green court and yellow shirt.

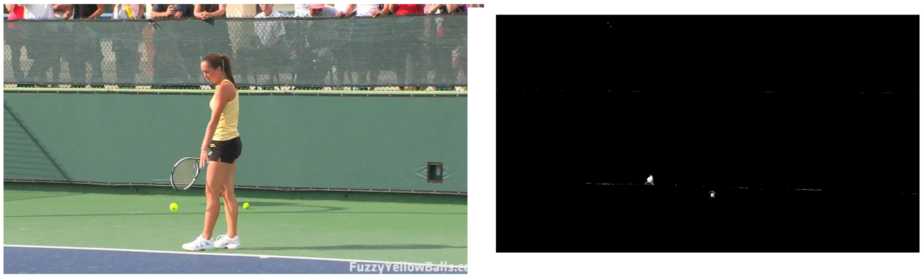

Figure 1: first frame of a video and the likelihood map of it, where the white contours means the color matching with the tennis ball.

However, if the color distribution of background are quite similar to the ball or the color of the ball changes during the video, the segmentation performance will drop with many false positives. Considering this, the moving distance and the change of radius of tennis ball in two frames is also taken into consideration. For those circles who fall into the thresholds of distance and radius requirement, the circle with the largest contour area is chosen as the **most probable position**, X_{most}, of the tennis ball in current frame.

$$P_{particle\ segmentation} == \begin{cases} 1; \text{if } distance(x, X_{most}) \leq 1.2 \cdot \text{initial radius} \\ 0.6; \text{if } 1 \cdot r_{\text{initial}} < distance(x, X_{most}) \leq 1.2 \cdot r_{\text{initial}} \\ 0.4; \text{if } 1.2 \cdot r_{\text{initial}} < distance(x, X_{most}) \leq 1.5 \cdot r_{\text{initial}} \\ \exp(\text{distance}/4); \text{if } \textbf{otherwise} \end{cases} \tag{1}$$

3.2 Ball detection

First, a rectangular area from the first frame of videos was cropped as the object. Then, we calculated the color bins histogram of the object. Each color channel was split uniformly into 8 bins. After generating the objective color histogram, we used it for computing the likelihood ratio of each pixel. By doing this, we can build a color likelihood map for every frame in order to detect the position of the objective ball.

$$P_{particle\ color} = arg(P_{pixel\ color}) \tag{2}$$

3.3 Particle filtering

By initialize the particles at the same time for extracting color, the likelihood of the particles are calculated based on appearance model and how close it is to the most probable position gained from the color segmentation described in 3.1 and then particles are sampled by using residue resampling[3].

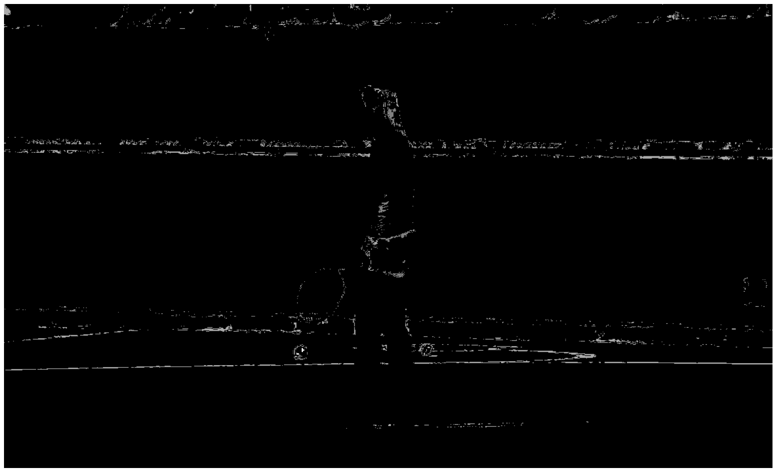

Figure 2: The likelihood of first frame of a video based on color distribution, where more white pixel means higher likelihood.

$$P_{particle} = P_{particle\ color} \cdot P_{particle\ segmentation} \tag{3}$$

3.3.1 Appearance model

Using the method as mentioned in 3.1, likelihood map, which contain the weight of every pixel in one image, was established for every frame of the video. We used these maps as the appearance model and updated the weight of pixels by calculating the average of the probability within the range of the particles.

3.4 Motion model

As for the motion model, the Gaussian distribution is assumed. New particles are gained using residue resampling based on the probability shown in Equation 3. The motion model is also force to vary around the possible position by assigning a portion of the particle equal to X_{most} before resampling step.

The variance of the Gaussian distribution is determined by distance change of X_most in previous frames in order to gain a proper prediction of the position of the tennis ball.

4 Experiment result

In the section, the tracking result will show tracking result of two kind of typical video. In all images, the green line show the contour found by mask using green color range. The purple circles indicate the particles in current frame. While the yellow circle shows the most probable position X_{most}. Red curve line indicate the moving path of X_{most}. The bold purple circle indicate the top high probability particle position.

