Available online at http://www.journalcra.com

International Journal of Current Research
Vol. 8, Issue, 01, pp.25277-25280, January, 2016

INTERNATIONAL JOURNAL OF CURRENT RESEARCH

ISSN: 0975-833X

RESEARCH ARTICLE

CCTV VIDEO ABSTRACTION AND OBJECT DETECTION FOR VIDEO SURVEILLANCE SYSTEM

*Hitesh Panchal

L. E. College (Polytechnic), Morbi, Gujarat, India

ARTICLE INFO

Article History:
Received 14th October, 2015
Received in revised form
05th November, 2015
Accepted 09th December, 2015
Published online 31st January, 2016

Key words:

CCTV video abstraction,
Object detection,
Key frame extraction,
Video surveillance system.

ABSTRACT

Millions of closed-circuit television (CCTV) cameras are installed in streets and businesses throughout the world with the stated goal of reducing crime and increasing public safety. This leads to large collection of video data. For searching particular object or person from whole CCTV video data is too much time consuming. The proposed algorithm extracts key frames from these videos to overlook videos. From these key frames required frames are easily detectable and video can be played from that frame to analyze further in details. Extracting small number of key frames that can abstract the content of video is very important for efficient browsing and detection. The amount of data in video processing is significantly reduced by using video segmentation and key-frame extraction. This paper presents algorithm for key frames extraction with matching difference between two consecutive frames is computed with different weights. Shot boundaries are detected with automatic threshold. Key frame is extracted by using reference frame-based approach.

Citation: **Hitesh Panchal, 2016.** "CCTV Video Abstraction and Object detection for Video Surveillance System", *International Journal of Current Research,* 8, (01), 25277-25280.

INTRODUCTION

CCTV (Closed-circuit Television) also known as video surveillance uses video cameras to transmit signal to specific place on limited set of monitors. CCTV is used for surveillance in areas that may need monitoring such as banks, railway stations, bus stations, shops, schools, airports etc. CCTV is spreading across the world specifically targeting public places. It is seen as a cheaper way to deter crime compared to increasing size of police departments. All these CCTV cameras collects large amount of video data. For analyzing or summarizing this video, a person requires watching whole video which is very time consuming. CCTV video abstraction and object detection techniques are needed to solve this difficulty (Fereshteh Falah Chamasemani, 2013). Shot boundary detection and key frame extraction are two bases for abstraction and summarization techniques. Researchers have actively developed different approaches for intelligent video management, including shot Transition detection, key frame extraction, video retrieval, etc (Hanjalic, 2002; Cernekova *et al.,* 2006; Babaguchi *et al.,* 2004; Hanjalic, 2003). Among these approaches, shot transition detection is the first step of content-based video analysis and key frame is a simple yet efficient form of video abstract. It can help users to understand the content at a glance which saves time (Cernekova *et al.,* 2006; Babaguchi *et al.,* 2004; Hanjalic, 2003).

Corresponding author: Hitesh Panchal
L. E. College (Polytechnic), Morbi, Gujarat, India

Many approaches used different kinds of features to detect shot boundary which includes histogram, shape information, motion activity. Among these approaches, histogram is the popular approach. But, in these histogram-based approaches, pixels' space distribution was neglected. Different frames may have the same histogram, In view of this, Cheng *et al.,* 2002 where each frame divided into *r* blocks, and the difference of the corresponding blocks of consecutive frames was computed by color histogram.

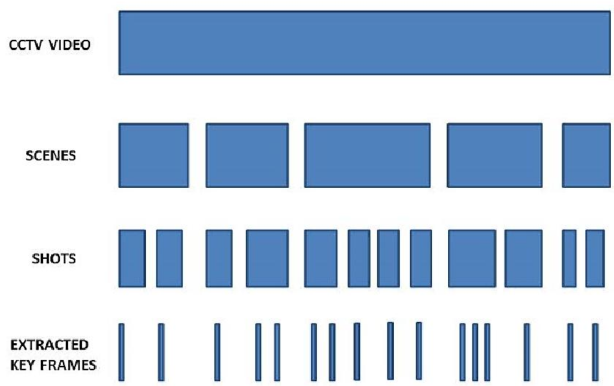

Fig. 1. Key frame Extraction from CCTV video

A shot is defined as the consecutive frames from the start to the end of recording in a camera. It shows a continuous action in an image sequence (Seung-Hoon Han, 2000). 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 boundaries show an abrupt change in image intensity or color, while those of fades or dissolves show gradual changes between frames.

