Effective Shot-based Keyframe Selection by using Image Quality Evaluation

Jong Min Lee, Hyeon Seok Yang, and Young Shik Moon†
Department of Computer Science and Engineering
Hanyang University
55 Hanyangdaehak-ro, Sangnok-gu, Ansan-si, Gyeonggi-do, Korea
{jmlee, hsyang}@visionlab.or.kr, ysmoon@hanyang.ac.kr

Seung-Hee Kim and Sun-Joong Kim
Broadcasting & Telecommunications Media Research Lab.
Electronics and Telecommunications Research Institute
218 Gajeong-ro, Yuseong-gu, Daejeon, Korea
{seung, kimsj}@etri.re.kr

***Abstract*— In this paper, we propose a fast and effective method for shot-based keyframe selection. The proposed method is based a visual image quality evaluation using image contrast and motion blur measures. The main purpose of this paper is to provide good keyframes to viewers fast and effectively. Therefore, we adopt fast and effective approaches instead of heavy time-consuming and complex approaches. Experimental results show that the proposed method fast and effectively selects a good keyframe in each video shot.**

Keywords— keyframe selection; image evaluation; motion; image contrast;

I. Introduction

Recently, as the development of techniques of digital video shooting equipment and the increment of personal content providers, digital video contents are extremely increased. In addition, viewers demand that the summarized contents are served as soon as possible after the broadcast ends. In order to provide main contents of the videos for viewers who need an outline of the video without watching the entire video, fast and effective keyframe extraction algorithms are necessary. Keyframe extraction is one of the famous conventional video abstractions. In spite of its drawback that is difficult to summarize whole sequence in one frame, it is widely used in various fields. In the early days, the first frame or the center frame of each shot is used as a keyframe because of the processing time. Then, a large amount of approaches are proposed using various concepts, such as low level features [1], object motion [2], salient objects [3], frame clustering scheme [4], entropy [5], etc [6]. However, these methods are not suitable for purpose of our application because of low quality or a large amount of processing time.

Therefore, in this paper, we propose a fast and effective keyframe selection method using our image quality measures to provide a keyframe as an outline of a shot for viewers. The rest of this paper is organized as follows. Section 2 describes the proposed method. Section 3 shows the experimental results and we conclude this paper in Section 4.

† Corresponding author

978-1-5386-2615-3/18/$31.00 ©2018 IEEE

II. Proposed Method

Due to select a good keyframe fast and effectively, we propose a novel shot-based keyframe selection scheme by using visual image quality measures. First, the proposed method adopts simple keyframe extraction algorithms to find an initial keyframe in a shot. Next, the proposed method evaluates its image quality by using our image quality measures for image contrast and motion blur. If the image quality is lower than the threshold values, the proposed method newly selects the best quality frame in the same shot. Fig. 1 shows the flowchart of the proposed method.

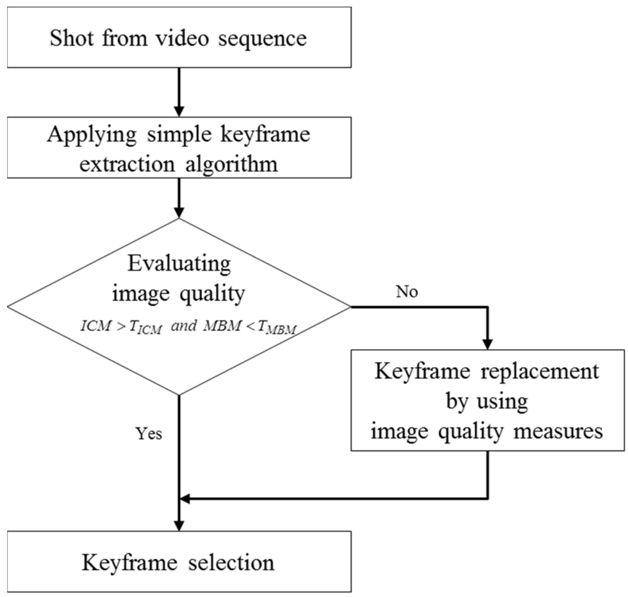

Fig. 1. Flowchart of the proposed method.

A. Image Quality Evaluation

We adopt image quality measure to select a good keyframe in each shot. The measure consists most important two factors that are image contrast and motion blur because viewers are sensitive to contrast and blur. We define two measures; *ICM* and MBM for image quality assessment. ICM is image contrast measure and MBM is motion blur measure.

1) Image Contrast Measure

In this paper, in order to measure for image quality of image contrast, we propose an image contrast measure; ICM. Equation (1) represents ICM.

$$ICM = \frac{I_{\max} - I_{\min}}{256} \quad (1)$$

Where, $I_{\max}$ is the 5% upper boundary value of the histogram, $I_{\min}$ is the 5% lower boundary value of the histogram, 256 is the maximum range of the histogram. We consider how wide the range except both 5% upper and 5% lower boundary value is. The value of ICM is in the interval [0, 1] and a higher value is better.

