A COMPARATIVE STUDY FOR HUMAN BASED KEYFRAME EXTRACTION FOR CRIME SCENE INVESTIGATION IN CCTV SURVEILLANCE

Rajeshwari D[1], Victoria Priscilla C[2]

Research Scholar[1], Head & Associate Professor[2]

PG Department of Computer Science, Shrimathi Devkunvar Nanalal Bhatt Vaishnav College for Women, Chromepet.

rajisowji2011@gmail.com[1], aprofvictoria@gmail.com[2]

ABSTRACT:

The crime scene investigation through surveillance has to recognize the human/object detection involved in the suspect. Due to the eruptive growth of the surveillance data, the requisite information of gathering the crime content from the surveillance effectively becomes a desperate problem. To solve this problem the content based video retrieval scores with the keyframe extraction provides an extensive attention for an effective solution. This paper reviews a comparative study on video human/object detection by various methods such as background subtraction, optical flow and spatio-temporal filter descriptors. The study proposes the detection by the pursue of obtaining the global and local motion state information of moving objects by the combination of spatio-temporal descriptors along with the methods of keyframe extraction results a summarized video. Thus, an optimal human based keyframe extraction method for crime investigations of suspecting the human by capturing their local motion state changes through spatio-temporal filter results with greater accuracy level.

KEYWORDS: *CCTV Surveillance, Keyframe Extraction, Human motion detection, Background Subtraction, Spatio-temporal*

1. INTRODUCTION

In current scenario, CCTV surveillance is one of the compact for preventing the crime or suspecting the crime. The potential barrier of CCTV surveillance is to identify the person involved in the crime either directly or indirectly from the recorded videos. The drawback of identifying the suspect by the cops is to analyse the recorded videos for long hours to detect the particular crime scene. To overcome these challenges the content based video retrieval is one of the configuration of retrieving the required content using shot boundary detection and keyframe extraction as deprived in Fig.1. As CCTV surveillance hoard a sequence of footages of a particular

distance so shot boundary detection doesn't prevails. Hence it retrieves the required content as keyframes from the recorded footages.
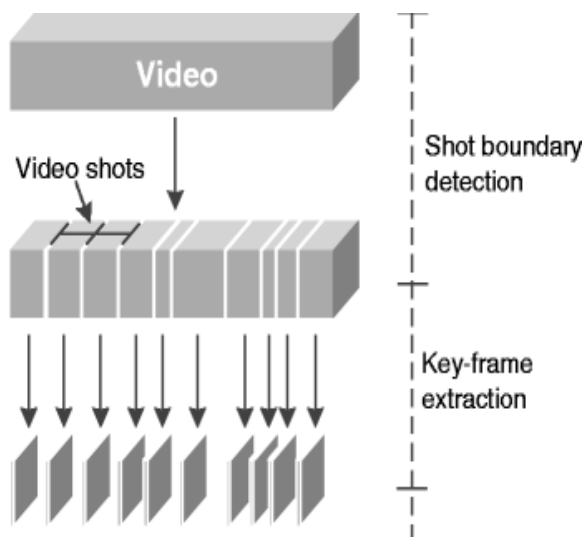

FIG. 1. CONTENT BASED VIDEO RETRIEVAL [1]

The keyframe extraction method for detecting the motion can be screwed by various algorithms such as spatio-temporal filter, optical flow and background subtraction. Still the keyframe extraction for crime scene detection lags to point out the particular person occurred in the crime. The combined work of spatio-temporal method (HOG and SVM) identifying the human as rectangular box with certain keyframe extraction methods evaporates the required keyframe with the person suspected in the crime.

The paper is systemized as: Section 2 explains the related work of the keyframe extraction related to video surveillance. Section 3 elaborates a detailed study on detecting the Human/Object by various studies. Section 4 followed by human classification resulting keyframe extraction ending up with the conclusion in Section 5.

2. RELATED WORK:

The survey reports for keyframe extraction using smart surveillance system in research area focus only on the traffic surveillance and the public safety. Zheng R.et.al., (2015) [2] proposed a novel approach on traffic surveillance to extract the keyframe based on the event or motion action based on GPU (Graphics processing Unit). In this perspective, the local maxima of the motion threshold value are set as keyframes in order to reduce the redundancy.

Suresh C.et.al., (2015) [3] proposed another novel approach on extracting a meaningful keyframes by the combination of entropy difference method, edge histogram descriptor method and spatial color distribution method with the reduced computational cost.

