Scientific Journal of Impact Factor (SJIF): 3.134

E-ISSN (O): 2348-4470
P-ISSN (P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 2,Issue 11,November -2015

Analysis of video sequences for anomaly detection using block based approaches

Puttegowda D[1], U N Sinha [2]

[1] Research Scholar, National Aerospace Laboratories, Bangalore, India
[2] Former Distinguished Scientist, CSIR-CMMACS, Bangalore, India

Abstract—*Presently it is very important both in private and public environments to monitor activities in Video surveillance applications . In this context, this paper presents a novel block-based approaches to detect abnormal event situations by analyzing the pixel-wise motion context. We proceed using motion estimation techniques to characterize the events at the pixel level. Optical flow is used to extract information such as density and velocity of motion. The proposed approaches identifies abnormal motion variations in regions of motion activity based on the entropy of Discrete Wavelet Transform and Discrete Cosine Transform coefficients. We will report successful results on the detection of abnormal events in video datasets.*

Index Terms—*Video surveillance, optical flow, Discrete Cosine Transform, Discrete Wavelet Transform event detection.*

I. INTRODUCTION

To ensue security at airports, banks, and institutions, Video surveillance has been a key component [1]. The goal of visual surveillance is not only to use cameras instead of human eyes, but also to accomplish the entire surveillance task as automatically as possible using video analysis. For automatic dynamic scene analysis, anomaly detection is a challenging task especially given a scene consisting of activities of multiple objects [2]. Anomaly detection techniques can be divided into two broad families of approaches, namely pattern recognition- based and machine-learning-based methods. The pattern recognition approaches are typically those where the type of abnormal activity or object is a priori known. But, the recognition methods require a list of objects or behaviour patterns that are anomalous. Unfortunately, this is not always possible, especially where suspicious activities cannot be known in advance. An alternative approach is based on learning normal" behavior from a video sequence exhibiting regular activity and then flag moving objects whose behavior deviates from normal behavior [3]. In the papers [4][5], many methods implement a general pipeline-based framework: at first moving objects are detected, then they are classified and tracked over a certain number of frames and finally, the resulting paths are used to distinguish normal" behavior of objects from the abnormal" [6] [7]. Although tracking-based methods have proven successful in different applications, they suffer from fundamental limitations. First, implementing such pipeline methods can result in a fragile architecture which may suffer from an error propagating through the subsequent processing stages. Secondly, tracking multiple objects at the same time requires complex algorithms and is computationally heavy. Therefore, multi-object tracking is not always efficient in crowded areas where objects regularly fully or partially occlude each other. This task is spatially hard in surveillance videos where quality and color information can be poor.

To address these limitations, some authors have recently proposed learning methods based on characteristics other than motion paths [8]. In such a case, there is no need for object tracking, instead, we consider pixel-level features. The main idea is to analyze the general motion context instead of tracking subjects one by one. We design a general framework based on features directly extracted from motion such as velocity at pixel-level. 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 or DCT coefficients. After successfully analyzing motion in each frame, we should understand the behaviour of the objects. Behavior understanding involves the analysis and recognition of motion patterns, and the description of actions and interactions at high level. 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.

The paper consisting of following sections. Section 2 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 3 presents our experimental results. Section 4 concludes the paper.

II. 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. Detecting moving regions provides features for later processing such as tracking and behaviour analysis. The computation of optical flow is not very accurate, particularly on coarse and noisy data. To deal with this, we use optical flow at each frame using the Lucas-Kanade algorithm [9]. Optical-flow-based motion estimation uses characteristics of flow vectors of moving objects over time to detect moving regions in an image sequence, relating each image to the next. Each vector represents the apparent displacement of each pixel from image to image [10]. The result of optical flow is the value of displacement of each pixel at both vertical and horizontal direction. We combine this displacement to obtain a motion magnitude vector. To process these motion vectors, we substitute pixel values for the estimated motion and we divide each frame into blocks. We expect 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 .

The figure 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, the output of the Discrete Wavelet Transform (DWT) block will be of some values, 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.

