Extraction of Epitome Key Frames from Video Frames

Mie Mie Tin
Department of Information Science Technology University (Yadanapon Cyber City)
Pyin Oo Lwin, Myanmar
miemietin1983@gmail.com

Khant Khant Win Tint
Department of Information Science Technology University (Yadanapon Cyber City)
Pyin Oo Lwin, Myanmar
khantkhantwintint2k17@gmail.com

Mie Mie Khin
Head of Admin Department
University of Computer Studies, Yangon
Yangon. Myanmar
miemiekhin@uscy.edu.mm

***Abstract*—This system considers surveillance video and processes to compress the video fames sequence size by the extraction of objects including frames from a section of that video. It extracts the objects that include frames from the video sequence, and this object refers to the relevant video with key information frames. To extract key information frames, the system considers all pixel values in a frame and calculates the frame differences. The system searches for the significant transition frames using the image HSV value. The analysis results for Compression Ratio (CR) and Computational Efficiency (CE) are good and can compress more than 90% using epitome keyframes.**

***Keywords*—HSV color, Shots. Euclidian Distance, Candidate key**

I. INTRODUCTION

A surveillance video has a large number of continuous frames and will generate large amounts of data for handling. It is very difficult to control to get an excellent result. Big data analysis supports the issues of social media data analytics, content-based analytics, text data analytics, video and audio data analytics. This paper supports the compression of a huge number of video frames into useful and efficient frame sequence lengths. To compress the length of the surveillance video, the frame sequences are classified into two types. The system neglects all non-transition frames as noise frames. To define transition frames, the system already takes the background-view frame for that surveillance video and stores it as a candidate keyframe. It supports extracting transition frames from the long video sequence frames.

II. RELATED WORKS

Aiswarya and Smitha [1] presented a review of a surveillance video summary of a domain-dependent system using the high-level features-based domain-adaptive video summary framework. Buczkowska et al. [2] presented a probabilistic mixture of hurdle poisson models. In the development of a probabilistic mix of hurdle-Poisson models for multiple industries of operation, distance measurements are regarded as two road distances with and without congestion, public transit distance, and the relevant travel times. Bendraou [3] has presented video shot boundary detection and key-frame extraction using mathematical models.

Srivastava, et al. [4] presented a hand gesture recognition system. It is based on the feasibility of principal component analysis. The eigenspace is used by processing the Eigenvalues and Eigenvectors of the images in the training data set. This approach used the principal component analysis (PCA) technique for the analysis.

III. PROPOSED APPROACH

The surveillance video files act as input files, and these files are converted to continuous frames sequences. These sequences are classified into two groups. The first group is a few transition frames, and these frames like the candidate frame. The second group objects that includes frames, and these frames are significant transition frames. These significant transition frames represent that video. All of the not-transition frames are excluded from the frame Department of Information Science sequence.

A. Processing steps

To compress a video sequence, it needs to omit some noise frames. This noise frame does not include any objects in vision. To find these frames, frame comparing techniques will be supported. The system uses Color Hue value, Saturation value, and value for each pixel in frames. To find a difference in a frame, consider the total value of HSV.

To extract significant frames in video sequences has three steps. The first step is to convert video to a continuous video frame sequence. The second step is to compare candidate keyframes and frames of sequences to find the frame difference value and transition frames. These continuous transition frame sequences are organized as a shot. Finally, it extracts epitome key frames from these shots. These epitome keyframes represent the relevant video sequence. The system selects a significant transition frame, by finding the difference value. between the candidate keyframe and other continuous frames. This frame value difference is less than the threshold value, it will become a non-transition frame. If the frame value difference value is greater than the threshold value, it will be defined as significant transition frames. The mean value of all frame value differences is defined as the threshold value.

$$\text{Threshold value T=} \quad \frac{1}{n}\sum_{1}^{n} f(x)dx \qquad (1)$$

n = total frames

The extraction of the epitome keyframe it has object information and this information represents that video. To find frame difference values, this paper uses the Euclidean distance method on each pixel value.

$$\text{Euclidean Distance of (x, y)} = \sqrt{\sum_{i=1}^{n}(x_{i-}\ y_i)^2} \qquad (2)$$

x_i = the value of point x from candidate key frame
y_i = the value of point y from other frame
n= pixel amount in a frame

To compare in a frame, the system considers hue value with eight bins, and saturation and value with four each. The frame difference value is considered as a combination of these three different values in a pixel. For frame value differences, the system considers the total value of all pixels in a frame.