Radha et.al.,(2013) [5] proposed a novel method working on Itti Koch method based on its color, intensity and the orientation of the image present in the frame. The main obstacle of this method is applied for the detection of unauthorized person entity through the keyframe.

Ujwalla G.et.al., (2020) [4] proposed the keyframe extraction using surveillance videos by Convolutional Neural Network(CNN) to eliminate the most irrelevant frames from the footages resulting in the reduction of time and computational complexity.

Liu X.et.al., (2018) [6] presented a paper for identifying the accident scene through traffic surveillance. In this paper, a deep feature has been recognised with CNN and along with Iterative Quantization (ITQ) encoding, the given frame condensed to form the keyframe with the summarized accident scene.

3. HUMAN DETECTION METHOD:

Human detection in CCTV surveillance aims to have distinct identification of moving objects in the scenario. The detection is classified by Background subtraction, Optical flow and spatio-temporal filter.

3.1. BACKGROUND SUBTRACTION:

Background subtraction strives to detect the moving object by the difference between the current frames to the previous frame either by pixel or block reference. The background subtraction approaches through various method such as adaptive Gaussian Mixture Model[7], Mixture of Gaussian(MoG)[8], Kalman Filter[9] etc., to achieve the best detection in various background scenes.

3.2. OPTICAL FLOW:

Optical flow establishes the motion occurred in the video by the matching points on the objects over frames. Optical flow detects the flow vectors of the moving objects at a particular distance. The motion detection through optical flow has a greater advantage on the dense areas especially in the public crowded places. This method is implemented by vector based algorithm [10] to achieve the greater accuracy

3.3. SPATIO-TEMPORAL FILTER:

Spatio- temporal motion-based method is to define the motion occurrence both with spatial and temporal information. The spatio-temporal feature reports the combined features of both the motion and shape of the individual human in a cluttered video sequence [11].

Table 1. Comparative study on Motion/detection by various methods

METHODS	CHALLENGES	ACCURACY	COMPUTATIONAL TIME	REVIEW
Background Subtraction using GMM[7]	Human Detection	Moderate	Moderate	Very simple implementation. Detecting of human without actions performed good but

52

				not well defined with dynamic background
Background Subtraction using GMM and KDE[12]	Motion Detection	Moderate	Moderate	Methods such as Mixture of Gaussians and KDE(Kernel density estimation) prove very good model accuracy. KDE requires a very high storage requirement for the order of about 100 frames this prevents the easy implementation on a very low memory devices.
Background Subtraction using VIBe [13]	Motion Detection	High	Moderate	Method using VIBe, can process upto 350frames per second and with the minimum amount of resources it works for any background subtraction algorithm.
Optical flow[10]	Human detection at crowded place	Moderate	High	Good in detection of object at crowded places but with high complexity
Mixed Optical flow combination with Background Subtraction[14]	Motion Detection	Moderate	High	Mixed optical flow field to perform least squares regression estimation and use our dedicated Constrained RANSAC Algorithm(CRA) to improve both the accuracy and the speed. Optical with background subtraction heels the good result
Spatio-Temporal filter[11]	Action Detection	Moderate	Moderate	Well performed in classifying the object with both the action and shape with low complexity
Spatio-temporal with three Frame Difference Method[15]	Object Detection in presence of dynamic background	High	Moderate	Spatio–temporal using three frame difference method efficiently extract the moving objects in the presence of dynamic

