IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

Comparative Analysis Of Video Summarization Methods

Aarti Kumthekar[*1], Mrs.J.K.Patil[2]

[*1,2] Department of Electronics and Telecommunication, Bharati Vidyapeeth's College of Engineering, Kolhapur, India

aartikumthekar@gmail.com

Abstract

This paper presents a comparative study on video summarization methods. Shot segmentation and key frame extraction are major phases of video summarization. Many segmentation methods have been brought forward to deal with Shot segmentation like Pair wise pixel comparison, likelihood ratio, Motion Activity Descriptor Based Algorithm, Shot boundary based approach and histogram difference. Among these methods histogram difference with automatic thresholding is the simple and important one. It is found that the transition effects occurred during shot segmentation can be well detected by histogram difference algorithm.

Introduction

The multimedia information indexing and retrieval are required to describe, store, and organize multimedia information and to assist people in finding multimedia resources conveniently and quickly. Dynamic video is an important form of multimedia information. Videos have the following characteristics: 1) much richer content than individual images; 2) huge amount of raw data; and 3) very little prior structure. Video summarization, aimed at reducing the amount of data that must be examined in order to retrieve a particular piece of information in a video, is, consequently, an essential task in applications of video summarization.

The growing interest of consumers in the acquisition of video and access to visual information has created a demand for new technologies to summarize the multimedia data. Very large databases of images and videos require efficient algorithms that enable fast browsing and access to the information pursued. A video summary is a sequence of still or moving images, with or without audio. These images must preserve the overall contents of the video with a minimum of data. Still images chronologically arranged form a pictorial summary that can be assumed to be the equivalent of a video storyboard. Video shot boundary detection has been deeply studied in recent years and has found application in various domains like video compression, video summarization and so on. Shot boundary detection and Key frame Extraction is a fundamental step for organization of large video data. Key frame extraction has been recognized as one of the important research issues in video information retrieval.

Video segmentation is of great importance in many applications, such as video databases, video compression and transmission, video retrieval and browsing, and so on. A video can be segmented into different units, such as frames, shots, or scenes. 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" [1], 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 shot is defined as a series of interrelated consecutive frames taken contiguously by a single camera and representing a continuous action in time and space [1]. In general, shots are joined together in a process called editing to produce a video. The unbroken image sequence in a shot usually has consistent content. While scene is a more semantic notion, which is essentially a story unit. Different shots are joined together in the editing stage of video. A video shot is defined as a series of interrelated consecutive frames taken contiguously by a single camera and representing a continuous action in time and space [1]. In general, shots are joined together in a process called editing to produce a video. The unbroken image sequence in a shot usually has consistent content. While scene is a more semantic notion, which is essentially a story unit. Different shots are joined together in the editing stage of video. There are several editing operations, with or without the use of the transition effect. Transitions between shots can be either 'abrupt' or 'gradual.

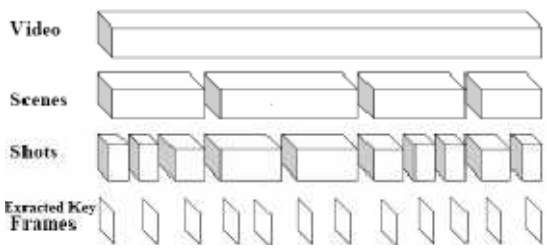

Fig 1:-Overview of shot boundary detection

During the video production process, first step is capturing the shots by using a single camera. Two consecutive shots are then attached together by a shot boundary that can either be abrupt or gradual. Abrupt shot boundaries are created by simply attaching a shot to another. While there is no modification in the consequent shots in an abrupt shot boundary, gradual transitions result from editing effects applied to the shots during attachment operation. According to the editing effect gradual transitions can be further divided into different subtypes. The number of possible transitions due to editing effect is quite high but most of the transitions fall into the three main categories: dissolve, fades (fade in, fade out), and wipes.

The main aim is to present an approach to segment a given video into different shots, find the beginning and end of each shot, and then extract a single frame representing the main content of each shot. Such a frame is so called key frame. The key frame can also be applied to the video content-based retrieval [6], content analysis [7], video summarization [5], editing and so on.

Key-frame based video representation views video abstraction as a problem of mapping an entire segment (both static and moving contents) to some small number of representative images. The features used for key frame extraction include colors (particularly the color histogram), edges, shapes, optical flow, camera activity, and features derived from image variations caused by camera motion.

The extraction of key frames must be automatic and content based so that they maintain the salient content of the video while avoiding all redundancy.

The number of possible transitions due to editing effect is quite high but most of the transitions fall into the three main categories: dissolve, fades (fade in, fade out), and wipes and that can be detected using this algorithm.

