ISSN: 2393-8528

IJICSE
Open Access
Journal Approved by UGC

Contents lists available at www.ijicse.in

International Journal of Innovative Computer Science & Engineering

Volume 4 Issue 3; May-June-2017; Page No. 160-165

Key frame based video summarization using frame difference

Smt.M.Tirupathamma

Assistant Professor, Department of ECE, JNTUHCEJ, Jagtial, Telangana, India

tirupathamma.jntu@gmail.com

Received 01 June 2017; Accepted 29 June 2017

ABSTRACT

The aim of this paper is to extract the video frames from the original video that would be the most informative and concise representation of the whole video. Such frames are referred as key frames. Key frames are extracted for video summarization using Frame Difference method. Key frames based video summarization works on frames so initially a video frame sequence is divided into frames. The redundant content from the extracted video frames is discarded by calculating the frame difference between the adjacent frames. The frames whose difference is greater than certain threshold are considered as Key frames. The extracted Key frames are combined to form a summarized video and played in MATLAB using few commands.

Keywords: *Video, key-frame, summarization, frame.*

INTRODUCTION

The twenty-first century is an age of information explosion. We are witnessing a huge growth in digital data [1]. The trend of increasing information access boosts the requirement for progress in multimedia technology. Video is the most effective media for capturing the world around us. But due to the huge amount of multimedia repositories, the browsing, retrieval and delivery of video contents very slow and even difficult tasks [2]. In order to extract valid information from video, video data must be processed efficiently, and the transfer stress of network must be reduced, more and more attention is being paid to the video processing technology.

Video summarization has been proposed to improve faster browsing of large video collections and more efficient content indexing and access. As the name implies, Video summarization is a mechanism to produce a short summary of a video to give to the user a synthetic and useful visual abstract of video sequence, it can either be key frames or video skims. In terms of browsing and navigation, a good video abstract will enable the user to gain maximum information about the target video sequence in a specified time constraint or sufficient information in the minimum time. Automatically generated summaries can support users in navigating large video archives and in taking decisions more efficiently regarding selecting, consuming, sharing, or deleting content.

The video summaries can be generated in different forms. The major challenge in video processing is the amount of data that has to be processed. The huge volume of data that undergoes processing can be reduced to a large extent. This is achieved by eliminating the redundant frames of the video by using efficient key frame extraction techniques [3].

Section 2 discusses about literature review, section 3 talks about proposed algorithm, section 4 projects the results, and section 5 presents the discussion and conclusion.

LITERATURE REVIEW

The rapid development of digital video capture and editing technology led to increase in video data, creating the need for effective techniques for video retrieval and analysis. Advances in digital content distribution and digital video recorders, has caused digital content recording easy. However, the user may not have enough time to watch the entire video. In such cases, the user may just want to view the abstract of the video instead of watching the whole video which provides more information

**Corresponding author: Smt.M.Tirupathamma |*

about the occurrence of various incidents in the video [4].

Video summarization is a summary which represents abstract view of original video sequence and can be used as video browsing and retrieval systems. The video abstraction process usually has three phases: Video information analysis, meaningful clip selection and output synthesis. The summarized video can be a highlight of original sequence which is the concatenation of a user defined number of selected video segments or can be a collection of key frames. There are various approaches for summarizing a video such as Key frame based summarization, Video skim based summarization etc.

Types Video summarization Methods

Video can be summarized by two different ways which are as follows [5].

Key Frame Based Video Summarization: Static video summary represents a video sequence in a static imagery form. One or more selected representative frames from the original video, or a synthesized image generated from the selected key frames. According to different sampling mechanisms, a set of key frames are extracted from shots of the original video These are also called representative frames, R-frames, still-image abstracts or static storyboard, and a set consists of a collection of salient images extracted from the underlying video source. Then, the selected key frames are arranged or blended in a two-dimensional space.

Following are some of the challenges that should be taken care while implementing Key frame based algorithm.

i. Redundancy: frames with minor difference are selected as key frame.

ii. When there are various changes in content it is difficult to make clustering.

Video Skim Based Video Summarization: Dynamic summarization consists in selecting the most relevant small dynamic portions (video skims) of audio and video in order to generate the video summary. This is also called a moving-image abstract, moving story board, or summary sequence. The original video is segmented into various parts which is a video clip with shorter duration. Each segment is joined by either a cut or a gradual effect. The trailer of movie is the best example for video skimming.