4.1 Tracking when mask is working

As shown in 1, mask is working the video in 3 very properly. Where we can see there are two tennis ball in the video with one of them is our tracking object, and another false positive is the shirt of one audience in the back.

The path of X_{most} show that we can track the tennis ball very properly even using only the mask image, compared rather good with the particles. The particle with most highest probability sometimes miss the tennis ball, this may due to the fact the variance may be too large sometimes. However, there is occlusion occurs in **(e)**, where the mask could only find the tennis ball at the right corner and all particle filters flees to that position. And when tennis ball show again in **(f)**, almost half of the particles stay around the false positive on the top, while the particle with highest probability stay close to the real tennis ball.

In this kind of situation, it may gain a better result if we assign more weight to the mask and use the weighted position combining the X_{most} and mean position of the particles.

4.2 Tracking when mask is not working

As shown below in 4, this video will gain a lot of false positive in background if we assign a wide range of color space since the grass has a similar color to the tennis ball, which may result in wrong tracking path for the mask. Thus only particle filtering could be used to predict the position when the mask is not working, where $P_{particle\ segmantation}$ is all 1 for all particles.

As show in (a) to (c) from 4, the mask can still track the tennis ball. But in (d), the tennis ball is under occlusion and bold purple circle predict a rather nice position for the tennis ball.

In this case, the particle filtering is stays close to where the tennis ball go missing, and find a position where the likelihood is highest. And when the mask can determined where the tennis ball is, it will obtained a more accurate prediction.

5 Conclusion

In this project, a modified particle filtering using residual resampling combined with the color segmentation using color mask is used to track the tennis ball in the video. As shown in the result, even if the tennis is of a few pixels, the result is still satisfying.

The algorithms can be improved by considering trajectory and velocity based on history frames. And more dataset can be used for testing besides tennis ball

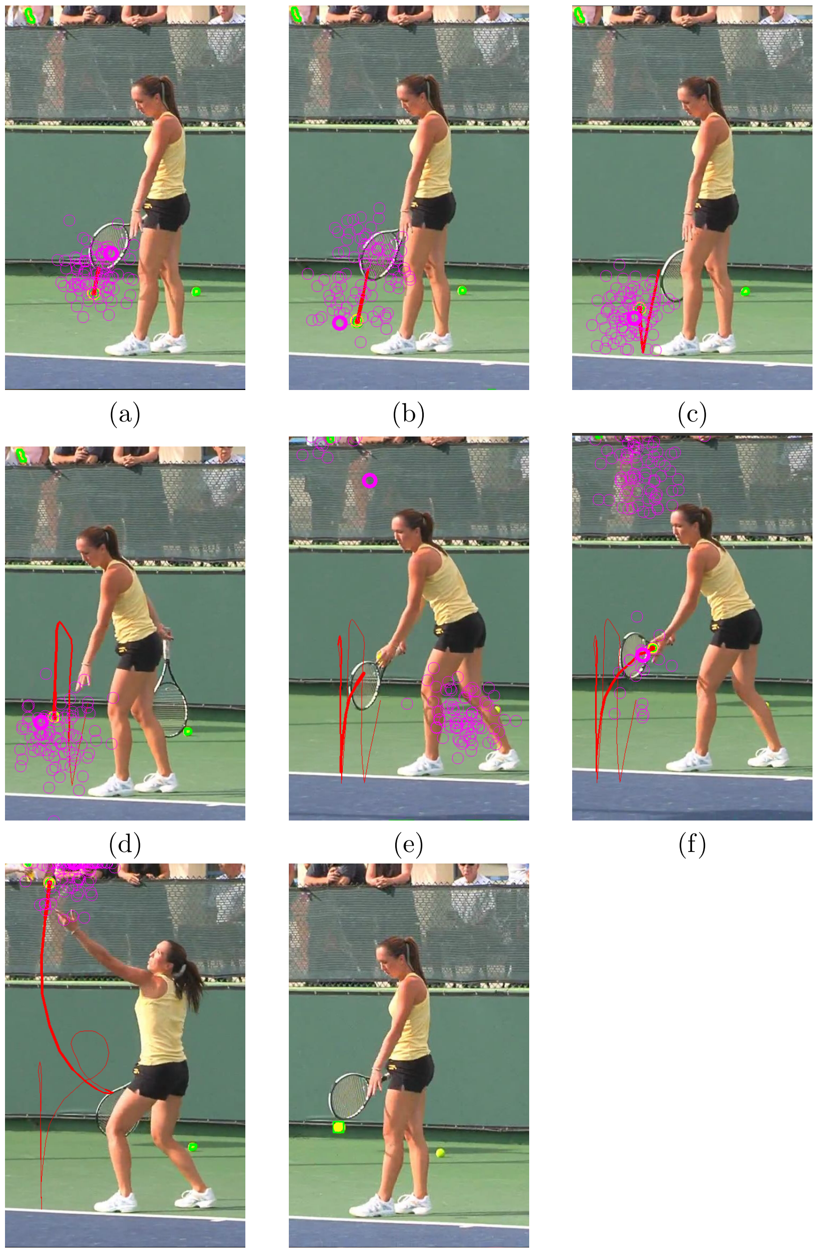

Figure 3: The tracking result of one video when mask is working, (a) to (g) showing the tracking result of tennis ball during serving, (h) with the last image showing the labeling of the tennis ball manually.

