SemArchy: Video Segment Retrieval Engine

Aarti Prabhu, Varun D N, Divya V, Prof. Ananthanarayana V S

Introduction

Video is composed of heterogeneous contents. Conventional video search engines retrieve complete videos relevant to the search queries using the metadata attached to videos. However, relevant content might be present only in a subsection of the video and the content of the videos are not understood through the metadata. This work is concerned with extracting relevant video segments by analysing the contents of the videos.

Methodology

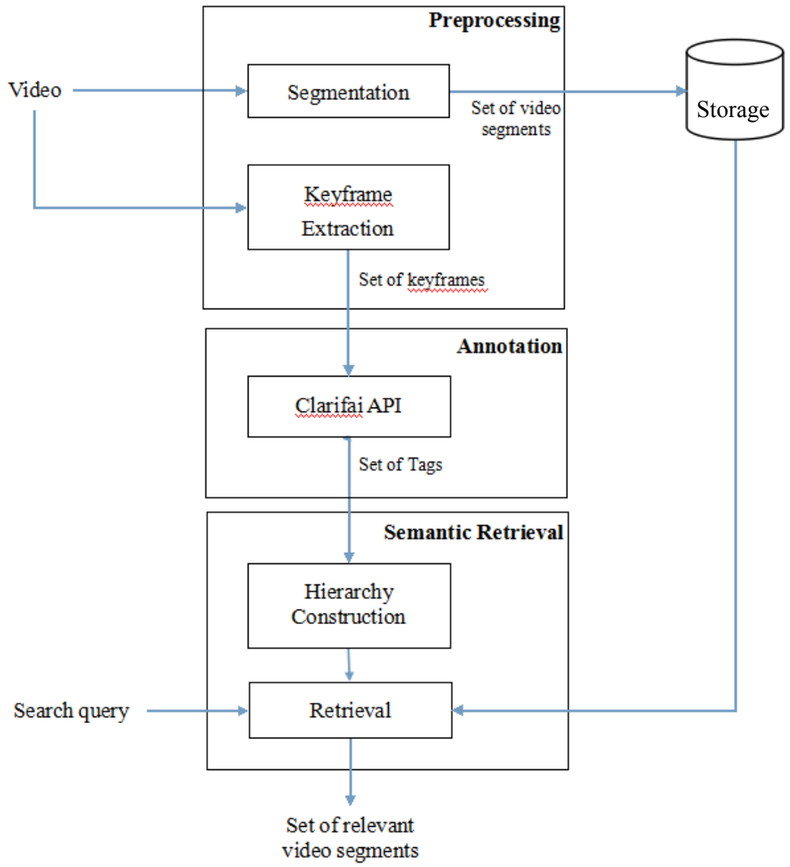

Figure 1: Flow Diagram

Objectives

- Key frames extraction and video segmentation
- Annotation of key frames using semantic properties
- Building a hierarchy over the tags annotated
- Search though the hierarchy for a user query

Unique Work

- Video segments are retrieved instead of whole videos
- Relationship between keywords that occur together is exploited

Results

702 video segments are extracted from a set of test videos. Number of comparisons through *linear search* is always 702.

Query	Number of comparisons using hierarchy
Landscape	87
Portrait	93
Nature, Sky	156
Landscape, Sky, Water	323

Table 1: Search Efficiency

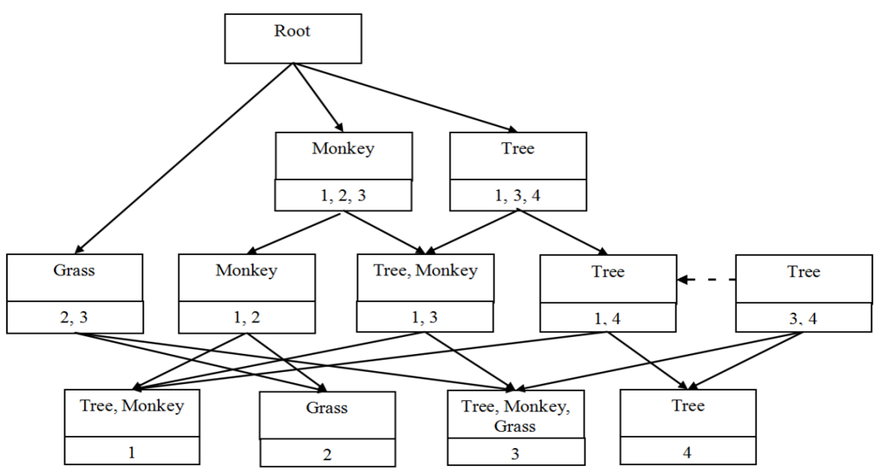

Figure 2: Sample Hierarchy

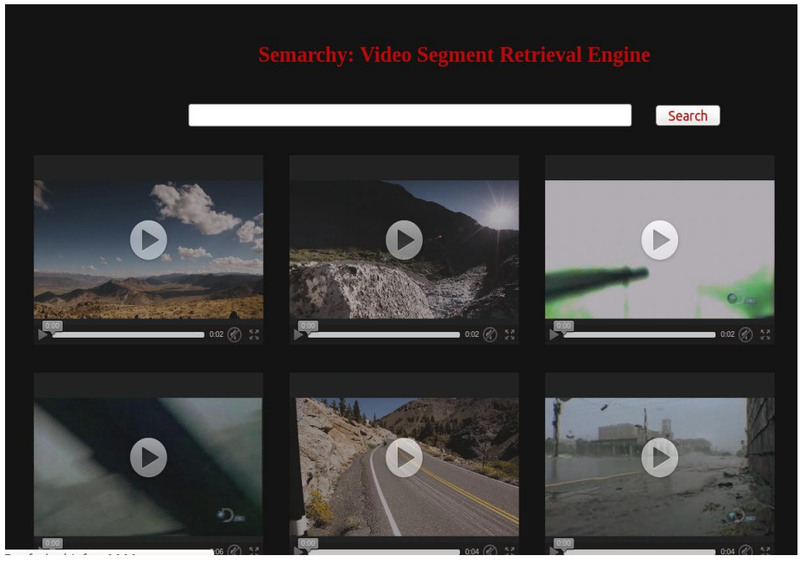

Figure 3: Output screenshot for the query 'landscape'

Conclusion

In this project, we developed a hierarchy for efficient retrieval of video segments for a given textual query. The hierarchy is built on tags annotated to the key frames extracted from the videos. The key frames are annotated using their visual and semantic properties. We observed that the hierarchy is most efficient when the query terms occur frequently and are related.

Currently, due to resource constraints, we are forced to contruct multiple hierarchies to accommodate large video sets. Faster retrieval can be achieved if a single hierarchy is constructed for the entire dataset. The work can also be improvised by parallelizing the search through the multiple hierarchies.

References:

- Yunchao Gong, Qifa Ke, Michael Isard, and Svetlana Lazebnik. A multi-view embedding space for modeling internet images, tags, and their semantics. International Journal of Computer Vision, 106(2):210-233, 2014.
- Li-Jia Li, Chong Wang, Yongwhan Lim, David M Blei, and Li Fei-Fei. Building and using a semantivisual image hierarchy. In Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on, pages 3336-3343. IEEE, 2010.
- Jia Deng, Alexander C Berg, and Li Fei-Fei. Hierarchical semantic indexing for large scale image retrieval. In Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on, pages 785{792. IEEE, 2011.