Different Types of Key-frame based Video summarization Methods

Key frames based video summarization can be classified in three different ways.

These are as follows.

Summarization based on sampling: This method chooses key frames uniformly or randomly under-sampling, without considering the video content. The summary produced by these methods does not represent all the video parts and may cause some redundancy of key frames with similar contents.

Summarization based on scene segmentation: This method extracts key frames using scenes detection, the scene includes all parts with a semantic link in the video or in the same space or in the same time. The disadvantage of these techniques is producing a summary, which does not take into account the temporal position of frames.

Summarization based on shot segmentation: This method extracts adapted key frames to video content. They extract the first image as shot key frames or the first and the last frames of the shot.

These methods are effective for stationary shot and small content variation, but they don't provide an adequate representation of shot with strong movements.

Different Types of Key-frame extraction methods

A key frame is a representative of each shot detected. The increased demand for intelligent processing and analysis of multimedia information has led to the development of different approaches for intelligent video management. The researchers have attempted to exploit various features for the extraction of key frames in videos. The majority of techniques focus on the extraction of key frames.

Different methods can be used to select key frames some of them are as discussed below [6].

By using triangle model of perceived motion energy (PME) motion patterns are modeled in video. The frames at the turning point of the motion acceleration and motion deceleration are selected as key frames. The key-frame selection process is threshold free and fast and the extracted key frames are representative [7].

In Visual frame Descriptors algorithm three visual features: color histogram, wavelet statistics and edge direction histogram are used for selection of key frames. Similarity measures are computed for each descriptor and combined to form a frame difference measure. Fidelity, Shot Reconstruction Degree, Compression Ratio qualities are used to evaluate the video summarization.

In Motion Attention Model shots are detected using color distribution and edge covering ratio that increase the accuracy of shot detection. Key frames are extracted from each shot by using the motion attention model. Here the first and last frame of every shot is considered as key frame and the others are extracted by adopting motion attention model. These key frames are then clustered and a

© 2017 All Rights Reserved.

priority value is computed by estimating motion energy and color variation of shots.

In Multiple Visual Descriptor Features algorithm, the key frames are selected by constructing the cumulative graph for the frame difference values. The frames at the sharp slope indicate the significant visual change; hence they are selected and included in the final summary.

Motion focusing method focuses on one constant-speed motion and aligns the video frames by fixing focused motion into a static situation. A summary is generated containing all moving objects and embedded with spatial and motion information. Background subtraction and min cut are mainly used in motion focusing.

In Camera Motion and Object Motion, the video is segmented using camera motion-based classes: pan, zoom in, zoom out and fixed. Final key frame selections from each of these segments are extracted based on confidence value formulated for the zoom, pan and steady segments.

For instance, Pal and Leigh used fuzzy geometrical and information measures to develop an algorithm to estimate the difference between two consecutive frames. The similarity between the frames was measured in terms of weighted distance in fuzzy feature space. Hanjalic et al. compared the difference in color histograms of consecutive frames with a threshold to obtain key frames.

Introduced by DeMenthon et al. The key frames were extracted by finding discontinuities on a trajectory curve, which represent video sequence.

In the Flexible Rectangles algorithm, the frame differences were used to build a ‘‘Content Development Curve’’ from a curve composed of a predefined number of rectangles through the use of an error minimization algorithm.

The Adaptive Sampling algorithm extracted key frames by uniformly sampling the y-axis of the curve of cumulative frame differences .The resulting sampling on the x-axis represented the key frames.

The Shot Reconstruction Degree Interpolation [18] selected the key frames based on the ability of frames to reconstruct the original video shot using frame interpolation.

Ciocca and Schettini extracted key frames by first finding the cumulative frame differences based on certain frame descriptors such as color histogram, histogram of edges and wavelets. Next, a curve of cumulative frame differences was sketched, and then the midpoints of two curvature points on this curve were selected as key frames. A curvature point is a point on the curve where the angle changes are drastic.

Some researchers used clustering for extracting key frames by treating video frames as points in the feature space. The core idea behind such techniques is to cluster the frames based on some similarity measure and then select one key frame from each cluster.

Though there are several methods for Key frame based video summarization, there is no any universally accepted method available that gives better output in all kinds of videos. In this paper we are going to deal with the Frame difference method for video key frame extraction.

