Object Tracking Based on CNN

Jinho Lee, Shouta Ide, Seiichi Uchida
Kyushu University

1 Introduction

Object tracking is one of the significant researches studied in computer vision and pattern recognition fields. In spite of many past researches, traditional object tracking methods still have constraints. For example, the traditional methods cannot deal with appearance changes of the target object. In addition, it is difficult for them to deal with occlusion.

The purpose of this research is to track the object more accurately, compared to traditional methods. In this research, a Convolutional Neural Network (CNN)[1] is used to evaluate the likelihood of the target object at every pixel in each video frame and Dynamic Programming (DP)[2] is used to obtaining the globally optimal tracking path.

2 Tracking Method

2.1 Object detection with CNN

A likelihood of the target is evaluated by using a CNN which is trained to classify a fixed-size block image into *target* or *non-target*. CNN can deal with various changes in target appearance by its high generalization ability acquired through training with a large dataset. It, therefore, can provide stable likelihood values to large appearance variations. Our CNN was made of 3 convolutional layers of 32, 32, and 64 nodes respectively with 5×5 kernels. We use the Softmax for output layer. The likelihood value of a pixel is obtained by inputting a $x \times y$ block centered at the pixel into the CNN. The likelihood map is given by sliding the block over the frame.

In the following experiment, we set automobile as the tracking target and thus the Cifar10 dataset is used for training the CNN, which consists of 6000 32×32 color images for each of 10 object classes, including automobile. The dataset is separated into two classes, i.e., automobile and the others, and CNN is trained for the two classes.

2.2 DP for the globally optimal tracking path

We obtain the globally optimal tracking path through all frames by using DP. DP is one of the most fundamental optimization techniques and has been used for object tracking. DP is a non-greedy algorithm and can guarantee the global optimality of the path over all frames. Because of this nature, DP-based tracking is robust to occlusion, which degrades the tracking performance of greedy algorithms.

In this research, the tracking path is evaluated by the likelihood values given by the CNN. DP searches for the highest likelihood values efficiently under a certain slope constraint for smoothing the tracking path.

3 Experimental results

Figure 1 shows the frames(440×200) and likelihood maps made by the matrices of the likelihood values. This video consists of 35 frames. It takes 1400 sec. for training and 75 sec. for DP tracking respectively. The white color of the likelihood maps means high value, in contrast with the black color. The red square indicates the tracked block image(80×60).

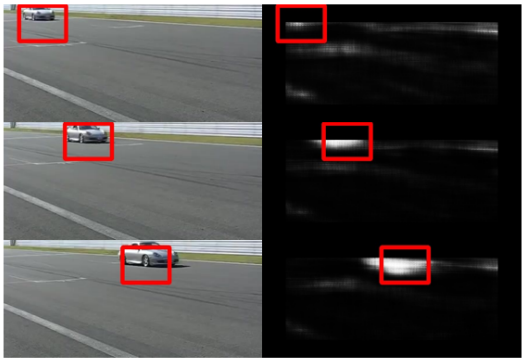

Figure 1: Original frames, tracking results, and likelihood maps. (*www.youtube.com/watch?v=4QqvcxV35pE*)

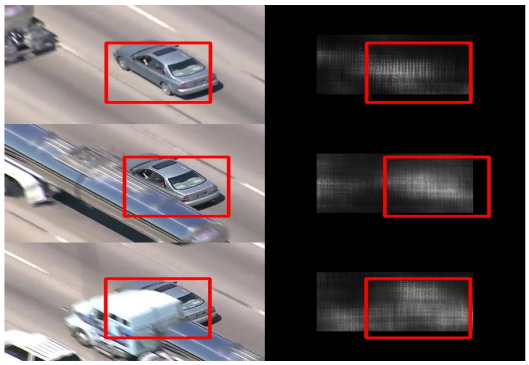

Figure 2: Occlusion frames, tracking results, and likelihood maps. (*www.youtube.com/watch?v=YoxkXLx1Hp8*)

Figure 2 shows the frames(500×300) included the occlusion and likelihood maps obtained by the same method. In the figure 2, the video consists of 14 frames and the tracked block image size is 200×150. It takes 542 sec. for training and 90 sec. for DP tracking.

As these results, we confirmed that the automobile is tracked appropriately in figure 1 and 2, despite occlusion and various changes.

4 Conclusion

We confirmed that our proposed tracking method can deal with occlusion and various changes. In the future, we will attempt to increase the accuracy of tracking and do multi-target tracking by setting the dataset of the CNN more efficiently and modifying the CNN and DP.

References

[1] Jialue Fan, Wei Xu, Ying Wu, and Yihong Gong "Human Tracking Using Convolutional Neural Networks," IEEE Transactions on Neural Networks, pp.1610–1623, 2010.

[2] Seiichi Uchida, Ikko Fujimura, Hiroki Kawano, and Yaokai Feng, "Analytical Dynamic Programming Tracker," ACCV 2010.