A CONTEXTUAL MODEL FOR SEMANTIC VIDEO STRUCTURING (ThuPmPO1)

Author(s) : Bruno Janvier Janvier (Université de Geneve, Switzerland)

Abstract : Semantic video structuring is a vital problem to solve for automated management of large video collections. The goal is to automatically extract from the raw data the inner structure of a video collection: so that a whole new range of applications to browse and search video collections can be derived out of this high–level segmentation. To reach this goal, we exploit techniques that consider the full spectrum of video content. It is fundamental to properly integrate technologies from the fields of computer vision, audio analysis, natural language processing and machine learning. In this paper, a multimodal feature vector providing a rich description of the audio, visual and text modalities is first constructed. Boosted Random Fields are then used to learn two kind of relationships : between features and labels and between labels concerning various modalities for improved consistency of the results. The parameters of this complex model are found iteratively by using two successive stages of Boosting. We experimented using the TRECvid corpus and show results that validate the approach over existing studies.

Menu