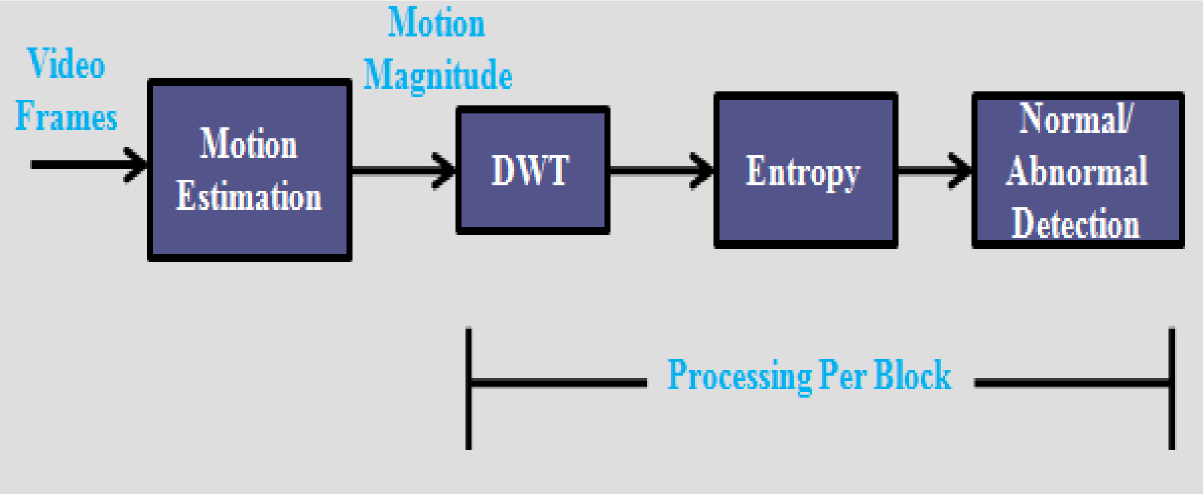

Figure 1: Proposed approach block diagram.

The figure 2 shows the work flow of our designed work, starting with the video data set selection from the folder, then the selected video is converted into number of frames. Motion estimation in video frames aims at detecting regions corresponding to moving objects, that is frame by frame such as vehicles, humans and other objects. Motion features are extracted with a Motion estimation technique. The output of the Motion estimation block will be the motion magnitude, which is given as the input to the DWT block.

The Discrete Wavelet Transform (DWT) is applied to the input video. Here only level one DWT is used. Now except LL part, remaining three parts i.e. LH, HL and HH of the Discrete Wavelet Transform 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 previous block will be the value of the combined high pass outputs, by comparing this value with the Entropy value, it is going to differentiate between the Normal or Abnormal events.

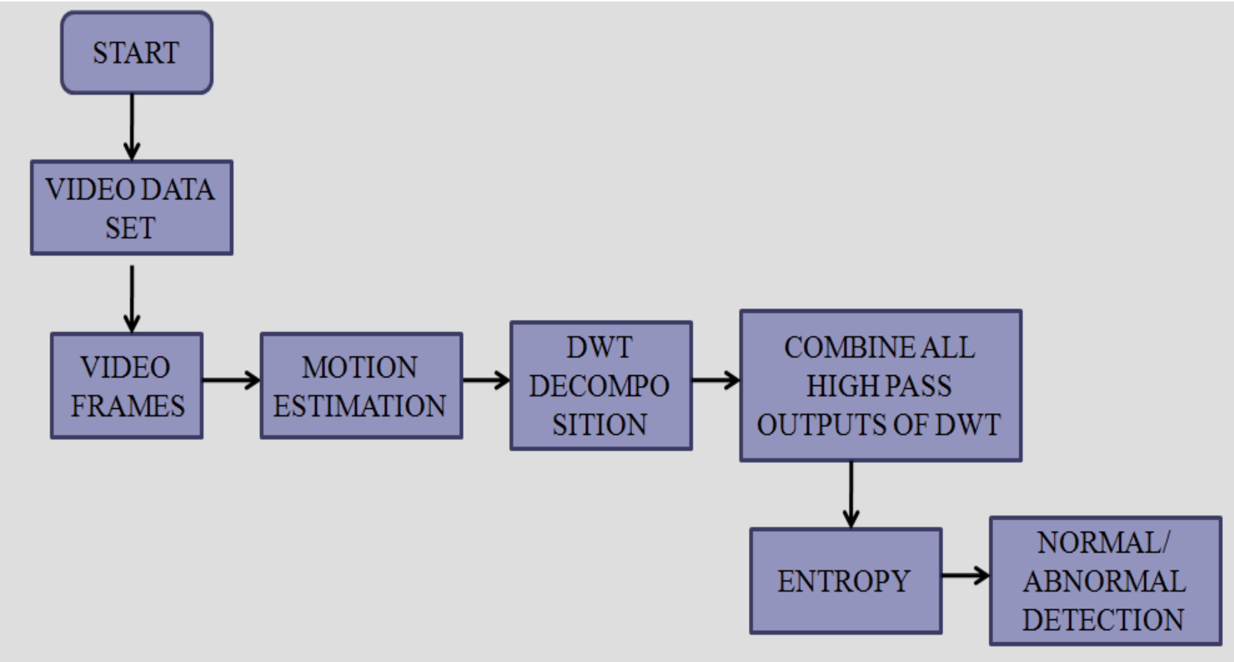

Figure 2: Work flow diagram.

