ICIC Express Letters
Part B: Applications
Volume **7**, Number **12**, December **2016**
 ISSN 2185-2766
pp. **2713–2717**

BLOCK BASED APPROACH FOR KEY FRAME EXTRACTION ON LARGE VIDEO SEQUENCES

Mie Mie Tin[1], Nyein Nyein Myo[2], Mie Mie Khin[1] and Thi Thi Zin[3]

[1]Myanmar Institute of Information Technology
P.O.Box 05053, 73rd Street, Ngu Shwe War Street, Chanmyathazi Township
Mandalay 05053, Myanmar
{ miemietin1983; miemie.khin9 }@gmail.com

[2]Information Science Department
University of Computer Studies, Mandalay
Pathangyi, Mandalay, Myanmar
Vecky@gmail.com

[3]Electrical and System Engineering
Faculty of Engineering
University of Miyazaki
Kibana Campus, 1-1, Gakuen Kibanadai-nishi, Miyazaki-shi 889-2192, Japan
thithi1972@gmail.com

Received May 2016; accepted August 2016

Abstract. *Video segmentation and key frame extraction are very important in real world video systems. Key frames are essential to analyze on large amount of video frame sequences. This paper emphasizes surveillance video and the aim is to extract some meaningful key frames from long video sequences. This purpose is to reduce weak transaction frames in a large video stream by using block base algorithm and cluster the transition video shots. Kullback-Leible divergence method is used in key frame extraction for strong transition video shot. For weak transition video shot, the system will find three candidate key frames and they are compared. Key frames are meaningful frames for video sequences. These frames which represent video streams can be analyzed. Duplicated key frames from the video stream are analyzed in order to be extracted from different shots. Finally, key frames have many assets such as stability, accuracy, and summarize information for a large video.*
Keywords: Video retrieval, Key frame extraction, Threshold value, Block base method, Information theory

1. **Introduction.** Key frame extraction has been recognized as one of the important research issues in video information retrieval [1]. There are some situations in which video information retrieval is needed at homes, especially for old people and babies to be watched over. Some places are needed to watch for security. Some patients at hospital need to watch out every time. However, we cannot serve all these services in every time for every person and every place. This system supports those services to make good decision by using key frames in long video sequences. This system will support patient care system, home security system, old-person and babies care system and so on.

This paper discusses the importance of key frame selection and proposes a block base approach for search transition frame. Introduce a new block based algorithm for finding video shots and key frame extraction based on unsupervised clustering [2]. The proposed algorithm is both computationally simple and able to solve for long frame sequence.

Objective of this system is to remove the redundant frames and select significant key frame from long video sequence. Key frame extraction is the fundamental step in any

of the video retrieval applications [3]. Key frame is useful for analysis in big video sequence. This system selected blocks for similarity measurement and reduced the number of stable frames in comparisons. This technique can reduce amount of comparison time and computational cost for all frames. In a video object segmentation system, tracking regions has many fundamental advantages based on tracking points or joint clustering of all pixels from all video frames. Circular dynamic-time warping method is capable of optimally matching closed region contours with linear complexity [8].

The process to classify video shots needs to choose candidate key frames. In this system, first frame will be chosen as candidate key frame and that frame is compared with other frames. These shots are classified as strong shots or weak shots. Two methods are used to extract key frame of these strong shots and weak shots.

Adaptive key frames allow users to quickly browse meaningful information about the video by viewing only a few highlighted frames [4,9]. Final analysis will take all key frames for their differences and it will reduce some similar key frames.

KLD has been demonstrated to perform better than ED, by shot-based computational mechanism using short vides stream [5]. The performance of simple edge change ratio (ECR) method, color histogram method and distance based methods can be used to extract key frames [6]. Various key frame extraction techniques using many parameters were analyzed and surveyed along with researcher's merits and demerits used in current video retrieval applications [7].

This paper is organized as follows. Problem statement is presented in Section 2. In this section, describe the block base technique and discuss the cluster technique for the video shot. Section 3 presents expected results of our research. Our research is ongoing stage. This paper is concluded in Section 4.

