

Object Mining for Large Video Data

R. Shah
ronakshah@ee.iitb.ac.in
R. Iyer
rishabh_nad@iitb.ac.in
S. Chaudhuri
www.ee.iitb.ac.in/~sc

Vision and Image Processing Lab,
Electrical Engineering Department,
Indian Institute of Technology, Bombay,
India.

The task of object mining has attracted a lot of attention of vision researchers. An efficient object mining scheme has many applications, including a mechanism for accessing data in an organized manner, selective video playback and summarization. Traditionally this task has been attempted in two ways. Several approaches cast this problem as one of extracting frequently occurring patterns. The over-all problem is then cast in the manner analogous to searching in a text corpus [7], [8], [5]. Such approaches fail to utilize temporal redundancy in video data. Another set of approaches attempt to organize a very large number of images in an ordered set through object mining [1], [4]. However, they either suffer from low recall rate or provide a very redundant representation.

In the proposed work, we perform object grouping to capture temporal redundancy in video data and provide a compact graph-based representation. Further, we utilize scripts and subtitles to construct a separate graph-based representation for each scene. Advantages of the proposed scheme include faster object retrieval, efficiency in further processing of video data and summarization as a spin-off.

Scripts and subtitles provide complimentary information for an underlying video. For combining such complimentary information robustly, we utilize Lavenshtein distance with dynamic time warping. This process is shown in Figure 1.