The dimensional representation of the HSV color space is hexacore and the vertical axis represents the image intensity. The angles in the range of 0 to 28 are represented with Hue and the angles 0 and 2S start with red color, green color at 2S/3. To compare pixel values between the candidate frame and other frames, the system considers the HSV color value in each pixel and uses different bin types.

IV. EXPERIMENTAL RESULT

Key frame extraction is one type of video compression for the large video sequences. This research supports video compression with epitome keyframes.

A. *Significant Frames Sequence*

The non-transition frames are extracted and are omitted from the frame sequence. The continuous significant frames are organized as shots. For a video sequence, have many shots with different shot lengths. The shot length is the number of frames in a shot, and the system neglects the five frame continuous shot length. The simple significant frame sequences are shown in Table 1.

TABLE.1 SIGNIFICANT FRAMES SEQUENCE

B. *Key Frames*

The keyframes are extracted from significant transition frames. These frames are organized into groups and called the "shot". The epitome keyframes are extracted based on the three most significant transition values from a shot. All of these frames represent the video storage system, video retrieves system and analysis system.

The experiment results with epitome keyframes from the relevant video are presented in Table 2. These keyframes are extracted with maximizing value frame differences and the object include rate is the highest. These keyframes represent o the relevant video.

TABLE 2. EPITOME KEY FRAME FROM DIFFERENCE DATASET

C. *Evaluation*

The compression rate supported the video transfer system, video extraction system and so on. In keyframes extraction processing, calculation of compression ratio needs to estimate the compression rate.

The compression ratio (CR) is defined by:

$$CR= 1- (\eta_{TSF} / \eta_{TNF}) \quad (3)$$

Where η_{TSF} = total no. of significant frames
η_{TNF} = total no. of frames in sequence

The compression rates are based on the significant number of transition frames. This amount become relevant important frames and are calculated the compression rate based on the total frame sequence.

TABLE 3. COMPRESS RATE WITH SIGNIFICANT TRANSITION FRAMES

No	Dataset Name	Frames Length	Significant Transition frames	Compression Ratio
1.	BOOTSTRAP	3055	1562	48.8707
2.	COMOFLUGE	352	38	89.2046
….	……………	……	……	……….

The compression rates with epitome keyframes give excellent results. The maximum compression rate is more than 99.6% and the minimal compression rate is more than 92.7%. That result shows that the epitome key frame based on the compression method gives excellent results. These results are shown in Table 4.

TABLE 4. COMPRESS RATE WITH KEY FRAMES

No	Dataset Name	Key Frames	Compression Ratio	Computational Efficiency
1.	BOOTSTRAP	45	98.527	0.1692
2.	COMOFLUGE	3	99.1477	0.2599
…	………..	…..	…..	…..

To evaluate the compression, the system calculates the compression ratio based on keyframe. The keyframe extraction depends on the object that includes the frames amount, and the compression rate depends on the number of significant transition frames and keyframes. To compress video, the keyframe extraction method supports getting an excellent result.

The calculation of Computational Efficiency (CE) based on compression ratio and processing time.

V. CONCLUSION

Using the HSV value base compression method gives excellent compression results. The compression rate with significant transition frames depends on objects that include frames. The compression rates with epitome keyframes gave excellent results. The maximum compression rate is more than 99.6% and the minimal compression rate is more than 92.6%. The results show the epitome key frame base compression method gives excellent results and supports the video compression research areas.

REFERENCES

[1]. N.R. Aiswarya, and P.S. Smitha, "A review on domain adaptive video summarization algorithm" In Networks & Advances in Computational Technologies (Net ACT), IEEE International Conference on pp. 412-415. 2017.

[2]. S. Buczkowska, N. Coulombel and M. De Lapparent, "Euclidean versus network distance in business location: A probabilistic mixture of hurdle-Poisson models." 2016

[3]. Y. Bendraou, "Video shot boundary detection and key-frame extraction using mathematical models." Université du Littoral Côte d'Opale. November. 2017.

[4]. T. Srivastava, R.S. Singh, S. Kumar and and P. Chakraborty, "Feasibility of Principal Component Analysis in hand gesture recognition system" *arXiv preprint arXiv:1702.07371*. 2017