**PROJECT DELIVERABLE 3**

**Semantic video retrieval**

Labelling large datasets of videos with relevant tags can be an extremely time consuming if done manually. One way to get around this is to use the technique of semantic video retrieval. It’s divided into three parts of 1) Semantic Concepts modelling 2) Automatic Semantic Labelling

3) Retrieval using Semantic Keywords. In the step of semantic concepts modelling a training dataset was annotated manually by us, for this purpose shots of the videos are generated using the Video-in-Video-Out algorithm’s part where shots were detected using automatic thresholding. Now, get a new concept from the list of annotated concepts, than get all the shots in the training set related to this concept and extract their DCT feature vectors. To get a DCT feature vector, divide the image in 16x16 blocks and apply dct2() function on them, than store these in an n x d array. Where n is the total number of blocks and d is the dimensionality. Apply Gaussian Mixture Model on the feature space, since, each image block is 16x16, this will result in 256 coefficients. Thus the feature vector will be of length [1- 256). GMM is on this high dimensional space.

GMM is only run once all features from few videos having multiple shots containing the concept you are modelling are obtained. Rare concepts have been discarded.



We go to each video in the dataset and extract each shot from the video. Then we compute DCT features for each shot (dividing the image into blocks of 16x16). For each concept in the database estimate the probability of this shot belonging to the concept. Assuming there were K concepts in the annotated set, you will end up having a vector of K probability values. Pick the top 4 concepts based on the obtained probability scores and assign them as tags to the video along with the obtained probability.