A Framework for Anomaly Event Detection by Analysing the Video Sequences

Puttegowda D, U N Sinha
Research Scholar, National Aerospace Laboratories , Bangalore
Former Distinguished Scientist, CSIR-CMMACS, Bangalore
puttegowda.77@gmail.com

ABSTRACT : Currently, it is very essential to monitor activities in Video surveillance applications both in private and public environments. In this context, our paper presents a novel framework to detect the normal or abnormal event situations by analyzing the pixel-wise motion context using block-based approaches. First motion estimation techniques are applied to characterize the events at the pixel level. Optical flow is used to extract information such as density and velocity of motion. The two different proposed approaches identifies abnormal motion variations in regions of motion activity based on the entropy of Discrete Wavelet Transform and Discrete Cosine Transform coefficients. The successful results of the detection of normal or abnormal events on different datasets are reported.

Keywords – Discrete Cosine Transform, Discrete Wavelet Transform, event detection, optical flow, Video surveillance.

I. INTRODUCTION

To ensue security at airports, banks, and institutions, Video surveillance has been a key component [1]. The main intent of visual surveillance is not only to use cameras instead of human eyes, but also the task of surveillance is carried out automatically by analyzing the video sequences. For automatic dynamic scene analysis, the challenging task for anomaly detection especially given a scene consisting of activities of multiple objects [2]. Anomaly detection techniques can be broadly classified into two broad families of approaches like pattern recognition-based and machine-learning-based methods. In pattern recognition methods, prior knowing the type of abnormal activity or object is must. But, the recognition methods require a list of objects or behavior patterns that are anomalous. Unfortunately, suspicious activities cannot be known in advance is not always possible. In machine learning based approach, by exhibiting regular activity in a video sequence, the behavior of the object is detected and then flag moving objects whose behavior deviates from normal behavior [3].

Recently, many authors have proposed learning methods based on characteristics other than motion paths to address these limitations [8]. In such a case, there is no need for object tracking, instead, features at pixel-level are considered. The main idea here, instead of tracking subjects one by one, general motion context are analyzed. We proposed a general framework based approach, where features like velocity at pixel-level are directly extracted from motion. This will lead to an image that expresses the motion in the scene. Then we analyze the information content of that image in the frequency domain by computing the entropy of the involved DWT(Discrete Wavelet Transform) or DCT(Discrete Cosine Transform) coefficients. After successfully analyzing motion in each frame, we should understand the behavior of the objects. To understand the Behavior of the objects, it involves the analysis and recognition of motion patterns, the description of actions and interactions at high level. We compare the entropies for each block to the median averaged values over time to classify events into normal and abnormal based on motion features.

The paper consisting of following sections. Section II describes the two proposed approaches for abnormal event detection including motion estimation, measuring entropy and then detecting abnormal events using DWT and DCT techniques. Section III gives experimental results and followed by Section IV conclusion and References.

II. RELATED WORKS

In the papers [4][5], a general pipeline-based framework is implemented. After detecting the moving objects initially, then by considering certain number of frames, objects are classified and tracked and finally, the activity of normal behavior of objects from the abnormal behavior are distinguished by using resulting path. Many different types of applications has proven successful different tracking-based methods, but they suffer from fundamental restrictions . First, implementing such a pipeline method can result in a delicate architecture which may go through the subsequent processing stages with an error. Secondly, tracking multiple objects at the same time requires complex algorithms [6][7]. Therefore, multi-object tracking is not always easy and efficient in crowded areas where objects occluded each other regularly fully or partially. This task is spatially hard where quality and color information can be poor in surveillance videos.

In the paper[14], temporal and spatial anomalies is proposed to detection and localization of anomalous behaviors in crowded scenes. The events in Surveillance video are extremely tedious to monitor which has very low probability. To enumerate the set of anomalies in crowds scenes that are possible in a given surveillance scenario is also infeasible. In this context, our paper presents a novel framework to detect the normal or abnormal event situations by analyzing the pixel-wise motion context using block-based approaches.