				background, shadows, bad weather etc.,

The above comparative study defines at the point of background subtraction clips the human unescorted with no meaning of sensing the crime suspect as due to removal of background. The optical flow holds the vector based representation of the motion only for a particular distance. Thus the human detection related to crime scene refers mostly in the spatio-temporal feature. This feature is the only representation of human by its motion and action along with the undesired changes in the background which supports for identifying the suspect with proof.
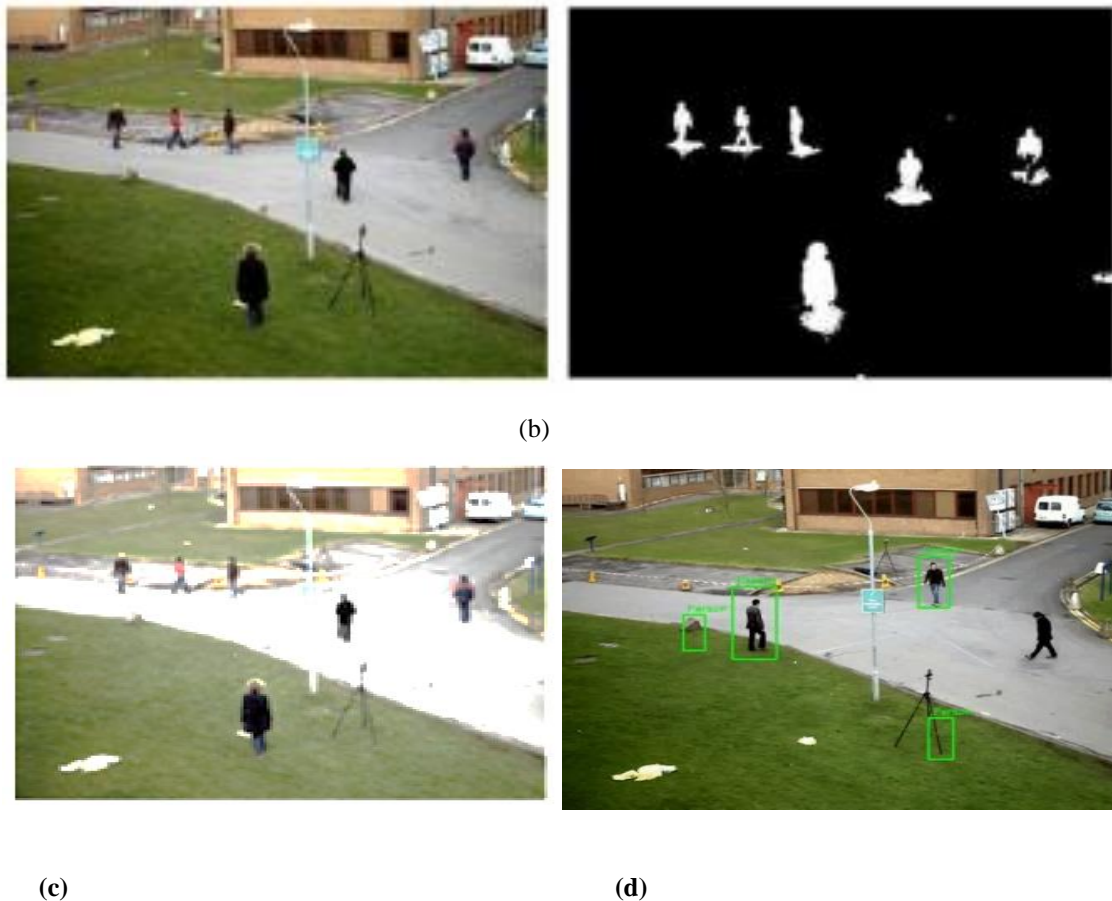

(a)   (b)

(c)   (d)

Fig. 2 (a)Original Frame (b)Background Subtracted Frame (c) Optical Flow Frame (d) Spatio-temporal Frame

4. HUMAN BASED KEYFRAME EXTRACTION METHODS

Keyframe extraction is the most efficient method used to express a clear summarized report as keyframes from the video content frames by the removal /deletion of the duplicated ones from the surveillance video [16]. The extracted keyframes reports as whole visual scene by the removal of duplication [17]. The keyframe approach is to reduce the burden and the amount of time required to gather a particular scene involved in the area suspect in a more convenient and an efficient manner [18]. Some of the features of keyframe extraction are motion detection and object detection. As to the existence the crime scene prevails only the motion occurred in the surveillance. These motion detection are extracted by various methods such as frame difference, three-frame difference, adaptive background subtraction, difference based spatio-temporal entropy image, background subtraction followed by many other self-organized background subtraction.

The human motion are deprived widely when the video in converted into frames using spatio-temporal features. These approaches are followed by the various motion detection methods to obtain the keyframe detection in combination with the humans to suspect the crime held by person in hand.

Fig.3 Human Based Keyframe Extracted

5. CONCLUSION AND FUTURE WORK:

The human detection in a video surveillance is still a challenging task in the research area of crime scene investigation. To acquire these obstacles the Content Based Video Retrieval survives a lot through keyframe extraction method. The crime suspects the human present in the surveillance. Hence the human has to detect without the elimination of background. The above comparative study discusses the different object detection algorithm for the detection of human in the surveillance video. The efficiency of obtaining the human in dynamic background has been reported in spatio-temporal feature. The frames co-presented with human

detected results a keyframe by the removal of duplication. This keyframe extraction helps to attain the possibility of recovering the required crime scene with a less time and complexity. At the end of the study, a discussion is made for the future work, in suspecting the human from a long duration video sequence with the combination of spatio-temporal method along various keyframe extraction methods to sustain a great accuracy level.