B. Abnormal event detection using DCT

The second method for abnormal event detection using DCT. Motion estimation in image sequences aims at detecting regions corresponding to moving objects. Detecting moving regions provides features for later processing such as tracking and behaviour analysis. The computation of optical flow is not very accurate, particularly on coarse and noisy data. To deal with this, we use optical flow at each frame using the Lucas-Kanade algorithm. Optical-flow-based motion estimation uses characteristics of flow vectors of moving objects over time to detect moving regions in an image sequence, relating each image to the next. Each vector represents the apparent displacement of each pixel from image to image. The result of optical flow is the value of displacement of each pixel at both vertical and horizontal direction. We combine this displacement to obtain a motion magnitude vector. To process these motion vectors, we substitute pixel values for the estimated motion and we divide each frame into blocks. We expect that during abnormal events the motion patterns and the energy of the images containing motion vectors change compared to normal behavior. Here we use DCT to find the coefficients. We apply a DCT to each block, as the DCT provides a compact representation of the signals energy. Then we compute the entropy of the DCT coefficients to measure the information content of the DCT coefficients.

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

Where

$$P = \cos\frac{(2x+1)u\pi}{2M} \quad \text{and} \quad Q = \cos\frac{(2y+1)v\pi}{2N}$$

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

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

The entropy is defined as [12]:

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

where N is the size of image and p contains the probability of the motion intensity value at a certain pixel location. For deciding whether the event is normal or abnormal, we compare the entropy value with thresholds which we learn per block in the beginning of the video sequence. This threshold is based on a median value of the entropies which we estimate during the first 100 frames of the video. An abnormal event is indicated when the value of the entropy for the current frame is higher than the threshold defined for that block. Figure 1 illustrates block-based processing framework using DCT in dynamic scenes.

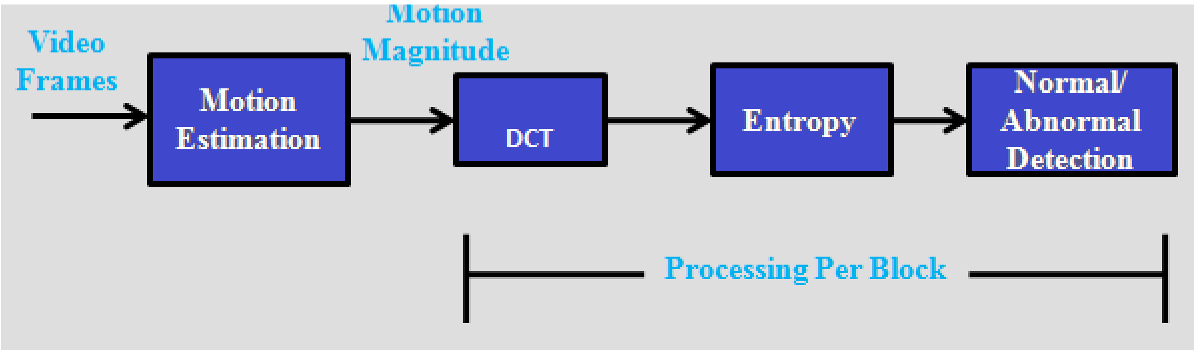

Figure 3: **DCT- Block based processing framework**

III. 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 experimentation was conducted by keeping the number of actions in the video. The number of actions served as the number of topics in the document. The video was divided into contiguous clips or frames each and each clip serving as a document in the model. 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. Figure 4 shows the frames of abnormal situation in dataset1 video. Figure 5 shows the motion estimation of abnormal event in dataset1 video.