2. **Problem Statement and Preliminaries.** Key frame extraction is very important in video processing. Shot boundary-based approaches select a key frame from a fixed position in the scene or several frames separated by a fixed distance. The visual content based approach uses multiple visual criteria to extract key frames. The motion analysis-based approach selects key frames at local minima of motion. The shot activity-based approaches, the local minima are selected based on the activity curve as the key frames.

Pixel-based difference: this method compares pixel difference between two consecutive frames in video sequences. It is very sensitive to camera motion.

Statistical based difference: this method divides all frames in video sequence into small regions and then compares some properties of every pixel between successive frames using statistical computation parameters.

Transform-based difference: this method represents compression difference computation with transformation methods.

Histogram-based difference: this method computes the histogram of each frame and compares it to detect shot boundaries. It is based on image color.

Edge based difference: in this method base edges detection method on each frame and then the edge pixels are paired with nearby edge pixels in consecutive frame [9].

Block based segmentation method is used to extract object in image [10,11]. There is one altered point of our approach.

In this system, we consider to control long video data and to extract complete and accurate key frames from these video sequences.

In the first step, long video data is divided to shot sequences based on time duration.

In the second step, choose first frame in video sequences as candidate key frame and identify $8*8$ blocks on candidate key frame and consequence frames. Compute a diagonal pixel based detection method for these blocks and find differences between these two blocks in consequence frame. It can get the distance between these two frames and compute this logic process until reaching transition frame and get shots.

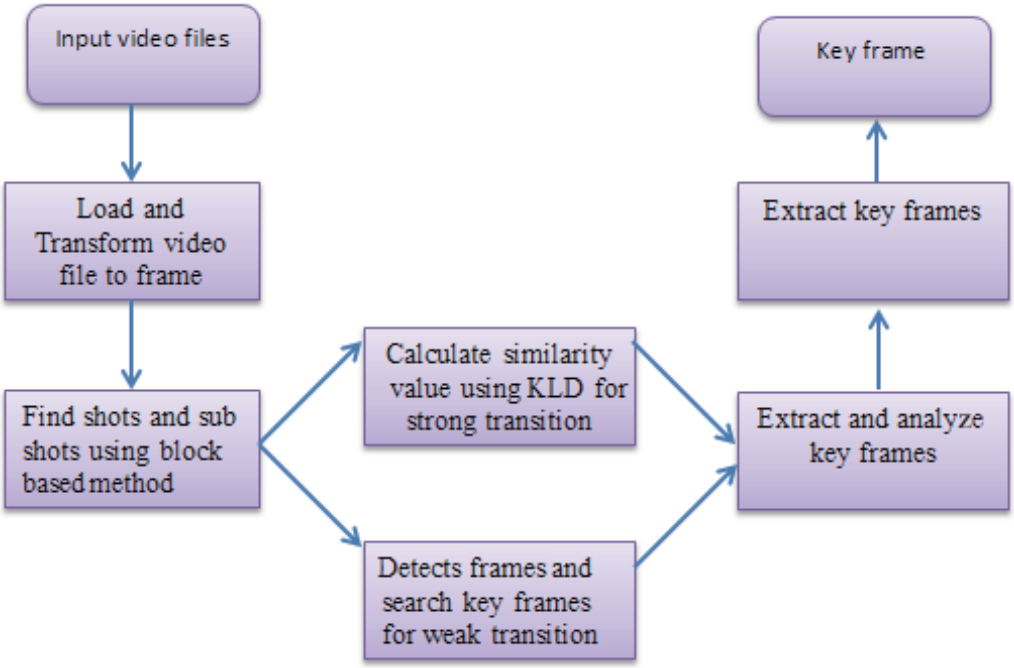

FIGURE 1. System design

Finally, we can cluster stable shots and transition shots in video sequences. This block based method just needs to detect half pixels of each frame and can reduce detection times for all video sequences.

We have two types of video sequence frames sets, one is weak transition frames sets and the other is strong transition frames sets or object moving frames sets. These frames need to analyze the frames differences and these values depend on threshold value.

In weak transition frames that are nearly stable frames that are included in frame sequence, choose the first frame as f_s, end of frame in sequence as f_e and f_m is middle of frame sequence. Find the difference between three frames and these different values are greater than or equal to threshold value, and choose f_m as key frame for that video shot. If these values are different from threshold value, use block based method repeatedly.

