Supplementary Material

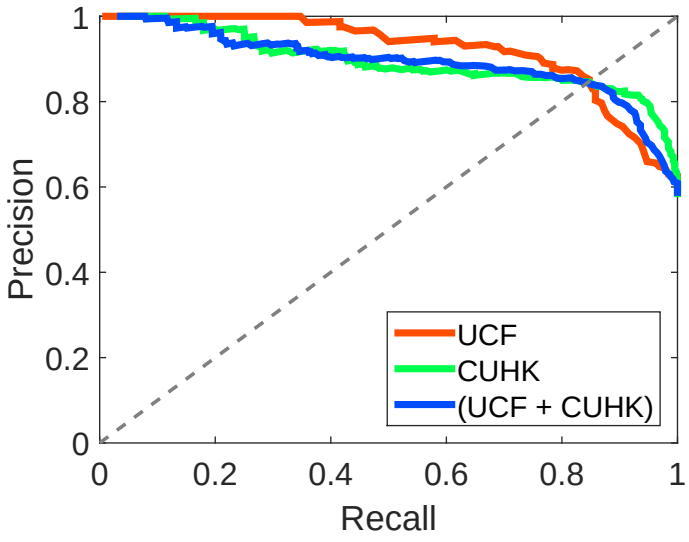

Fig. 1. Precision-Recall curves. Red line: the result on UCF dataset. Green line: the result on CUHK dataset. Blue line: the result on the combination of UCF and CUHK datasets.

The threshold for specifying significance determines whether a cluster should be considered or ignored. When the value of this parameter is large, some crowd motions will be undetected. Otherwise, there will be some false detections. In order to evaluate how much this parameter affects the results, we conducted an experiment by varying it from 100 to 20000 with step size of 100, and Precision-Recall curve and Area Under Curve (AUC) are used as the evaluation metrics. The detection results on the UCF and CUHK crowd datasets are shown in Figure 1 in this manuscript. The performances on different datasets are comparable with each other (the values of AUC on the UCF, CUHK and UCF + CUHK datasets are 0.924, 0.896 and 0.893 respectively). Due to the large variations of image size and motion pattern, the same parameter setting for the threshold would not produce satisfying detection results (i.e., high recall with high precision). So, in our previous experiments, the threshold for specifying significance is manually set to a different value for each image sequence. It should be noted that this manually setup is the same for the Baseline III, IV and our methods, and thereby the comparisons between the proposed CDT descriptors and related crowd features are still convincing. A probable way to automatically set the parameter is that we could use the proportion (e.g. 80%) of moving particles as the threshold for identifying significant clusters, which would produce favorable detection results despite the large differences of image size and motion pattern. This issue will be deeply explored in the future.

As the supplementary material, this parameter setting and discussions along with the source code can be found at the website: http://github.com/shuangseu/CDT_crowd_descriptor.