Distribution Learning for Moving Object Segmentation

Chenqiu Zhao(zhao.chenqiu@ualberta.ca)

Abstract

Is there any better way to classify two distributions, **can we make the deep learning network learn and classify distributions automatically and accurately,** which should have wide ranges of applications in computer vision, such as moving objects segmentation. In our previous work [1], we already found that the convolutional neural network can be forced to focus on statistical distribution by randomly permutating the input of network. Such property is utilized to segment moving objects in a video, since the distribution generated by moving objects has significant difference with the ones of background scenes. In this proposal, we want drag deeper into this field to figure out what exactly learned by the network, why the network acquire the ability to classify the statistical distribution theoretically. We will focus on improving the performance as well the generality of distribution learning. In particular, the theoretical procedure will be a key task in this proposal, since it is very important for academia. Also, once we figure out the reason of network's ability to classify distributions, it is also helpful to devise a better network for distribution learning.

Introduction

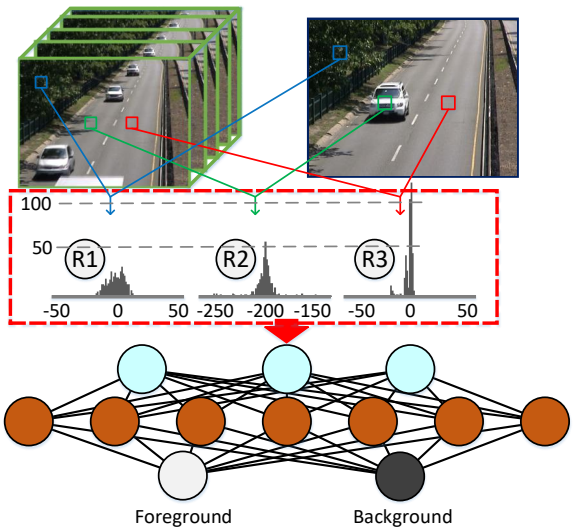

Figure 1: Deep distribution Learning.

Motion is a powerful cue for image and scene segmentation in human visual system. It has wide range of applications such as video monitoring, optical motion capturing and multimedia applications. The human ability to detect, segment, analyze and understand motion is nearly instantaneous, and work well in diversely complex scenes. While there has been much recent progress related to motion, it still appears we are far from human capabilities. Therefore, as an essential and fundamental research topic related to motion of computer vision, background subtraction has been attracting increasing attention in the recent years.

The main aim of background subtraction is to separate object with motion from the background in a video, in which each pixel is classified into two classes, namely, foreground class which represents the moving objects, and background class which corresponds to the background scenes. Traditionally, background is subtracted by the analysis of pixels' observations, and plenty of previous work address the problem by approximating the distribution of observations via statistic model. However, due to the complexity and the

diversity of natural scenes, the distribution generated by pixels is so hard to be approximated by artificial model perfectly, as shown in the Fig. 1.

Essentially, background subtraction is the classification of pixels in a sequence of image frames, typically under the assumption of a static camera, wherein each pixel in a particular frame is classified as foreground or background by comparing its current measurement with historical observations. When dealing with diverse and complex scenes, it is a challening problem as the pixel measurements take the form of complex distributions for the pixels in both foreground and background. For example, as shown in the R1, R2 and R3 of Fig. 1, the distributions of the pixel values in the pixels belonging to a dynamical background, moving objects and a static background are completely different, and manually-tailored models have limited ability to cope with these contrasting distributions. In this proposal, we focus on learning the distribution automatically instead of devising the sophisticated model, and deep learning network is naturally be selected to learn the distribution, due to it is excellent learning ability shown recently years.

Moreover, the distribution itself has a wide range of applications. Therefore, the distribution learning can not only be used in background subtraction, but also in some other research problems, such as tracking, object detection, or even some challenging problem outside of the computer vision, in which tons of research work can be done. Hence, like the hierarchical features learned by deep learning network, which work better than all these artificial features, **is there any hierarchical representation of distribution implied which is not be discovered yet? Will this high level distribution work better than the existed techniques, e.g. histogram, density estimation and so on, related to distribution?** Solving this problem would be extremely interesting, which is the **main purpose of "Deep Distribution Learning".**

References

[1] C. Zhao, T. Cham, X. Ren, J. Cai, and H. Zhu. Background subtraction based on deep pixel distribution learning. In *2018 IEEE International Conference on Multimedia and Expo (ICME)*, pages 1–6, July 2018.