

001  
002  
003  
004  
005  
006  
007  
008  
009  
010  
011  
012  
013  
014  
015  
016  
017  
018  
019  
020  
021  
022  
023  
024  
025  
026  
027  
028  
029  
030  
031  
032  
033  
034  
035  
036  
037  
038  
039  
040  
041  
042  
043  
044  
045  
046  
047  
048  
049  
050  
051  
052  
053  
054  
055  
056  
057  
058  
059  
060  
061  
062

---

Egocentric video and wearable computing have become increasingly prevalent in the past decade, resulting in a huge explosion in the amount of available video content. In this paper, we present a novel approach for egocentric activity recognition using the UC Irvine ADL (Activities of Daily Living) dataset [1]. Existing work in activity recognition uses predefined binning schemes, which may fail to capture important spatio-temporal relationships between features. We propose to partition video clips into sets of 3-dimensional cuboids based on many different multi-level randomized partitioning schemes, then concatenate object histograms over multiple levels to form feature vectors which we then use to train a pool of weak SVM classifiers. Finally, we use a boosting algorithm to learn which partitioning schemes are most discriminative and form a final strong classifier with accuracy that improves upon the current state of the art. Our main novel contribution is a method for creating biased partition schemes based on observed distributions of active object locations across each spatial and temporal dimension of the video clips. We found that partitions which cut through spatio-temporal regions that tend to contain active objects are often more discriminative than unbiased partitions and partitions that cut around such active object regions.

[1] H. Pirsiavash and D. Ramanan. Detecting activities of daily living in first-person camera views. *Computer Vision and Pattern Recognition, 2012.*, pages 2847–2854.