Medical Video Summarization using Central Tendency-Based Shot Boundary Detection

G. G. Lakshmi Priya, Department of Computer Applications, National Institute of Technology, Tiruchirappalli, Tamil Nadu, India

S. Domnic, Department of Computer Applications, National Institute of Technology, Tiruchirappalli, Tamil Nadu, India

ABSTRACT

Due to the advancement in multimedia technologies and wide spread usage of internet facilities; there is rapid increase in availability of video data. More specifically, enormous collections of Medical videos are available which has its applications in various aspects like medical imaging, medical diagnostics, training the medical professionals, medical research and education. Due to abundant availability of information in the form of videos, it needs an efficient and automatic technique to manage, analyse, index, access and retrieve the information from the repository. The aim of this paper is to extract good visual content representatives – keyframes. In order to achieve this, the authors propose a new method for video shot segmentation which in turn leads to extraction of better keyframes as representative for summary. The proposed method is experimented and evaluated using publically available medical videos. As a result, better precision and recall is obtained for shot detection when compared to that of the recent related methods. Evaluation of video summary is done using fidelity measure and compression ratio.

Keywords: Feature Extraction, Keyframe Extraction, Medical Videos, Shot Clustering, Video Summary

1. INTRODUCTION

Recent development in multimedia systems resulted in enormous availability and usage of digital videos. More specifically, in medical fields, the availability of multimedia data (videos) goes on increasing. Some developing countries suffer from lack of facilities in accessing this information and are in situation providing inadequate training to the medical professions. Since medical videos are mainly used for educational purposes, video content is usually recorded or edited using the style formats. In order to make use of this huge amount of data for various medical related applications there is a need for tools that can effectively analyse, categorize, search and retrieve the relevant video material. This paved way for various researches in video summarization.

In general, video summary is a sequence of still (representative frames) or motion pictures (video skimming) that provides concise infor-
mation about the content of the video. Representative frames are the set of keyframes extracted from the video, whereas video skimming is the summary composed of set of shots. Compared to video skim, keyframe sets are more flexible in terms of browsing and navigation purpose.

Various works have been proposed for video summarization, which are categorized into sample based, cluster based and shot based approaches (Truong & Venkatesh, 2007). In the sampling based approaches, keyframes were extracted by randomly choosing of uniformly sampling from the original video. It is the straightforward and easy way to extract keyframes. It may fail to capture real video contents because of the random extraction of keyframes.

Most of video summarization approaches are based on clustering techniques (Hadi et al., 2006; Mundur et al., 2006; Chen et al., 2009; Herranz & Martinez, 2009; Furini et al., 2010). The basic idea of clustering approaches is to generate summary by grouping similar frames / shots and extracting frames from the groups (usually one frame per group). However, it is important to select the features in order to consider and group the similar frames. There are some techniques that produce summaries of accepted qualities but computationally expensive (Mundur et al., 2006, Furini et al., 2010, Amiri & Fathy, 2010). In Borth et al. (2008), Borth uses shot boundary detection to segment the video into shots and the k-means algorithm to determine cluster representatives for each shot that are used as keyframes. In paper Besiris et al. (2009), the author proposed an automatic video summarization technique based on graph theory methodology and the dominant sets clustering algorithm. Ali Amiri and Mahmood Fathy (2010) proposed a Hierarchical keyframe based video summarization system using QR-decomposition. Based on the dynamicity of the shots the number of keyframes is selected. Avila et al. (2011) presented VSUMMM, a methodology for production of video summary, is based on color feature extraction from frames and k-means clustering.

In the shot-based approaches, the video is segmented into separated shots and one or more keyframes extracted from each shot. In most of the literature, shot boundary detection is widely used as first step in generation of video summaries. The shot boundary detection is segmentation of the video into its basic meaningful and manageable element, called a shot. A shot is defined as a sequence of frames captured uninterruptedly by a single camera in time and space. In general, Shot boundaries are classified into abrupt and gradual transitions. If an immediate transition takes place from one shot to another then it is considered as abrupt shot or cut transition. When the shot changes gradually with few intermediate frames, then it is gradual transition which includes fade-in, fade-out, dissolve and wipe.

The detection process involves on three steps: visual content representation (feature extraction), construction of continuity signal (similarity between the consecutive frames) and the classification of continuity values (Transition identification). The video frames itself contains various information about the visual content and also possible to obtain more valuable information by processing the frame contents. Automatic detection of shot change is still a difficult problem, due to the variety of transitions that can be used between shots (Lienhart, 1999) and occurrence of the fast motion in the video sequence. The selection of features, similarity metrics and the classification methods should be in such a way that it should yield better performance by reducing the influences of the SBD issues. On identifying the transition type and region of transition the frames are segmented into shots. Using various techniques, keyframes are extracted from the shot as its representative. In papers (Zhu et al., 2004; Li et al., 2005; Cernekova et al., 2006; Hadi et al., 2006; Chang & Chen, 2007) shot changes are detected in various ways and the extraction of video summary rely on the correctly detected shots.
Related Content

Classifying Behaviours in Videos with Recurrent Neural Networks

www.igi-global.com/article/classifying-behaviours-in-videos-with-recurrent-neural-networks/195006?camid=4v1a

Enhancement of Hybrid Multimodal Medical Image Fusion Techniques for Clinical Disease Analysis

www.igi-global.com/article/enhancement-of-hybrid-multimodal-medical-image-fusion-techniques-for-clinical-disease-analysis/212374?camid=4v1a
Two-Dimensional Canonical Correlation Analysis of the Logically Concatenated Cross Binary Pattern for Cross Pose Face Recognition
[www.igi-global.com/article/two-dimensional-canonical-correlation-analysis-of-the-logically-concatenated-cross-binary-pattern-for-cross-pose-face-recognition/212376?camid=4v1a](www.igi-global.com/article/two-dimensional-canonical-correlation-analysis-of-the-logically-concatenated-cross-binary-pattern-for-cross-pose-face-recognition/212376?camid=4v1a)

Selected Shape and Texture Features for Automatic Detection of Acute Lymphoblastic Leukemia
Vanika Singhal and Preety Singh (2018). *Biomedical Signal and Image Processing in Patient Care* (pp. 162-183).
[www.igi-global.com/chapter/selected-shape-and-texture-features-for-automatic-detection-of-acute-lymphoblastic-leukemia/188156?camid=4v1a](www.igi-global.com/chapter/selected-shape-and-texture-features-for-automatic-detection-of-acute-lymphoblastic-leukemia/188156?camid=4v1a)