REFERENCES:

1. Ferreira L., D. Silva C., Luis A., Assuncao P., (2016). Towards Keyframe extraction methods for 3D Video: a review. *EURASIP Journal on Image and Video Processing*, 28- doi: 10.1186/s13640-016-0131-8.

2. Zheng R., Yao C., Jin h.,Zhu L., Zhang Q., Deng W., & Zhou F.,(2015). "Paralled Key Frame Extraction for Surveillance Video Service in Smart city". *PLOS ONE*, 10(8), e0135694, doi:10.1371/journal.pone.0135694.

3. Suresh C. R., Charul B., & Anand S. J., (2015)."A Novel Framework for Efficiednt Extrction of Meaningful Keyframes from Surveillance Video". *International Journal of System Dynamic Applications*. Vol. 4, Issue 2. PP. 56-73.

4. Ujjwalla G., Kamal H., & Yogesh G., (2020). "Deep Learning Approach to Keyframe Detection in Human Action Videos". *Recent trends in Computation Intelligence*. 10.5772/intechopen78839(Chapter 7)

5. Radha D., Amudha D., Ramyashree P., Ranju R., & Shalini S., (2013). "Detection of Unauthorized Human Entity in Surveillance Video". *International Journal of Engineering and Technology (IJET).* Vol 5. No. 3 Jun-Jul 2013, ISSN: 0975-4024.

6. Liu X., Liu J., Lingchen G., & Yannan R., (2018). "Keyframe Based Vehicle Surveillance Video Retrieval". *International Journal of Digital Crime and Forensics*, 10(4), 52-61. Doi:10.4018/IJDCF.2018100104.

7. Kanangamalliga S., Vasuki, & Shanmugapriya M. (2016). "Foreground Object Detection Using Expectation Maximation Based Effective Gaussian Mixture Model". *Innovations in Information, Embedded and Communication Systems*, Issue 24, pp.51-57, 2016.

8. Tian YL., Feris RS., Liu H., Hampapue A, & Sun MT., (2011). "Robust Detection of abandoned and removed objects in complex surveillance videos", *Syst. Man Cybem. Part C Appl. Rev. IEEE Trans*. 41(5), 565-576(2011).

9. Wang J, Bebis G., & Miller R., "Robust Video-based Surveillance by integrating target detection with tracking", *IEEE Computer Vision and Pattern Recognition Workshop (CVPRW '06), (IEEE, Piscataway, 2006)* pp.137.

10. Candama J., Shreve M., Goldgof DB., Sapper DB., & Kasturi R.,(2010) ."Understanding transit scenes: A survey on Human Behaviour recognition algorithms". *IEEE Trans. Intell. Transp. Syst.* 11(10), 206-224 (2010).

11. Sun J., Wu X., Yan S.C., Cheong L.F., Chua T.S., & Li J.T., (2009) "Hierarchial Spatio-Temporal Context Modeling for Action Recognition ", *CVPR*, 2009.

12. Piccardi M., (2004). "Background subtraction techniques: a review". *IEEE International conference on Systems, man and cybemetics*, 4, 3099-3104.

13. Bamich O., & Van Droogenbroeck M., (2011). "ViBe: A universal background subtraction algorithm for video sequences", *IEEE Transactions on Image Processing*. 20(6), 1709-1724.

14. Junjie H., Wei Z., Zheng Z., & Jiagang Z., (2018). "An Efficient Optical Flow Based Motion Detection method for Non-stationary Scenes". *ariXiv*:1811.08290v2 [cs.CV]

15. Paul N., Singh A., Midya A., Roy P., & Dogra D., (2017). "Moving object detection using modified temporal differencing and local fuzzy thresholding". *The Journal of Supercomputing*, 73(3), 1120-1139. Doi: 10.1007/s11227-016-1815-7.

16. Paul A., Milan K., Kavitha J., Rani J., & Arockia P., (2018). "Keyframe Extraction Techniques: A Review", *Recent Patents on Computer Science*, Vol. 11, no. 1, pp. 3-16.

17. Gharbi H., Bahroun S., & Zagrouba E., (2016). "A Novel Key Frame Extraction Approach for Video Summarization", in *VISIGRAPP (3: VISAPP*), pp. 148-155.

18. Sheena C.V., & Narayanan N., "Key-frame Extraction by analysis of histograms of video frames using statistical methods", *Procedia Computer Science*, vol. 70, pp. 36-40.