Intro for Phase-2

Automatic categorization of human actions in video sequences is very interesting for a variety of applications: detecting activities in video surveillance, indexing video sequences, content based browsing etc. We would try to categorize six classes of human actions, namely running,jogging,walking,boxing,handclapping,handwaving as the time taken to perform the computations for a video is significant.

Image structures in video are not restricted to constant velocity and/or constant appearance over time. On the contrary, many interesting events in video are characterized by strong variations of the data in both the spatial and the temporal dimensions. As example, consider scenes with a person

entering a room, applauding hand gestures, a car crash or a water splash.

More generally, points with non-constant motion correspond to accelerating local image structures that might correspond to the accelerating objects in the world. Hence, such points might contain important information about the forces that act in the environment and change its structure. In the spatial domain, points with a significant local variation of image intensities have been extensively investigated. Such image points are frequently denoted as “interest points” and are attractive due to their high information contents. we detect interest points in the spatiotemporal domain and illustrate how the resulting spacetime features often correspond to interesting events in video data. To detect spatio-temporal interest points, we used the Harris interest point operators.

Intro for Phase -1

Detecting humans in images is a challenging task owing

to their variable appearance and the wide range of poses that

they can adopt. The first need is a robust feature set that

allows the human form to be discriminated cleanly, even in

cluttered backgrounds under difficult illumination. We study

the issue of feature sets for human detection, showing that locally

normalized Histogram of Oriented Gradient (HOG) descriptors. For simplicity

and speed, we use linear SVM as a baseline classifier

throughout the project.

SVM:

Support vector machines (SVM) are a group of supervised learning methods which constructs a hyperplane or set of hyperplanes in a high- or infinite-dimensional space, which can be used for classification, regression, or other tasks. Intuitively, a good separation is achieved by the hyperplane that has the largest distance to the nearest training-data point of any class (so-called functional margin), since in general the larger the margin the lower the generalization error of the classifier.