ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 11, November 2013

Key Frame Extraction using Edge Change Ratio for Shot Segmentation

Azra Nasreen [1], Dr Shobha G [2]

Assistant Professor, Department of Computer Science and Engineering, R V C E, Bangalore [1]

Professor & Head, Department of Computer Science and Engineering, R V C E, Bangalore [2]

Abstract: With an ever growing amount of video content available, digital TV, easy access to the Internet and the availability of cheap mass storage there is a huge demand for efficient indexing and content-based retrieval on large video collections. 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. Key frames present the most salient features of the shot eliminating any redundancy. In this paper edge change ratio algorithm is used for detecting shots of the video and key frames are extracted from these shots. The extracted key frames represent the video as a whole and can be used in variety of applications such as content based video retrieval, video summarization etc.

Keywords: key frame extraction, edge change ratio, shot segmentation.

I. INTRODUCTION

Key frame extraction is the basis of content-based video retrieval. It is an essential part in video analysis and management, which can be used for suitable video summarization, video indexing, browsing as well as retrieval. The use of key frames reduces the amount of redundant data present in a video. Many algorithms for key frame extraction have been proposed with high computational complexity and processing inefficiency. Key frame is the frame which can represent the salient content and information of the shot. The key frames extracted must summarize the characteristics of the video, and hence the image characteristics of a video can be tracked by all the key frames in time sequence. It is necessary to discard the frames with repetitive or redundant information during the extraction. Some of the features that are used for key frame extraction are: color histogram, edges, shapes, optical flow, MPEG motion descriptions, MPEG discrete cosine co-efficient, motion vectors, camera activity, Motion-presenting actions or events in video, triangle model of perceived motion energy (PME) etc. A shot represents a sequence of frames captured from a unique and continuous record from a camera. Shot segmentation mainly refers to detecting the transition between successive shots.

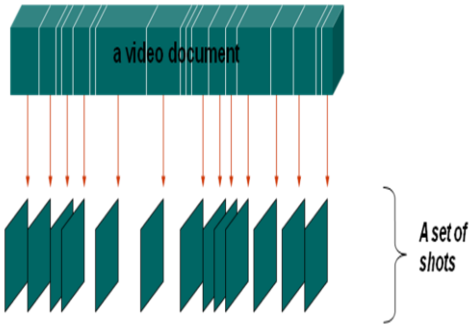

Fig1. Shot Segmentation

Figure 1 shows the segmentation of video into shots. The first step towards key frame extraction regardless of the algorithm followed is to segment the videos into temporal "shots," each representing an event or continuous sequence of actions. The detection methods can be an abrupt transition detection or gradual transition detection [2]. The commonly used methods of shot transition detection are: pixel-based comparison, template matching and histogram-based method [3]. The pixel-based methods are highly sensitive to motion of objects. So it is suitable to detect obvious segmentation transition of the camera and object movement. Template matching is apt to result in error detection if you only simply use this method. In contrast to pixel-based methods, the Histogram-based methods completely lose the location information of pixels. Consequently, two images with similar histograms may have completely different content. An algorithm of key frame extraction from compressed video data is presented in paper [4]. The features of compressed data are analyzed to obtain the key frames. Firstly, let the first frame as a key frame, and the ratios are calculated according to equation (1) and equation (2).

$R_P = no\_com / com$ (1)

$R_B = back / forw$ (2)

Where no_com denotes the number of macro blocks without motion compensation, and com stands for the number of macro blocks after motion compensation. When R_p peak appears, the P frame can be selected as a key frame. In (2) back is the number of the backward motion vectors and forw denotes the number of the forward motion vectors. The frame where a ratio peak occurs is extracted as a key frame. If there are no transitions in a shot, the frames in the shot have high similarity, and there is no significant change among the characteristic curves. Then the first frame can be extracted as a key frame, and finally the key frames of the video can be obtained. It has

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 11, November 2013

high computational complexity. Edge change ratio algorithm can be used to detect the shots of the video and is found to be efficient.

Pre-processing: The frames of the video need to be processed before ECR algorithm can be applied. It involves converting the color frame into gray scale frame and then gray frame into an edge frame. All the edges of the gray image can be detected using the sober method. This method detects an edge where the gradient I is maximum. After detecting the edges in the image they are filled to enhance the input image. In order to get an edge frame with black pixels on edges the image can be inverted. Image dilation is performed in the input binary image as with the dilated image it is easy to operate on subsequent modules to calculate the entering and exiting edge detection.