The frame difference based methods are intuitive and simple in nature. These properties make them suitable for many real-time and/or online applications. However, for extracting a particular key frame, these techniques only consider sufficient content change between the consecutive frames. Therefore, a key frame that is extracted by these methods does not fully represent the portion of the video preceding it.

A video can be segmented into different units, such as frames. The complete moving picture in a video can be discretized to a finite image sequence, i.e., many still images. Each still image is called a “frame”, which is the basic unit of the video. The image sequence is naturally indexed by the frame number. All the frames in one video have a same size and the time between each two frames is equal, typically 1/25 or 1/30 seconds. A video is defined as a series of interrelated consecutive frames taken contiguously by a single camera and representing a continuous action in time and space.

The objective is to represent the most “important” or “meaningful” scenes of the large amount of visual information by only a few images, i.e., the key frames. In order to achieve the objective compute the frame differences based on some criteria and then discard the frames whose difference with the adjacent frames are less than a certain threshold.

PROPOSED ALGORITHM

Key frame based video summarization method consists of a collection of salient images extracted from the underlying video source. Key frames based video summarization works on frames so first step is to extract frames from original video frame sequence. Later the redundant content from the extracted video frames is discarded and the selection of key frames is done. Finally the extracted key frames are combined to form a summarized video.

Algorithm and Flow chart for the implementation of extraction of key frames is discussed in this chapter. Here is the algorithm for extraction of Key frames

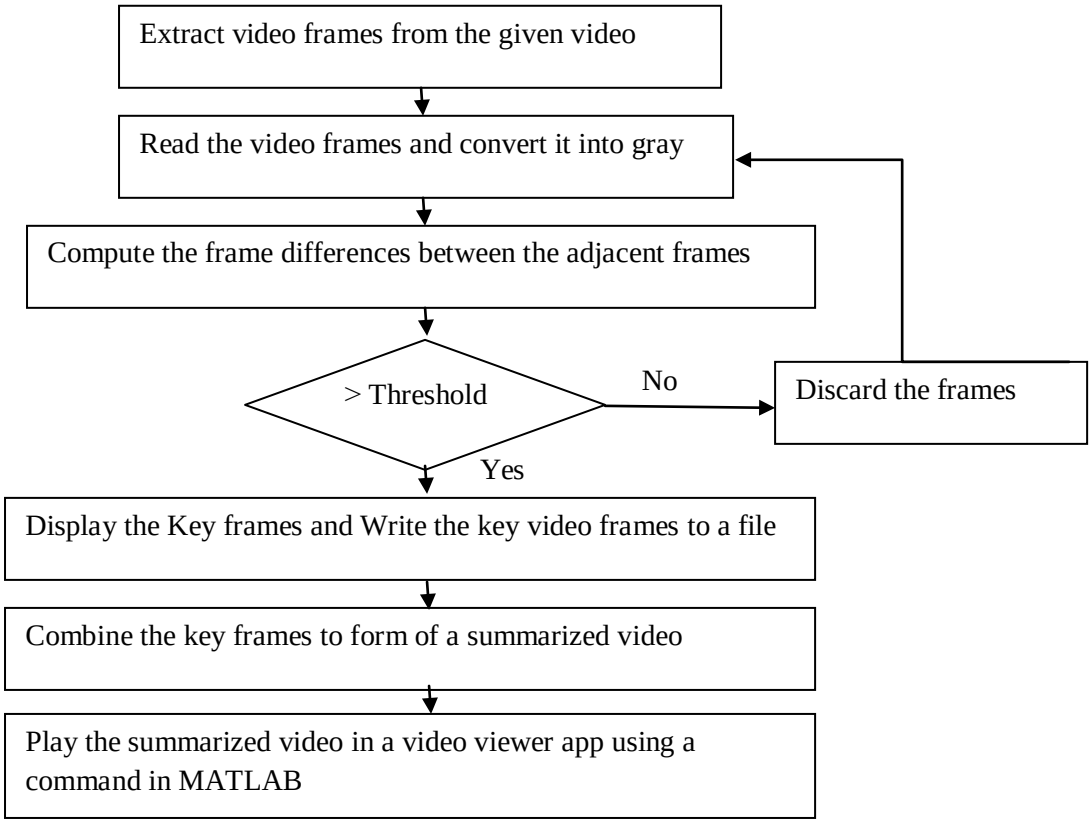
© 2017 All Rights Reserved.

for video summarization using Frame difference method.