III .PROPOSED APPROACHES

A. Abnormal event detection using DWT

Our abnormal event detection is based on motion features extracted with a motion estimation technique. Motion estimation in image sequences aims at detecting regions corresponding to moving objects. For tracking and behavior analysis in the later stage, the features of moving object is detected. For noisy and coarse data, the computation of optical flow is not accurate. To deal with this, we use optical flow at each frame using the Lucas-Kanade algorithm [9]. To detect the moving regions in an image sequence relating each image to the next is used in Optical-flow-based motion estimation is considered by the characteristic of flow vectors of moving objects over time. The perceptible displacement of each pixel from image to image is represented in each vector [10]. The result of optical flow at both vertical and horizontal direction is the value of displacement of each pixel. To obtain a motion magnitude vector, we combine this displacement. To progression these motion vectors, we replace with pixel values for the estimated motion and each frame is divided into blocks. We anticipate that during abnormal events the motion patterns and the energy of the images containing motion vectors change compared to normal behavior. Here we use DWT to find the coefficients. We apply a DWT to each block, as the DWT provides a compact representation of the signals energy. Then we compute the entropy of the DWT coefficients to measure the information content of the DWT coefficients.

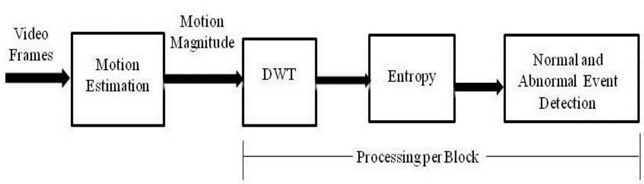

Fig. 1: Proposed approach block diagram

The Fig.1 shows how actually the design of our proposed work is, where the processing is done block by block. Firstly, the video will be given as a input, that video will be divided into number of frames, these video frames are given as input to the Motion Estimation block.

The output of the Motion Estimation block will be the motion magnitude, this motion magnitude will be the input to the Discrete Wavelet Transform (DWT) block. In DWT block only level one is used. The four parts of the DWT are LL, LH, HL, HH. Out of these four parts, only three parts like LH, HL and HH are used for object detection. In the next block the high pass outputs i.e., HH, HL and LH parts are combined together which results in a form in which most of the image pixels becomes dark except the outline of the object which is moving in the input video. The output of the Discrete Wavelet Transform (DWT) block will be of coefficient values and these values will be the input to the Entropy block.

The output of the Entropy block will decide whether the event present in the given video data set is Normal event or Abnormal event. This is the output of this below block diagram that is output of our proposed system. Processing of this below block diagram will be done per block.

B. Abnormal event detection using DCT

The second method for abnormal event detection using DCT. Using the Motion estimation of image sequences are used to detecting regions corresponding to moving objects. For tracking and behavior analysis in the later stage, the features of moving object is detected. . For noisy and coarse data, the computation of optical flow is not accurate. To deal with this, we use optical flow at each frame using the Lucas-Kanade algorithm. To detect the moving regions in an image sequence relating each image to the next is used in Optical-flow-based motion estimation is considered by the characteristic of flow vectors of moving objects over time. The perceptible displacement of each pixel from image to image is represented in each vector. The result of optical flow at both vertical and horizontal direction is the value of displacement of each pixel. To obtain a motion magnitude vector, we combine this displacement. To progression these motion vectors, we replace with pixel values for the estimated motion and each frame is divided into blocks. We anticipate that during abnormal events the motion patterns and the energy of the images containing motion vectors change compared to normal behavior. To find the coefficients, here we used Discrete Cosine Transform. DCT provides a compact representation of the signals energy when we apply DCT to each block. Then we compute the entropy of the DCT coefficients to measure the information content of the DCT coefficients.

$$T = (u, v)\alpha_u \alpha_u \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) P.Q \tag{1}$$

Where

$$P = \cos\frac{(2x+1)u\pi}{2M}$$

and

$$Q = \cos\frac{(2y+1)v\pi}{2N}$$

$$\alpha_u = \alpha_v = \begin{cases} \sqrt{\frac{1}{M}} & u = v = 0 \\ \sqrt{\frac{2}{N}} & u \neq v \neq 0 \end{cases}$$

f(x,y) = input image with size M x N, where M=row and N=Column

The entropy is defined as [12]:

$$E = -\sum_{i=1}^{N} p \log p \qquad (2)$$

where N is the size of image and at a certain pixel location, probability of the motion intensity value is stored in p. We compare the entropy value with thresholds which we learn per block in the beginning of the video sequence to decide whether event is normal event or abnormal event. By considering the first 100 frames of the video, the threshold value is estimated based on a median value of the entropies. When the value of the entropy for the current frame is higher than the threshold defined for that block, then an abnormal event is indicated . Fig. 2 illustrates block-based processing framework using DCT in dynamic scenes.