II. EDGE CHANGE RATIO

A feature that is proved to be useful in detecting shot boundaries is edges. Zabih et al. proposed the approach of ECR [6]. Temporal visual discontinuity usually comes along with structural discontinuity. i.e the edges of the objects in the last frame before the hard cut cannot be found in the first frame after the hard cut, and the edges of objects in the first frame after the hard cut in turn usually cannot be found in the last frame before the hard cut. The basic idea of edge change ratio method is summarized as follows:

- Detect edges in two contiguous frames fn and fn+1 respectively.
- Count the number of edge pixels in frame fn and fn+1.
- Define the entering and exiting edge pixels En+1 and En.

Suppose there are two images lm(n) and lm(n+1), the entering edge pixels En+1 are the fractions of edge pixels in lm(n+1) which are farther than a fixed distance 'r' away from the closest edge pixel in lm(n). Similarly, the exiting edge pixels En are the fraction of edge pixels in lm(n+1) which are further than r away from the closest edge pixel in lm(n+1). Then the edge change ratio between the frames fn and fn+1 can be computed as

$ECR(n, n+1) = max(En+1in/ \delta n+1, Einout/ \delta n)$

If the edge change ratio is larger than the predefined threshold it is considered as a cut between the frames. This is repeated for all the frames in the video and the hard cuts can be detected.

Adaptive threshold: Choosing a right threshold is an important problem in any algorithm such as color histogram comparison or edge change ratio algorithm. It is impossible to find single global threshold that works with all kinds of video materials. Hence a constant (global) threshold is avoided. Instead, an adaptive threshold is a better alternative to enhance the detection precision. It uses the local thresholds of the feature or similarity function to be compared which in the above mentioned cases are histogram similarity and ECR respectively. This threshold determines whether a shot transition has occurred. Once the shots of the video are detected correctly, the next step is to extract key frames of each shot.

III. Key frame extraction

Key frames are those frames of the video that contain the salient features and represent the entire video.

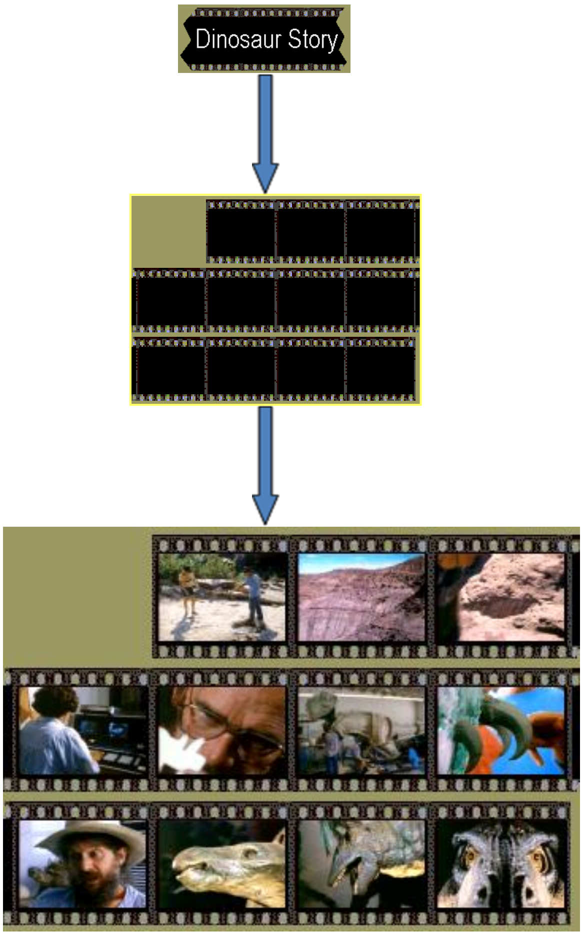

Fig 2: An example of key frames sequence of a dinosaur video

As shown in figure 2, the entire video of dinosaur story can be summarized using only those frames which are very different from other frames of the same shot.

After the hard cuts are detected in a given video the first frame of each shot can be taken as a key frame. This frame is the starting frame of the shot and hence is very different from all previous frames and may also contains the most salient features of the current shot. Along with this, the middle and the last frame of each shot are considered as candidate key frames. These candidate frames are again compared among each other for their correlation coefficients. From these frames the redundant ones are eliminated and the one which are unique are selected as key frames. In case, if all the three candidate frames are not correlated closely then all the three are considered as key frames. This extraction of key frames plays a very important role in content based retrieval applications as it reduces the size of the video greatly and is suitable for video summarization. Compression techniques exploit these features, reduces the amount of data that has to be processed thereby reducing the transfer stress on networks.