Algorithm

Step 1: Read the input video and get the properties of the video.
Step 2: Convert the input video into frames.
Step 3: Open the video writer object file to write summarized video data.
Step 4: Read the video frames of the input video and convert it to gray.
Step 5: Calculate the difference between the adjacent frames.
Step 6: Quantize the difference for better result.
Step 7: If the difference between the adjacent frames is greater than threshold, consider the frame as Key frame and write it to the video writer object file.
Step 8: Else discard the frames and compute the difference of next two frames.
Step 9: When all the frames of input video are read and difference is computed close the writer object file.
Step 10: Key frames extracted are combined to form a new video which is the summarized video of the given input video.
Step 11: The summarized video is played in video viewer app using a command in MATLAB.

Flow-chart



RESULTS

The proposed method is implemented in MATLAB Version R2014a. Image and video processing tools are employed here. When the algorithm discussed in the previous chapter is executed, key frames are extracted from the given input video using Frame difference method.

The resultant five Key frames extracted from the input video are shown below.

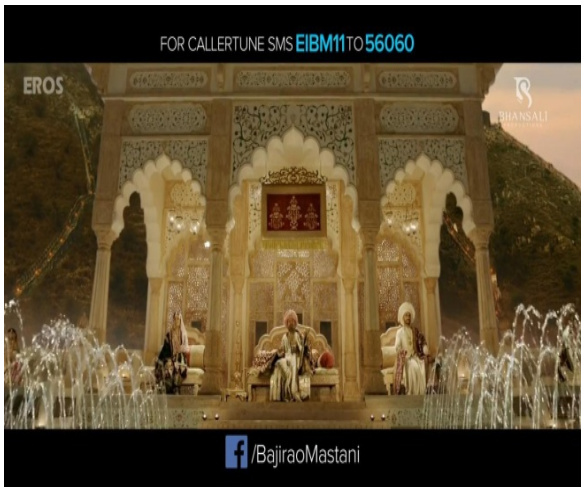


Fig 1: 1st Key frame

© 2017 All Rights Reserved.

Fig 2: 2nd Key frame

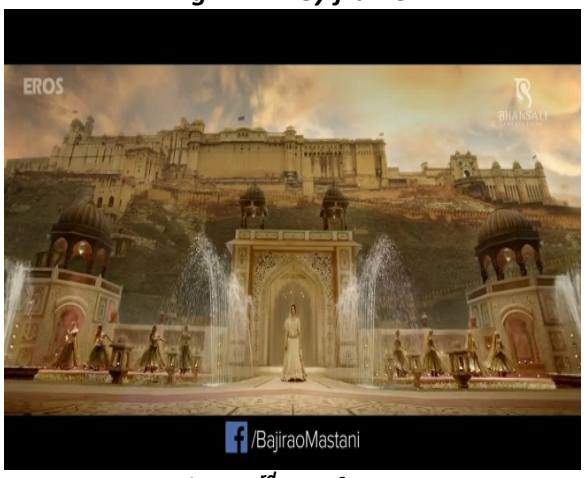


Fig 3: 3rd Key frame

Fig 4: 4th Key frame

Fig 5: 5th Key frame

These five Key frames are summarized and played in the form of a video using implay command in MATLAB as shown.

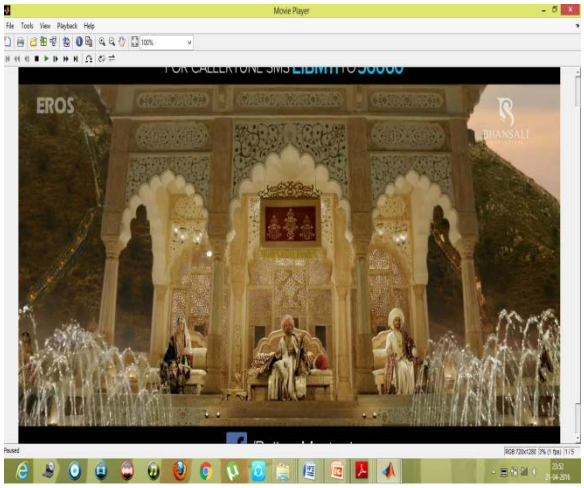


Fig 6: Summarized video played in MATLAB

DISCUSSION AND CONCLUSION

Many professional and educational applications that involve generating or using large volumes of video and multimedia data are prime candidates for taking advantage of video content analysis techniques. The developed techniques in video summarization touch various domains; we find in, three categories presented: Consumer video applications, Image-Video databases management and surveillance. For each category, some of the exemplar applications are listed.

