Generative Model Based Video Shot Boundary Detection for Automated Surveillance

Biswanath Chakraborty, RCC Institute of Information Technology, Kolkata, India
Siddhartha Bhattacharyya, RCC Institute of Information Technology, Kolkata, India
Susanta Chakraborty, IIEST, Howrah, India

ABSTRACT

Video shot boundary detection (SBD) or video cut detection is one of the fundamental processes of video-processing with respect to semantic understanding, contextual information accumulation, labeling, content-based information retrieval and many more applications, such as video surveillance and monitoring. In this work, the authors have proposed a generative-model based framework for detecting shot boundaries in between the frames of a video segment. To generate a model of shot-boundaries, the authors have applied the concepts of Support Vector Machine to estimate the distance between any two images, and then, have generated a Gaussian Mixture Model from the estimated distances. Next, a Bayesian Estimation process checks the presence of boundaries in between the images by exploiting the Gaussian Mixture-based boundary model. Further, the authors have used the principles of Compressive Sensing to reduce the overhead of boundary detection process without losing of important information.

KEYWORDS
Compressive Sensing, Support Vector Machine, Surveillance, Video Shot Boundary Detection

INTRODUCTION

Video shot boundary detection or video cut detection technique has huge demand in software for various kinds of surveillance system as well as postproduction of videos, as it is the initial step of video image processing specifically applied to content based video retrieval, video summarization and video segmentation. Video shot boundary detection is still a challenging proposition in various types of video analysis applications. A shot can be defined as a similar in nature to a sequence of frames taken in one camera shot, and a scene can be expressed as a collection of one or more consecutive shots that focus on a particular event(s). For an example in a high road no traffic, normal traffic, heavy traffic or any accidental scenarios may be defined as different. In any high road crossing, more than one camera shots showing different cars running in different direction which may be treated as a scene (Dey et al., 2016). There are different types of cuts and transitions between the shots. A hard cut is an abrupt change between two consecutive video image frames and transitions me be defines as with slow changes in brightness. A dissolve occurs when the images of any shots looks less bright and the next consecutive image frames looks brighter, hence the transition is take a superimposition. The problem of video image analysis has been carried out using several techniques ranging from classical probabilistic thresholding to intelligent classification techniques. Based on the types of transitions

DOI: 10.4018/IJACI.2018100105

Copyright © 2018, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.
between two consecutive shots boundaries, it may be categorized into two types: cut (CUT) and gradual transition (GT). Moreover, gradual transition may further be classified into dissolve, wipe and fade out/in (FOI). Video shot boundary detection (SBD) is aimed to identifying two adjacent shots. As far as the benchmark databases are concerned, NIST has a huge evaluation database. The inception of TRECVID (http://trecvid.nist.gov/) has rapidly advanced the research in SBD. As far as by study, it is found that researchers have been able to detect CUTs to some remarkable extent, but the detection of GT still remains a challenging position. For detection of GT, it is really difficult to select an optimal feature set and preparation of cluster which may clearly indicate GT in any type of videos.

The proposed approach for video shot boundary detection can be broadly partitioned into two sub-processes: (1) Model Generation, and (2) Boundary Detection. The proposed work has been compared with (Mohanta, Saha & Chanda, 2012) which is based on Model-Based Shot Boundary Detection Technique Using Frame Transition Parameters. Experimental results showing remarkable success to detect GT as feature set has been selected in such a manner so that proposed model can easily classify boundary and no-boundary shots very efficiently.

Next, briefly describe our proposed approach, which consists of two sub-procedures: (a) Model Generation and (b) Boundary Detection in section Proposed Methodology.

**MOTIVATION**

As on date several methods have been published for video shot boundary detection. But it is very difficult to compare with the available methods. Because of implementation published method is not available always and preparing of full system is very difficult also. Another reason is most of the methods tested on similar kind of video sequences. During experiment it is found that most of the method works on a specific type/category of video but not for all type of videos or wide range of videos. Here the author has proposed a framework which can be applied on all type/category of video shot boundaries i.e. to detect hard cut as well as gradual changes of any kind of video scene. Another intension is to develop such framework which will work fast with minimal resource and computing facilities. Proposed framework can be applied to identify point of changes in multiple images of the similar type of scene recorded at different time for video surveillance. Say for an example, in a smart city where entire city is surveillance by CCTV cameras. Any unnatural activity alerts the city administration based on some pre-known pattern of incidents or abrupt changes in scene. For example, someone is being chased by goons or two gangs of goons are getting into confrontation or an ATM is being looted by some hooligans, or a fire break out incident in a locality, abnormal traffic in high road etc. all can alert the city administration automatically and inform respective departments of concerned administration without any manual intervention (Haribaabu & Joseph James, 2016).

**BACKGROUND/SURVEY OF EXISTING APPROACHES**

A detailed survey on VSBD is available in the literature (Rowe et al., 2005, Lienhart, 2001, Gargi et al., 2000, Boreczky et al., 1996, Lienhart, 1999; Choubey et al., 1997). In this section, it is tried to discuss mainly on three categories of related works on SBD based on (A) Methods based on visual content representation (B) Methods of classification (C) Methods of transition detection and (D) SBD System based on Support Vector Machine.

**Methods Based on Visual Content Representation**

Techniques like pixel-based, histogram, and mean and standard deviation of intensities have been successfully introduced and it is found that simple histogram feature shows satisfactory result compared to other complicated features such as edges (Hu et al., 2011). Also, it is found that the pixel-based method is sensitive to local/global movement of pixels. Zhang et al. (Zhang, Low & Smoliar, 1995;
The Application of Rough Set Theory and Near Set Theory to Face Recognition Problem
www.igi-global.com/chapter/the-application-of-rough-set-theory-and-near-set-theory-to-face-recognition-problem/140462?camid=4v1a

Mathematical Programming for Modelling Green Supply Chains Under Randomness and Fuzziness
www.igi-global.com/article/mathematical-programming-for-modelling-green-supply-chains-under-randomness-and-fuzziness/171653?camid=4v1a