A General Framework for Edited Video and Raw Video Summarization

Objective:

Main objective of the project is to learn a compact dictionary from the original video as the video summary.

Abstract:

In this paper, we build a general summarization framework for both of edited video and raw video summarization. Overall, our work can be divided into three folds: 1) Four models are designed to capture the properties of video summaries, i.e., containing important people and objects (importance), representative to the video content (representativeness), no similar key-shots (diversity) and smoothness of the storyline (storyness). Specifically, these models are applicable to both edited videos and raw videos. 2) A comprehensive score function is built with the weighted combination of the aforementioned four models. Note that the weights of the four models in the score function, denoted as property-weight, are learned in a supervised manner. Besides, the property-weights are learned for edited videos and raw videos, respectively. 3) The training set is constructed with both edited videos and raw videos in order to make up the lack of training data. Particularly, each training video is equipped with a pair of mixing-coefficients which can reduce the structure mess in the training set caused by the rough mixture.

Introduction:

NOWADAYS, with the popularity of camera devices, a sheer amount of videos are captured and shared online Every day, vast video data floods the Internet social networking platform. It provides users a convenient way to access to video data. However, it also makes data browsing time-consuming. So we urgently need an efficient way to handle these huge video data. Fortunately, video summary can be a good assistance in the data explosion era . It offers viewers the video gist by generating a compact version of the video content. There are mainly two categories of video summary. One is storyboard, which consists of key-frames . The other is video skim, which is composed of video segments, namely key-shots.

Usually, video shots are generated by uniform cutting or segmentation models. The two versions of video summary have individual advantages, e.g., storyboard represents the video with just a few frames, while video skim can retain the dynamic characteristics of the original video. Both of them can not only provide a viewer friendly way to video browsing, but also have a wide range of applications, such as activity recognition , event detection , and video embedding etc..