Related Key Frame extraction Techniques

In the past decade, there has been significant work done in the area of video processing to partition a given video into separate shots. Shot segmentation is the first step of the key frame extraction, which mainly refers to detecting the transition between successive shots. The basic idea of most of the techniques is to measure and compare the similarities between consecutive frames. The following are some of the technique-

Pair Wise Pixel Comparison

Pair wise pixel comparison [3] is a straightforward and simplest way, in which the number of pixels changed from a frame to the next is counted. When the total percentage of the pixels has changed, a shot is detected. In this algorithm individual pixels from frames are compared to find out frame difference. Pair-wise comparison evaluates the differences in intensity or color values of corresponding pixels in two successive frames. In this algorithm the pixel-wise difference algorithm gives quite acceptable results with adaptive thresholding. By considering difference between the difference signal values of adjacent frames is a worthwhile approach. In practice, it is observed that it is useful to reduce the difference signal with a threshold derived from the maximum and minimum difference signals over a small aperture. Even with the adaptive thresholding, the algorithm produces false alarms, if the shot before/after the shot boundary includes high motion activity. The reason can be explained as follows: The weakness of the pixel based features is the high sensitivity to the video content. It is difficult for this algorithm to understand whether the change in the continuity signal is due to shot boundary or due to disturbances/motion. In order to enhance the algorithm, adaptive thresholding can be used. However, the high level of activity in the images around shot boundary produces a larger difference signal than expected. As a result adaptively obtained threshold is larger. A threshold that is larger than expected results in missed shot boundary.

The main disadvantage of this method is its inability to distinguish between a large change in a small area and a small change in a large area. It is observed that cuts are falsely detected. The other disadvantage of this method is that it is quite sensitive to fast object movements and the camera motion - fast camera panning or zooming. Also it is sensitive to noise.

Likelihood Ratio

Likelihood ratio [3] is a region-based technique. It is a typical statistical difference method, which can be regarded as an extension to pixel difference. It can solve the problem of false detection due to small camera motions. Instead of comparing individual pixel, it compares the statistical characteristic, the so-called *likelihood ratio*, of the corresponding regions (i.e. blocks) in two successive frames. If the likelihood ratio is larger than a preset threshold, the

region is regarded as being changed. A shot can be declared if a certain number of regions have changed. A shot boundary is found if more than a certain number of blocks have changed. It is less sensitive to camera and object motion and noise.

Motion Activity Descriptor Based Algorithm

The motion activity is one of the motion features included in the visual part of the MPEG-7 standard [11]. It also used to describe the level or intensity of activity, action, or motion in that video sequence. The main idea underlying the methods of segmentation schemes is that images in the vicinity of a transition are highly dissimilar. It then seeks to identify discontinuities in the video stream. The general principle is to extract a comment on each image, and then define a distance (or similarity measure) between observations. The application of the distance between two successive images, the entire video stream, reduces a one-dimensional signal, in which we seek then the peaks (resp. hollow if similarity measure), which correspond to moments of high dissimilarity. In this work, the extraction of key frames method based on detecting a significant change in the activity of motion is used. To jump 2 images which do not distort the calculations but we can minimize the execution time. First the motion vectors between image i and image i+2 is extracted then calculates the intensity of motion, we repeat this process until reaching the last frame of the video and comparing the difference between the intensities of successive motion to a specified threshold.

Motion Capture Data By Curve Saliency

In this paper, Eyuphan Bulut & Tolga Capin proposes a new approach to find key frames in a motion captured sequence. Treat the input motion sequence as a curve, and find the most salient parts of this curve which are crucial in the representation of the motion behavior. We apply the idea of saliency to motion curves in the first part of our algorithm. Then in the second part, we apply key frame reduction techniques in order to obtain the most important key frames of the motion. This method is effective to a certain extent.

Proposed Work

We are proposing a novel method of extracting efficient key frame for video summarization based on the block based Histogram difference and edge matching rate.

The block diagram of this proposed work is shown in fig 2. The shot segmentation is the first step of the key frame extraction, which mainly refers to detecting the transition between successive shots. A shot represents a sequence of frames captured from a unique and continuous record from a camera. The domain of video shot segmentation falls into two categories: uncompressed and compressed. There are two different types of transitions that can occur between shots, abrupt (discontinuous) also referred as cut, or gradual (continuous) such as fades, dissolves and wipes. The cut, fades, dissolves can be detected using this algorithm.

After shot segmentation, key frames are to be extracted. Shot segmentation is the premise of key frame extraction, and key frames are the salient content of the video. Thus summarization can be done using key frame extraction.