Shot Boundary Detection

Image Segmentation

First, each frame is divided into nine blocks, B(1,1), B(1,2), B(1,3), B(2,1), B(2,2), B(2,3), B(3,1), B(3,2), B (3,3). Then the difference of the corresponding blocks between two consecutive frames is computed. Finally, the final difference of two frames is obtained by adding up all the differences through different weights. Different position's pixels have different contribution to shot boundary detection: pixels on the edge are more important than others. Thus, different weights are given to blocks of different position. Here more weights are assigned to corner blocks compared to other blocks.

Matching Difference

There are six kinds of histogram match (Cheng *et al.,* 2002). Color histogram was used in computing the matching difference in most literatures. However, through comparing several kinds of histogram matching methods, Nagasaka (He and Geng, 2003) reached on conclusion that x^2 histogram outperformed others in shot boundary Recognition. Hence, x^2 histogram matching method is referred in this paper.

Algorithms Description

Shot boundary detection (ZHAO Guang-sheng, 2008)

Let $F(k)$ be the k^{th} frame in video sequence, $k = 1, 2\ldots, Fv$ (Fv denotes the total number of frames in video).

$$D_B(k,k+1,i,j)=\sum_{l=0}^{L-1}\frac{[H(i,j,k)-H(i,j,k+1)]^2}{H(i,j,k)} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

Where, $H(i, j, k)$ and $H(i, j, k+1)$ stand for the histogram of blocks at (i, j) in the k^{th} and $(k+1)^{th}$ frame respectively and L is the number of gray in an image.

$$D(k,k+1)=\sum_{i=1}^{m}\sum_{j=1}^{n}w_{ij}D_B(k,k+1,i,j) \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (2)$$

Where m=n=3, w11=2, w12=1, w13=2, w21=1, w22=1, w23=1, w31=2,w32=1, w33=2.

$$MD=\sum_{k=1}^{F_v-1}\frac{D(k,k+1)}{F_v-1} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (3)$$

$$STD=\sqrt{\frac{\sum_{k=1}^{F_v-1}(D(k,k+1)-MD)^2}{F_v-1}} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (4)$$

.... (4)

Threshold, T=MD + a × STD .... (5)

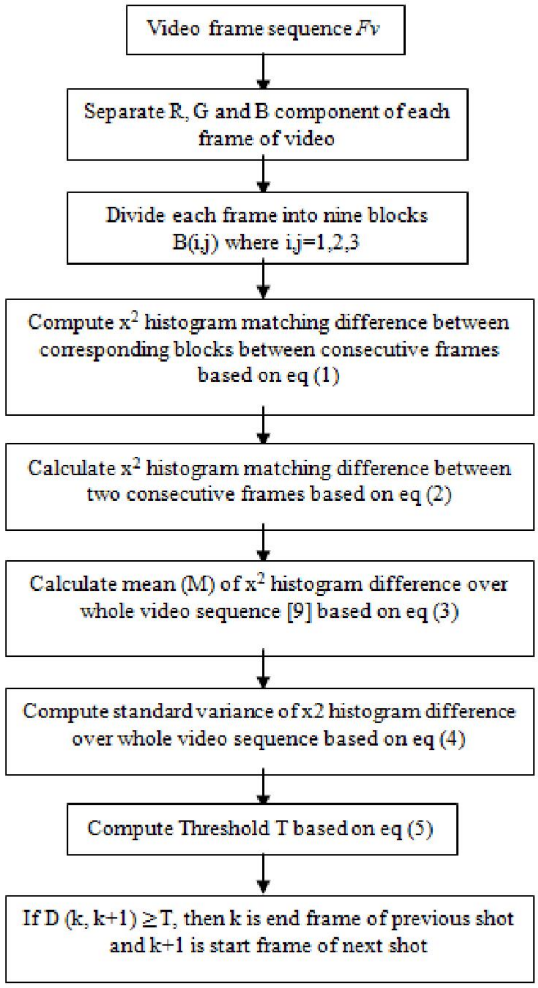

Fig. 2. Shot boundary detection

Key frame extraction (ZHAO Guang-sheng, 2008)

Step 1: First frame of each shot is reference frame and all other frames within shots are general frames. Computing the difference between all the general frames and reference frame in each shot with the above algorithm:

$$D_C(1,K)=\sum_{i=1}^{m}\sum_{j=1}^{n}w_{ij}D_{CB}(1,k,i,j) \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (6)$$

Where k=1, 2, 3…F_{CN} (k) (where F_{CN}(k) is total number of frame in current shot).