In strong transition frames sequences, the difference value between two frames is not stable. In this case, we use Kullback-Leible divergence (KLD) algorithm to find key frame for this sequence.

3. **Expected Results.** This paper proposes a key frame extraction on large video sequences by using block based method. Image sequence consecutive frames are analyzed on a block-by-block and analyze these pixels with diagonal method. We can detect half of pixels and miss four pixels rounded area. However, this area is very small and it can be omitted for decision that frame is transition or not. By detecting all frames with this method, system can search transition frame and can classify all shots. When we obtain sub shots, if it is a weak transition shot, find three candidate key frames and measure differences to decide these shots are transition or not.

If these candidate key frame differences $\delta 1$ and $\delta 2$ are greater than or equal to predefined threshold value, the middle frame can be extracted as key frame for that video shot.

If the differences between $\delta 1$ or $\delta 2$ are less than predefined threshold value, we will divide these frames sets as sub shots and will use three candidate keys frames technique until key frames are extracted.

For transaction frames such as object movie frames shot, some steps from Kullback-Leible divergence (KLD) methods can be used. It is needed to calculate to find the

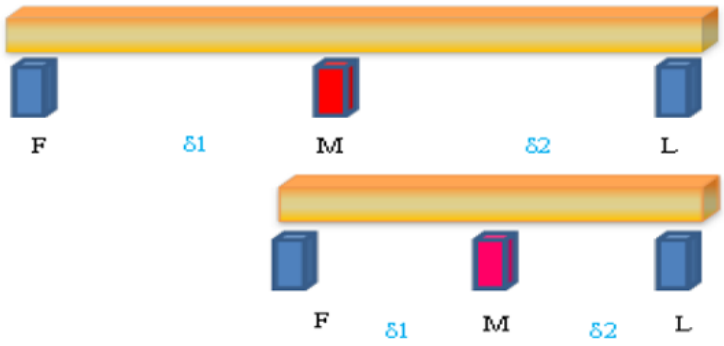

FIGURE 2. Three candidate key frames

differences of all these frames and need to find effective and efficient information in video sequences.

The Kullback-Leibler divergence of Q from P is defined to be

$$D_{KL}(P \parallel Q) = \sum_i P(i) \log \frac{P(i)}{Q(i)} \tag{1}$$

So we can easily neglect some of the shots including stable frame and can calculate for strong and weak transition frame for key frame extraction. For strong transition frames, we can find key frames based on KLD approach [5,12].

$$KLD(f_{i-1}, f_i) = \sum_{j=0}^{n} P_{f_{i-1}}(j) \log \left(\frac{P_{f_{i-1}}(j)}{P_{f_i}(j)} \right) \tag{2}$$

$P_{f_{i-1}} = \{P_{f_{i-1}}(1), P_{f_{i-1}}(2), \ldots, P_{f_{i-1}}(n)\}$ and $P_{f_i} = \{P_{f_i}(1), P_{f_i}(2), \ldots, P_{f_i}(n)\}$ are respectively probability density functions.

To find key frames in moving object shots threshold value (δ^*) is used to find frames difference.

$$\frac{KLD(f_{i-1}, f_i)}{KLD_u(f_{i-1}, f_i)} \geq \delta^* \tag{3}$$

Finally, many key frames represented many video shots. Some key frames are similar to different video shots. It is needed to analyze and to reduce these similar frames from key frame group. So it needs to calculate similarity between these key frames and reduce these frames. It supports to get accurate key frame from long video sequences. Sample data sets for experiments are as shown in Table 1 and Figure 3. Our research is ongoing stage and testing with some datasets to get weak transaction shots and strong transaction shots. Finally, we can find stable and accurate key frames from these shots.