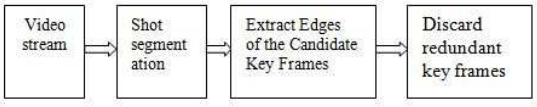

Fig 2: Block diagram of key frame extraction

The Histogram difference of every two consecutive frame is calculated, and then the edges of the candidate key frames are extracted by Prewitt operator. Then, the edges of adjacent frames are matched. If the edge matching rate is above average edge matching rate, the current frame is deemed to the redundant key frame and should be discarded. The edge detection can remove the irrelevant information and retain important structural properties of the image, so it helps to eliminate redundant data.

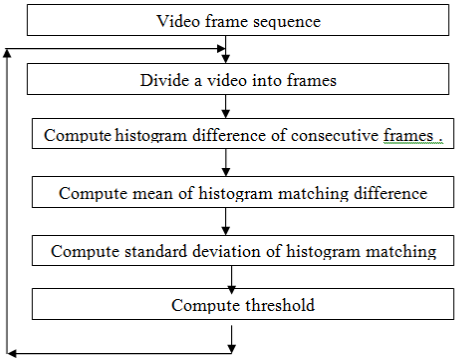

Fig 3: Flowchart for key frame extraction

The algorithm of key frame extraction is described as follows.

Let F(k) be the kth frame in video sequence, k = 1,2,…., Fv （ Fv denotes the total number of video). The detail step for video summarization technique is as follows:-

Step 1*:* Partitioning a frame into blocks with m rows and n columns.

Step 2: Computing histogram matching difference between the corresponding blocks between consecutive frames in video sequence.

Step 3: Computing histogram difference between two consecutive frames.
Step 4: Computing threshold automatically.
Step 5: Computing the difference between all the general frames and reference frame.
Step 6: Searching for the maximum difference within a shot
Step 7: Depending on shot type, selecting key frame.

The most critical activity in the shot segmentation process is the selection of the thresholds in any shot boundary detection step. The performance of the algorithm mainly remains in the thresholding phase. The computation of threshold is done automatically by calculating the mean and standard deviation of histogram difference over the whole video sequence. With the edge detection and edge matching, eliminate redundant key frames, improve the accuracy rate of the key frame extraction and reduce the redundancy.
Some key characteristics of the proposed approach are as follows:

- Efficiency: Easy to implement and fast to compute.
- No Redundant Keyframe: As Redundant Key Frames can be removed using edge detection algorithm and Information per Keyframe can be improved.
- Higher Recall & Precision: Algorithm provides higher Recall Rate and Precision Rate.

Conclusion

In this study, a novel algorithm of key frame extraction based histogram difference with automatic threshold selection is proposed. Key frame extraction aims to reduce the amount of video data, and the frame sequence must preserve the overall contents of the original video. The transition like cut, fade and dissolve can be detected using this algorithm**.** It compensates for the shortcomings of other algorithm such as Pair wise pixel comparison, likelihood ratio, Motion Activity Descriptor Based Algorithm, Shot boundary based approach and improves the techniques of key frame extraction.

Referencese

[1] Guozhu Liu and Junming Zhao "Key Frame Extraction from MPEG Video Stream" Third International Symposium on Information Processing, 2011
[2] Mika Rautiainenand DavidDoermann,"Temporal color correlograms for video retrival**",** 2002
[3] HaiyanXie," Key Frame Segmentation in Video Sequences", 2008
[4] Irena Koprinska Sergio Carrato,"Temporal video segmentation: A survey",feb 2000
[5] G. Ciocca1, R. Schettini"An innovative algorithm for key frame extraction in video summarization",
[6] Weiming Hu, NianhuaXie, Li Li, XianglinZeng, and Stephen Maybank ,"A Survey on Visual Content-Based Video Indexing and Retrieval".IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART C: APPLICATIONS AND REVIEWS, Vol. 41, no. 6, november 2011
[7] Janko Calic and Ebroul Izquierdo,Multimedia and Vision Research Lab, Queen Mary, University of London,"Efficient key-frame extraction and video analysis",
[8] M. Ceccarelli , A. Hanjalic , R.L. Lagendijk'**"**A sequence analysis system for video databases",
[9] ISIL BURCUN KAYAALP,"Video segmentation using partially decoded mpeg bitstream",DECEMBER 2003
[10] "Keyframe Based Video Summarization Using Automatic Threshold & Edge Matching Rate." Mr. Sandip T. Dhagdi, Dr. P.R. Deshmukh, 2012.International Journal of Scientific and Research Publications, Volume 2, Issue 7, July 2012 ISSN 2250-3153
[11] Abdelati Malek Amel, Ben Abdelali Abdessalem and Mtibaa Abdellatif "Video shot boundary detection using motion activity descriptor". JOURNAL OF TELECOMMUNICATIONS, VOLUME 2, ISSUE 1, APRIL 2010 ,54