Step 2: Searching for the maximum difference within a shot:

Max (i) = {D_c (1, k)}, k= 2, 3…F_{CN}(k) .... (7)

Step 3: Determining shot type according to the relationship between Max (i) and MD: static shot (0), dynamic shot (1).

Shot type = 1 if Max (i) ≥MD
= 0 otherwise(8)

Step 4: Determining the position of key frame: if Shot Type=0, with respect to the odd number of a shot's frames, the frame in the middle of shot is chose as key frame; in the case of the even number, any one frame between the two frames in the middle of shot can be chose as key frame. If Shot Type=1, the frame with the maximum difference is declared as key frame.

Simulation Results

Here two CCTV videos are simulated.

Case 1: Video A
Video format: AVI
No. of frames/sec: 30/sec

Fig. 3. Extracted key frames from sample video A

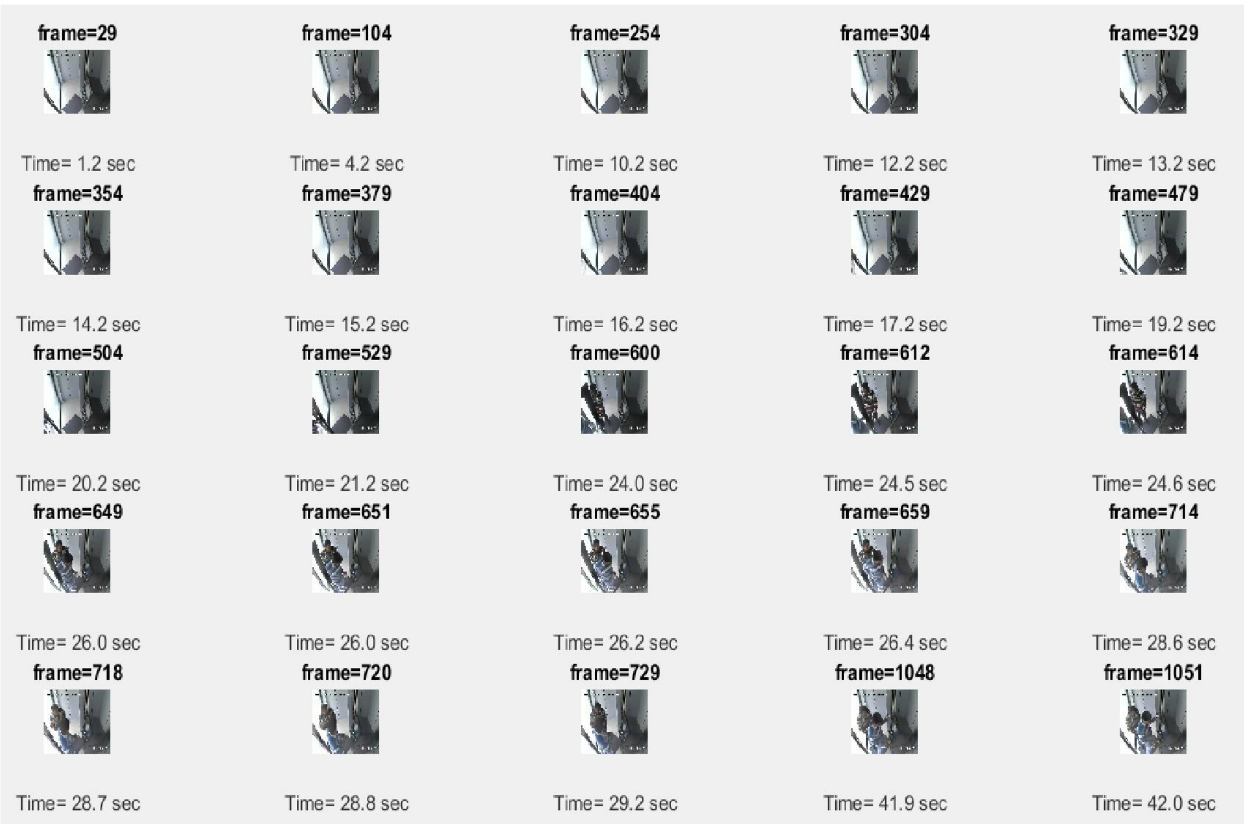

Fig. 4. Extracted key frames from video B

Total frames: 283
Duration: 10sec
Algorithm execution time: 7.275 sec

For the simplicity here we have taken very small duration video. But it can also applicable to large size/large duration video. The video here contains total 283 frames. Executing the proposed algorithm we are getting 11 key frames as shown in fig 3. The figure also shows extracted key frame number and time at which it occurs in video. From these frames we can abstract and retrieve features of video. Also we can directly move to the required content or frame from frame number or frame timing, e.g. if we need to see video when two person comes near, we can directly start video from frame no. 133, time 4.4second. It saves lots of time when we go for long duration video and want to see only particular part or object or content.

Case 2: Video B

Video format: AVI
No. of frames/sec:- 25/sec
Total frames:- 1143
Duration:- 45sec
Algorithm execution time:- 127.816 sec

After executing this video with proposed algorithm, 25 key frames are extracted with frame number and frame timing from total 1143 frames as shown in fig. 4. It eliminates so many duplicate frames within only 128 seconds.

Conclusion