2) Motion Blur Measure

In order to measure for image quality of motion blur, we propose a motion blur measure; MBM. Equation (2) represents MBM.

$$MBM = \frac{\frac{1}{N}\sum_{1}^{N} motion\ value}{maximum\ motion\ value} \quad (2)$$

Where, N is the number of keypoints to calculate motion between a current frame and a previous frame, $motion\ value$ is a mount of motion from each selected pixel to calculate motion, and $maximum\ motion\ value$ is the theoretical maximum value of motion in a frame. In this paper, we determine that the $maximum\ motion\ value$ is the diagonal distance of a frame. We consider how much the average movement is. The value of MBM is in the interval [0, 1] and a lower value is better.

B. Keyframe Replacement Strategies

In order to extract initial keyframes from each shot, we apply simple keyframe extraction algorithms, such as the first frame extraction algorithm in a shot, the center frame extraction algorithm in a shot, etc. Then we evaluate image quality of the initial keframes by using proposed image quality measures; ICM and MBM. Experimentally, in this paper, we adopt 0.25 as the T_{ICM} and 0.01 as the T_{MBM}.

If the both value are satisfied by threshold values, the initial keyframe is determined as a final keyframe of the shot. Otherwise, if the value of ICM is lower than threshold value; T_{ICM}, we replace a new keyframe which has the highest value of ICM and lower MBM value than T_{MBM} among all frames from a same shot. Likewise, if the value of ICM is higher than threshold value; T_{ICM} and the value of MBM is higher than threshold value; T_{MBM}, we replace a new keyframe which has the highest value of ICM and lower MBM value than T_{MBM} among all frames from a same shot. Because viewers are more sensitive to image contrast than motion blur, we give priority to image contrast over motion blur.

III. Experimental Result

In order to evaluate the performance of the proposed method, we apply to shots from the video sequence. The experimental video consists of 492 shots, which has a keyframe. In the video, the propose method replaced 63 keyframes from the first frame extraction algorithm. Among the 63 keyframes, 60 keyframes have been improved, but 3 keyframes have not been improved. Also, the propose method replaced 55 keyframes from the center frame extraction algorithm. Among the 55 keyframes, 50 keyframes have been improved, but 5 keyframes have not been improved. It takes about three minutes to select keyframe for 60 minutes length video by using our proposed method. It means we are able to serve the keyframes to viewers as soon as possible after the broadcast ends.

Fig. 2 shows some of the experimental results. Fig. 2 (a) and (b) are results of conventional keyframe selection algorithms, while (c) is a newly selected keyframe by the proposed method in the shot.

(a) The 1st frames of the shot as keyframes (b) The center frames of the shot as keyframes (c) Replaced keyframes by the proposed method

Fig. 2. Experimental Results.

IV. Conclusion

In this paper, we propose the method to select a good keyframe in each video shot fast and effectively. To this end, keyframes extracted by using simple keyframe extraction methods are evaluated by the proposed image quality measure. That is consists of image contrast and motion blur measures. As experimental results, the proposed method evaluates the image quality appropriately and selects a good keyframe in each shot effectively. However, because the proposed method focuses on speed, some of the replaced keyframes are unsatisfactory in terms of content, even if the image quality is improved. So our future research will focus on both content and speed.

ACKNOWLEDGMENT

This work was supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government(MSIT) (2015-0-00219, Development of smart broadcast service platform based on semantic cluster to build an open-media ecosystem).

REFERENCES

[1] Y. Li, T. Zhang, and D. Tretter, "An overview of video abstraction techniques," Technical Report HPL-2001-191, HP Laboratory, 2001.

[2] W. Wolf, "Key frame selection by motion analysis," IEEE International Conference on Acoustics, Speech, and Signal Processing, pp. 1228-1231, 1996.

[3] S. Marat, M. Guironnet, and D. Pellerin, "Video Summarization Using a Visual Attention Model," 15th European Signal Processing Conference, pp. 1784-1788, 2007.

[4] Y. Zhuang, Y. Rui, T. S. Huang, and S. Mehrotra, "Adaptive key frame extraction using unsupervised clustering," IEEE International Conference on Image Processing, pp. 866-870, 1998.

[5] Q. Xu, Y. Liu, X. Li, Z. Yang, J. Wang, M. Sbert, and R. Scopigno, "Browsing and exploration of video sequences: A new scheme for key frame extraction and 3D visualization using entropy based Jensen divergence," Information Sciences , Vol. 278, pp. 736-756, 2014.

[6] B. T. Truong and S. Venkatesh, "Video abstraction: A systematic review and classification," ACM Transactions on Multimedia Computing, Communications, and Applications, Vol. 3, No. 1, pp. 1-37, Feb. 2007.