4. **Conclusions.** This paper aims to get best key frame from surveillance long video data. This result can be useful to image's key information by using keys frame extraction in image information retrieval systems. It is found that a new technique that emphasized the best key frame from large video data sets. We use two difference methods to find frames differences. KLD technique and three candidate key frames technique are used to

TABLE 1. Testing data sets

No	File name	Length (s)	Size (M)	Frames
1	CAMOUFLAGE	12	32.6	352
2	WAVING TREES	9	30.7	286
3	MOVEDOBJECT	57	173	1744

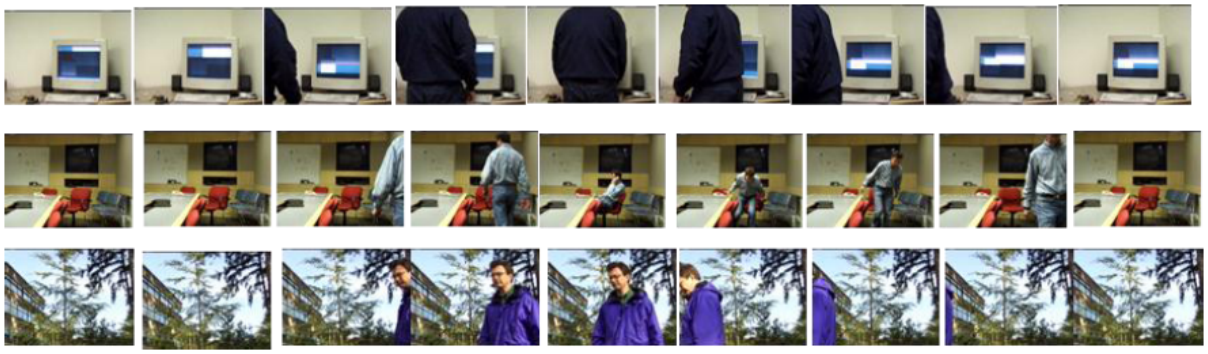

FIGURE 3. Testing data sets

extract key frames from video shots. Future work is to find and extract key frames in long video sequences for complex objects and to analyze these objects.

Acknowledgment. This work is partially supported by University of Computer Studies, Mandalay. We thank a lot to rector for supporting this research. We also gratefully acknowledge the helpful comments and suggestions for our paper of the reviewers from ICICIC2006 International Conference because they supported to improve this research.

REFERENCES

[1] Y. Zhuang, Y. Rui, T. S. Huang and S. Mehrotra, Adaptive key frame extraction using unsupervised clustering, *International Conference on Image Processing*, 1998.

[2] S. Yang and X. Lin, Key frame extraction using unsupervised clustering based on a statistical model, *Tsinghua Science and Technology*, pp.169-173, 2005.

[3] Y. H. Sharath Kumar, N. Manohar and H. K. Chethan, Animal classification system: A block based approach, *International Conference on Advanced Computing Technologies and Applications*, pp.336-342, 2015.

[4] S. Nikita and M. Ravi, A survey based on video shot boundary detection techniques, *International Journal of Advanced Research in Computer and Communication Engineering*, vol.3, no.4, pp.6279-6282, 2014.

[5] L. Li, Q. Xu, X. Luo and S. Sun, Key frame selection based on KL-divergence, *IEEE International Conference on Multimedia Big Data*, pp.337-341, 2015.

[6] S. A. Begum and A. Askarunisa, Performance analysis of various key frame extraction methods for surveillance applications, *International Journal of Emerging Technology and Advanced Engineering*, 2014.

[7] A. Nasreen and G. Shobha, Key frame extraction from videos – A survey, *International Journal of Computer Science & Communication Networks*, pp.194-198, 2014.

[8] W. Brendel and S. Todorovic, Video object segmentation by tracking regions, *IEEE International Conference on Computer Vision*, pp.833-840, 2009.

[9] C. T. Dang and M. Kumar, Key frame extraction from consumer videos using epitome, *The 19th IEEE International Conference on Image Processing*, pp.93-96, 2012.

[10] E. R. Vimina and K. P. Jacob, A sub-block based image retrieval using modified integrated region matching, *International Journal of Computer Science Issues*, vol.10, no.1, pp.686-692, 2013.

[11] V. Reddy, C. Sanderson and B. C. Lovell, Improved foreground detection via block-based classifier cascade with probabilistic decision integration, *IEEE Trans. Circuits and Systems for Video Technology*, vol.23, no.1, pp.83-93, 2013.

[12] C. Lang, D. Xu, W. Cheng and S. Feng, Automatic key-frames extraction to represent a video, *The 7th International Conference on Signal Processing*, 2014.