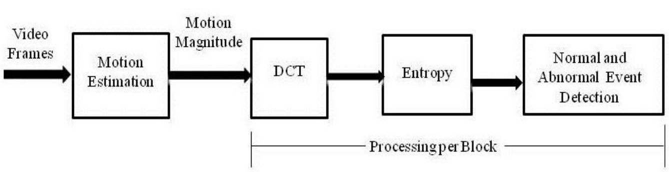

Fig. 2: DCT- Block based processing framework

IV. EXPERIMENTAL RESULTS

In our experiments, we use video dataset where several abnormal situations are simulated by a group of volunteers. These situations include running of several people to the middle and from the middle of the scene For an event to be considered normal or abnormal based on motion features, we compare the entropies for each block to the median averaged values over time to classify events into normal and abnormal using applying either DCT or DWT technique.

The video was divided into contiguous clips or frames. Anomalous video clips were separated from the rest of the video clips for testing. In the test data, anomalous clips were considered as positive examples and the non-anomalous clips were considered as negative examples.

In our implementation, we divide each frame into 4 blocks. For each block we calculate the entropy of the DCT or DWT coefficients of the motion vector magnitudes and then compute the median value over the first 100 frames. Based on experiments and evaluation, the threshold for the median entropy to classify an abnormal event is empirically set to 3 times the median value. Abnormality is detected for the whole frame is raised if abnormality is detected in any of the blocks.

Data Set 1: This video is of 25 sec, it is named as ―Mov_0350‖ which is a footage recorded in college campus where we can observe that many people talking to each other and laughing, which actually indicates the normal event but all of a sudden some abnormal event took place and everyone started running in all the directions and this shows the abnormal event occurred in the video.

The video is divided into 95 frames. The video can be divided into any number of frames and it is depend on the user‘s choice. And the movement occurred in the video which indicates the abnormal is shown below and it is in the form of frames. Fig. 3a and 3b shows the frames of abnormal and normal situation in dataset1 video. Fig. 4a and 4b shows the motion estimation of abnormal and normal event in dataset1 video.

Fig. 3a : abnormal situation frame - DWT block based framework

Fig. 3b : Normal situation frame - DWT block based framework

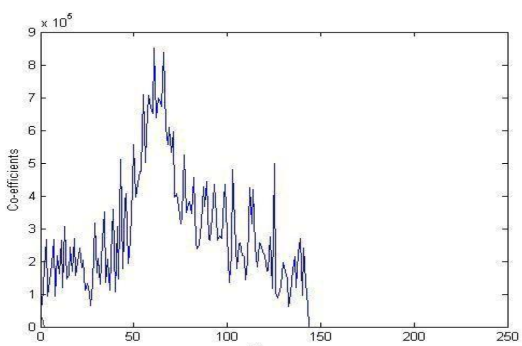

Fig. 4a: Graph indicates the motion estimation- Abnormal

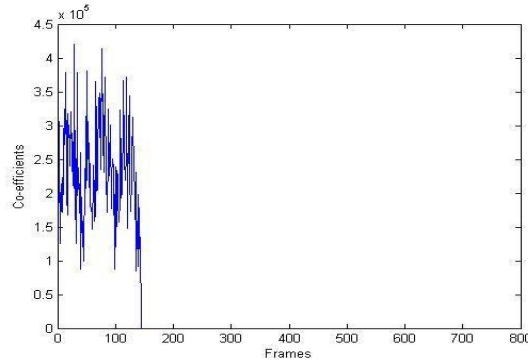

Fig. 4b: Graph indicates the motion estimation- Normal

Data Set 2: This video is of 1 min 40 sec, it is named as –Panic video‖ which is a footage recorded in train where we can observe that many people getting inside the train and settled down, which actually indicates the normal event but all of a sudden some event took place and everyone started running and get down from the train and after some time again everyone get inside the train and this shows the abnormal event occurred in the video.

The video is divided into 95 frames. The video can be divided into any number of frames and it is depend on the user‘s choice. And the movement occurred in the video which indicates the abnormal is shown below and it is in the form of frames. Fig. 5 shows the frames of abnormal situation in dataset 2 video. Fig. 6 shows the motion estimation of abnormal event in dataset2 video.