We proposed algorithm for CCTV video abstraction and object tracking using shot boundary detection and key frame extraction for two types of videos. We detect shot boundary with the help of x^2 test histogram matching difference between consecutive frames with automatic threshold. After that key frames are extracted based on reference frame approach. Here key frames indicates frame number and frame timing. By using frame number and timing we can easily detect and track any object.

REFERENCES

Babaguchi, N., Y. Kawai.T. Ogura, and T. Kitahashi, 2004. “Personalized abstraction of broadcasted American football video by highlight selection”, IEEE Transaction On Multimedia, Vol.6, No.4, August, pp. 575-586.

Cernekova, Z., I. Pitas, and C Nikou, 2006. “Information Theory-based shot cut/fade detection and Video Summarization”, IEEE Transaction on Circuits and System for Video Technology, Vol.16, No.1, January, pp. 82-91.

Cheng, Y., X. Yang, and D. Xu, 2002. “A method for shot boundary detection with automatic threshold”, TENCON’02. Proceedings, IEEE Region 10 Conference on Computers, Communication, Control and Power Engineering[C], Vol.1, October 2002: 582-585.

Fereshteh Falah Chamasemani, Lilly Suriani Affendey, 2013. Systematic Review and Classification on Video Surveillance Systems, I.J. Information Technology and Computer Science, 07, 87-102

Hanjalic, A. 2002. “Shot-boundary detection: Unraveled and resolved?”, IEEE Transaction on Circuits and System for Video Technology., Vol.12, No.2, February, pp. 90-105.

Hanjalic, A. 2003. “Multimodal approach to measuring excitement in video”, Proceedings of International Conference on Multimedia and Expo ICME 03[C].Vol.2, July, pp. 289-292.

He, D. J., N. Geng, 2003. “Digital image processing”, Press of Xidia University. July, 2003, First Edition, pp. 104-106.

Seung-Hoon Han, Kuk-Jin Yoon, and In So Kweon, 2000. ”A new technique for shot detection and key frames selection in histogram space.”*12th Workshop on Image Processing and Image Understanding,* pp 475-479.

ZHAO Guang-sheng, 2008. “A Novel Approach for Shot Boundary Detection and Key Frames Extraction”, 2008 *International Conference on Multimedia and Information Technology*, *IEEE.*

Zhuang, Y., Y. Rui, T.S. Huang, and S. Mehrotra, 1998. “Adaptive key frame extraction using unsupervised clustering”, *Proceeding.ICIP’98[C], Chicago, IL*, Vol.1, pp. 866-870.

****