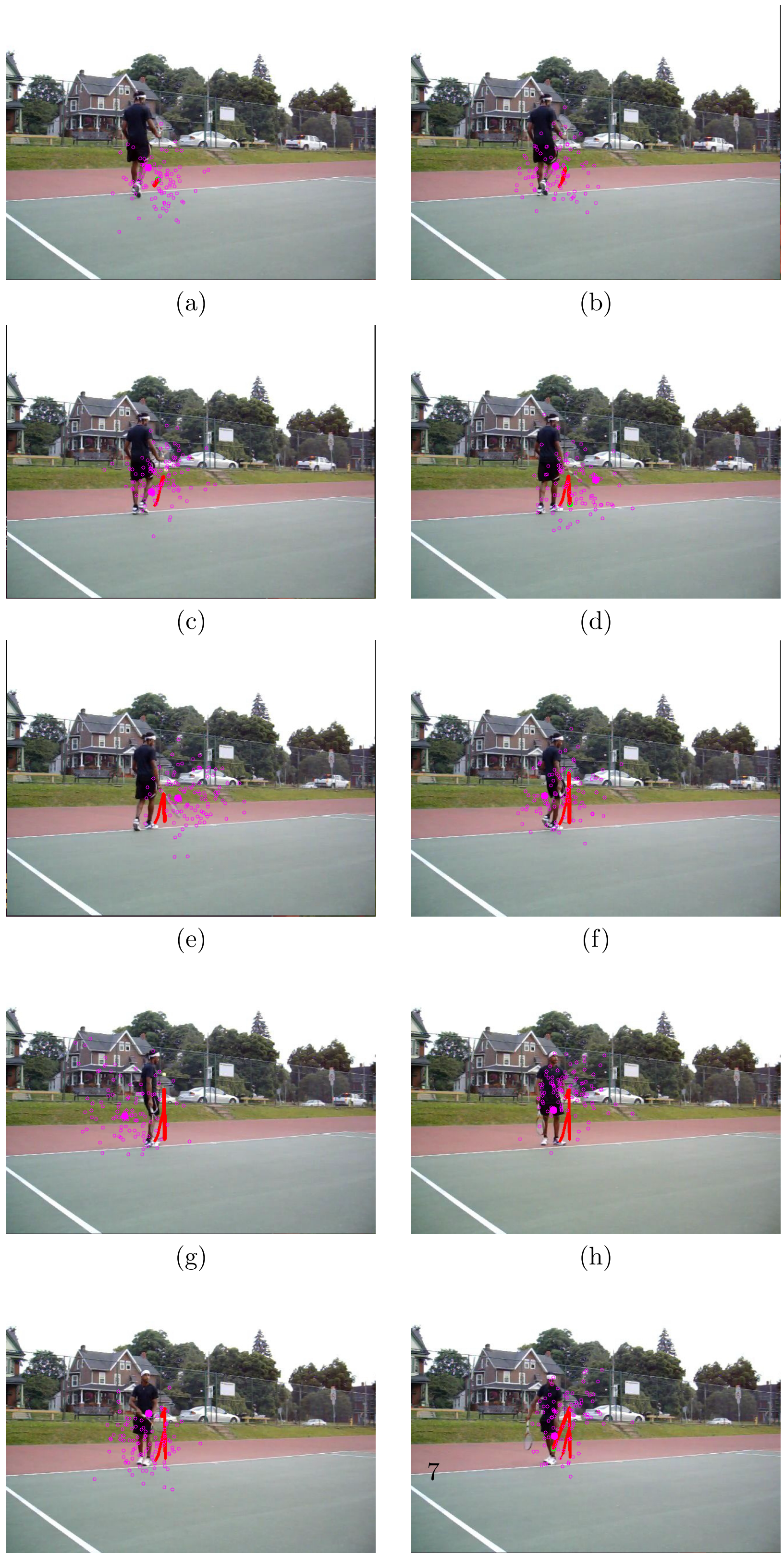

Figure 4: The tracking result of one video where mask is not working, (a) to (j) showing the tracking result of tennis ball during serving.

serving. And more quantitative experiment result can obtained if using labeled video.

References

[1] Y. Huang and J. Llach. Tracking the small object through clutter with adaptive particle filter. In *Audio, Language and Image Processing, 2008. ICALIP 2008. International Conference on*, pages 357–362. IEEE, 2008.

[2] Y. Huang, J. Llach, and C. Zhang. A method of small object detection and tracking based on particle filters. In *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*, pages 1–4. IEEE, 2008.

[3] P. Kozierski, M. Lis, and J. Ziktkiewicz. Resampling in particle filtering-comparison. 2013.