Fig. 5: abnormal situation frame - DCT block based framework

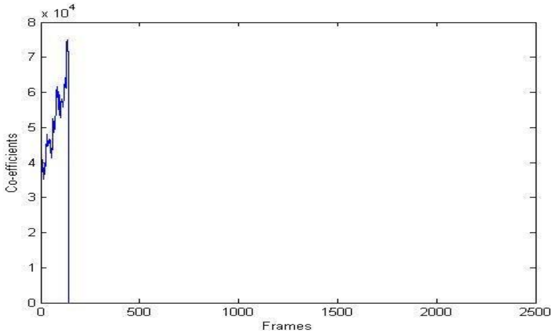

Fig. 6: Graph indicates the motion estimation

V. CONCLUSIONS

In this paper, we have developed a motion-context-based algorithm by applying DWT and DCT techniques to detect abnormal events in surveillance videos of a public place. We are implemented informative features based on motion and using threshold to detect abnormal events. In our approach, for classifying whether the current activity in the video is normal or abnormal, the entropy of the DWT or DCT-transformed motion magnitude is a reliable measure. Because the proposed methods are block-based, we can indicate exactly in which part of the frame the abnormal event takes place. Our proposed system detects the anomalies successfully and discriminate between Normal and Abnormal events in the given video data set.

REFERENCES

[1] L. Ovsenk, A. Kolesrov, J. Turn, Video Surveillance Systems‖, Acta- Electrotechnica et Informatica, Kosice: FEI TU, vol. 10, no. 4, pp.46-53, 2010.

[2] J. Li, Sh. Gong, T. Xiang, Global Behaviour Inference using Probabilistic Latent Semantic Analysis‖, the 19th British Machine Vision Conference (BMVC), pp. 193-202, Leeds, UK, September, 2008.

[3] Y. Benezeth, P.-M. Jodoin, V.Saligrama, C. Rosenberger, Abnormal Event Detection based on spatio-Temporal Co-occurence‖, in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, pp. 2458- 2465, 2009.

[4] T. Chen, H. Haussecker, A. Bovyrin, et al., Computer vision workload analysis: case study of video surveillance systems‖, intel. Technology journal, vol. 9, no. 2, pp. 109-118, 2005.

[5] H.Boxton, Learning and understanding dynamic scene activity: A review‖, Image and Vision Computing, vol. 21, no. 1, pp. 125-136, 2003.

[6] N. Johnson and D. hogg, Learning the distribution of object trajectories for event recognition‖, Image and Vision Computing, vol. 14, no. 8, pp. 609-615, August, 1996.

[7] I. Junejo, O. javed and M. Shah, Multi feature path modeling for video surveillance‖, In Proc. of ICPR, no. 04, pp. 716-719, 2004.

[8] N. Ihaddadene, C. Djeraba, Real-time Crowd Motion Analysis, 19th International Conference on Pattern Recognition (ICPR), pp. 08-11, Tampa, Florida - USA, December, 2008.

[9] B. D Lucas and T. Kanade, An iterative image registration technique with an application to stereo vision‖, In DARPA Image Understanding workshop, April, 1981.

[10] A., Wali, A.M, Alimi, Event Detection from Video Surveillance Data Based on Optical Flow Histogram and High-level Feature Extraction‖, 20th International Workshop on Database and Expert Systems Application, 2009.

[11] Syed Ali Khayam, (2003) The Discrete Cosine Transform : : Theory and Application.

[12] Gonzalez, R.C., R.E. Woods, S.L. Eddins, Digital Image Processing Using MATLAB, Gatesmark Publishing, 2nd edition, 2009.

[13] Chen T.s, Chang C. C.,and Hwang, K. S. F.,2002.Digital Image Processing, Taipei : Flag

[14] Weixin Li, Student Member, IEEE, Vijay Mahadevan, Member, IEEE, and Nuno Vasconcelos, Senior Member, IEEE, ―Anomaly Detection And Localization In Crowded Scenes‖, January 2014.

Authors Biography

Mr. Puttegowda D, M.Tech (CSE) working as Associate Professor, Department of computer science and Engineering, ATME College of Engineering, Mysuru, Karnataka, currently pursuing research work on Video Mining in NAL research centre, Bangalore, affiliated to University of Mysore and has eleven years of teaching experience and two years of industry experience and published six international paper.

Dr. U N Sinha, the scientist, is very well known in the NAL, CSIR and the Indian scientific circles for his breadth and depth of knowledge in mathematics, fluid dynamics thermodynamics, parallel computing, atmospheric science and Sanskrit. Dr. U N Sinha obtained his engineering degree engineering in 1967 and his PhD from IIT Kanpur in 1976. In 1986 Flosolver project was born. It is creditable that Dr. Sinha and his team developed Flosolver Mk1, Indiaís first parallel computer and in 1993 Dr.U N Sinha and his team was first to complete the parallelization taking up of DSTís project to parallelize a global weather prediction model which was being used by NCMRWF for operational forecasts. Apart from his scientific achievements, Dr. Sinha has many other accomplishments to his credit: building a very good book collection in the library, teaching a very large number of students, taking good care of his colleagues.