With the increasing in the storage and computational capacity of consumer electronic devices such as personal video recorders (PVR), consumer video applications enables the end user of browsing the recorded content in efficient ways and view the interesting parts quickly.

On the other hand, Image and video databases management includes different application areas like video search engine, digital video library, object indexing and retrieval, automatic object labeling and object classification.

Consequently, Media organizations and TV broadcasting companies have shown considerable interest in these applications, especially in organizing and indexing large volumes of video data to facilitate efficient and effective use of these resources for internal use. These large video libraries create a unique opportunity for using intelligent media analysis techniques to create advanced searching and browsing techniques to find relevant information quickly and inexpensively. Intelligent video segmentation and sampling techniques can reduce the visual contents of the video program to a small number of static images.

© 2017 All Rights Reserved.

We can browse these images to spot information and use image similarity searches to find shots with similar content and motion analysis to categorize the video segments. Higher level analysis can extract information relevant to the presence of humans or objects in the video. Audio event detection and speech detection can extract additional information to help the user find segments of interest.

Recent years have witnessed an enormous increase in video data on the internet. This rapid increase demands efficient techniques for management and storage of video data. Video summarization is one of the commonly used mechanisms to build an efficient video archiving system. Video summarization plays important role in many video applications. A survey on various methods for key frame based video summarization has been carried out. But there is no any universally accepted method available for video summarization that gives better output in all kinds of videos. The summarization viewpoint and perspective are often application-dependent. The semantic understanding and its representation are the biggest issues to be addressed for incorporating diversities in video and human perception. Depending upon the changes in contents of the video, the key frames are extracted. As the key frames need to be processed for summarization purpose, the important contents must not be missed.

The video summarization methods generate summaries of the videos which are the sequences of stationary or moving images. Video abstraction is an integral part of many video applications, including video indexing, browsing, and retrieval. The summarized video decreases the required storage area and increases the computational efficiency in extracting the key frames from the given video. The key frames are nothing but the meaningful part of the entire video eliminating the redundant frames of the frames. The purpose is to remove the redundant frames, reduce the computational complexity and improve recognition efficiency.

In this paper key frames are extracted for the video summarization by using Frame difference. The method used is computationally simple and dynamically determines the number of key frames and has high accuracy rate and low error rate. Threshold value is dynamic and will be changing during the course of each video so that key frames can be retrieved efficiently. The extracted key frames can satisfactorily represent the content of video.

This summarization tool can be further extended to for feature extraction for the video retrieval and indexing mechanism. The main future enhancement to this paper is the end-end single package development from the video summarization to video indexing and retrieval mechanism.

With higher level analysis this paper can extract information relevant to the presence of humans or objects in the video.

REFERENCES

1. Rong, Jiawei, Wanjun Jin, and Lide Wu. "Key frame extraction using inter-shot information." Multimedia and Expo, 2004. ICME'04. 2004 IEEE International Conference on. Vol. 1. IEEE, 2004.
2. Ahanger, Gulrukh, and Thomas DC Little. "A Survey of Technologies for Parsing and Indexing Digital Video1." Journal of visual Communication and image representation 7.1 (1996): 28-43.
3. Kuehne, Hildegard, et al. "HMDB: a large video database for human motion recognition." Computer Vision (ICCV), 2011 IEEE International Conference on. IEEE, 2011.
4. Dimitrova, Nevenka, et al. "Applications of video-content analysis and retrieval." IEEE multimedia 9.3 (2002): 42-55.
5. Gygli, Michael, et al. "Creating summaries from user videos." Lecture Notes in Computer Science. Vol. 8695. Springer, 2014.
6. Chitra, Dhawale, and S. Jain, “A novel approach towards key frame selection for video summarization,” Asian Journal of Information Technology, vol. 7, no. 4, pp. 133–137, 2008.
7. Liu, Tianming, Hong-Jiang Zhang, and Feihu Qi. "A novel video key-frame-extraction algorithm based on perceived motion energy model." IEEE transactions on circuits and systems for video technology 13.10 (2003): 1006-1013.

© 2017 All Rights Reserved.