ISSN (Print) : 2319-5940
ISSN (Online) : 2278-1021

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 2, Issue 11, November 2013

IV. CONCLUSION

The key frames are extracted after shot switches are detected in the video. ECR method detects the edge changes in the frames and adaptive threshold value is used for edge changes in the frames. Threshold value is dynamic and will be changing during the course of each shot so that key frames can be retrieved efficiently. The extracted key frames can be used to represent the video as a whole and summarizes the salient content of the video. The method is of good feasibility, high efficiency, and high robustness. It avoids any processing inefficiency and computational complexity with low error and high robustness. Shot change using Edge change ratio was found to be efficient as the key frames can be extracted based on global adaptive threshold.

ACKNOWLEDGMENT

Authors of the paper are thankful to department of Computer Science & Engineering of R V College of Engineering for providing the needful infrastructure to carry out this work.

REFERENCES

[1] Weiming Hu, Senior Member, IEEE, Nianhua Xie, Li Li, Xianglin Zeng, and Stephen Maybank, "A Survey on Visual Content-Based Video Indexing and Retrieval ", IEEE Transactions on Systems, Man, and Cybernetic: Applications and Reviews, VOL. 41, NO. 6, November 2011. pp no 797-812.

[2] Ullas Gargi, Rangachar Kasturi, and Susan H. Strayer, "Performance Characterization of Video-Shot-Change Detection Methods", IEEE Transactions On Circuits And Systems For Video Technology, Vol. 10, No. 1, February 2000.

[3] T. Liu, H. Zhang, and F. Qi, "A novel video key frame-extraction algorithm based on perceived motion energy model," IEEE Transactions On Circuits And Systems For Video Technology, vol. 13, no. 10, pp. 1006-1013, 2003.

[4] Guozhu Liu, and Junming Zhao, "Key Frame Extraction from MPEG Video Stream", Proceedings of the Second Symposium International Computer Science and Computational Technology(ISCSCT '09) Huangshan, P. R. China, 26-28,Dec. 2009, pp. 007-011.

[5] Tianming Liu ; Hong-Jiang Zhang ; Feihu Q, "A novel video key-frame-extraction algorithm based on perceived motion energy model ",IEEE Transactions on Circuits and Systems for Video Technology, vol 13, issue 10, Oct 2003.

[6] R Zabih, J Miller, K Mai, "A featured based algorithm for detecting and classifying scne breaks" , ACM Multimedia 05, San Francisco, pp 189-200.

[7] Janko Calic, Ebroul Izquierdo, "Efficient key frame extraction and video analysis", Multimedia and Vision Research Lab, Queen Mary, University of London.

[8] Costas Cotsaces, Nikos Nikolaidis, and Ioannis Pitas, "Video shot detection and condensed representation: a review," IEEE Signal Processing, vol. 23, no. 2, pp. 28-37, 2006.

[9] S. M M Tahgoghi, Hugh E Williams, James A Thom, and Time Walker, "Video cut deection using frame windows", Proceedings of the Twenty-eight Australian Conference on Computer Science, NewCastle, Austraila, Jan 2006, pp 193-199.

[10] Sujatha C, Mudenagudi, U. "A Study on Keyframe Extraction Methods for Video Summary", IEEE Conference on Computational Intelligence and Communication Networks (CICN), 7-9 October 2011,Gwalior, India.

[11] Luo, C. Papin, and K. Costello, "Towards extracting semantically meaningful key frames from personal video clips: From humans to computers", IEEE Trans. Circuits Syst. Video Technol. , vol. 19, no. 2, pp. 289-301, Feb. 2009.

[12] Vasileios Chasanis, Aristidis Likas and Nikolaos Galatsanos,"Video rushes summarization using spectral clustering and sequence alignment", Proceedings of the 2nd ACM TRECVid Video Summarization Workshop, 2008.

[13] Jafarpour S, Cevher V, Schapire R.E, "A game theoretic approach to expander-based compressive sensing", IEEE International Symposium on Information Theory Proceedings (ISIT) , August 2011, page(s):464-46.