Figure 4: abnormal situation frame - DWT block based framework

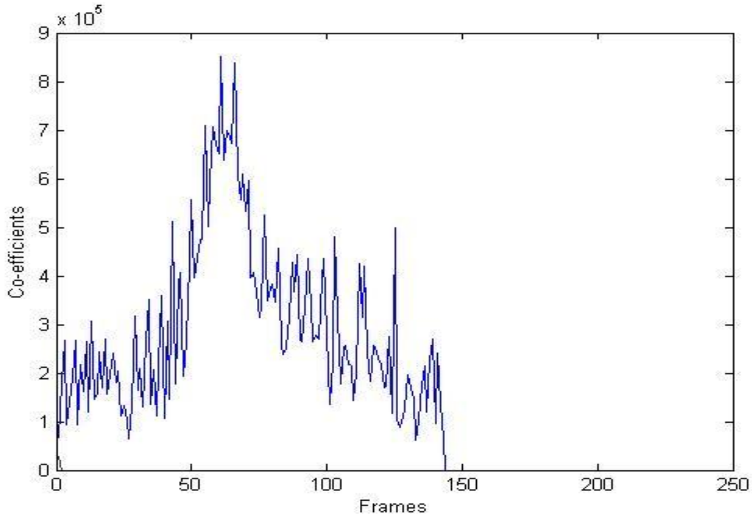

Figure 5: Graph indicates the motion estimation

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. Figure 6 shows the frames of abnormal situation in dataset 2 video. Figure 7 shows the motion estimation of abnormal event in dataset2 video.

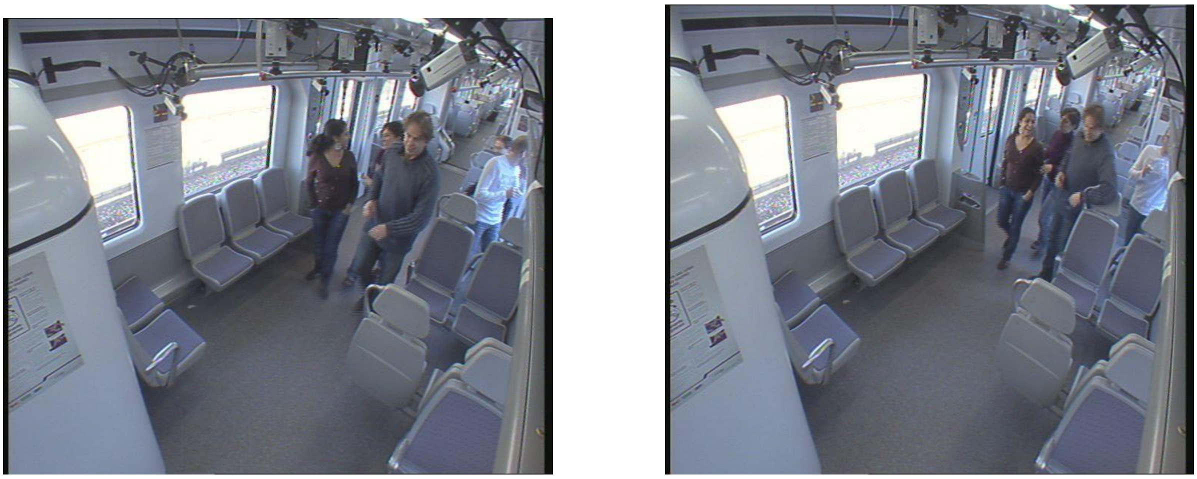

Figure 6: abnormal situation frame - DCT block based framework

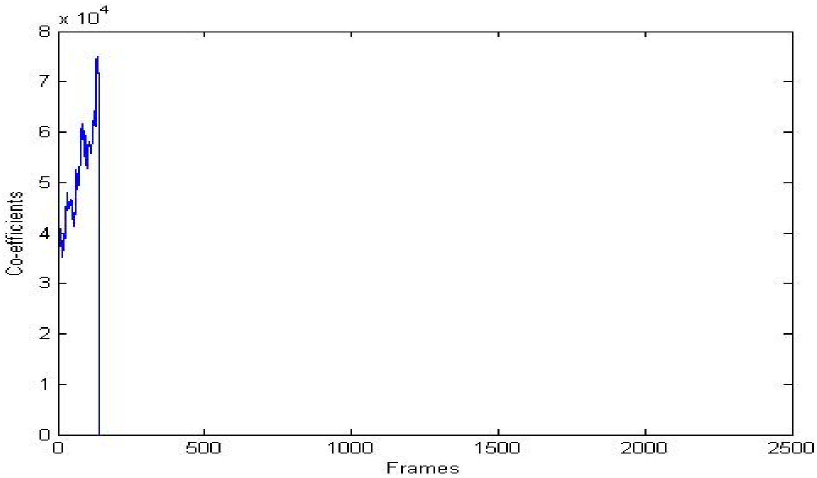

Figure 7: Graph indicates the motion estimation

IV. 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. We have discovered that the entropy of the DWT or DCT-transformed motion magnitude is a reliable measure for classifying whether the current activity in the video is normal or not in our approach. Because the proposed methods are block-based, we can indicate exactly in which part of the frame the abnormal event takes place. We are overcoming some of the drawbacks found in the existing systems such as, occlusions, DCT method, normal environment, crowd regions etc. In the papers the others have mentioned that if occlusions occur their system would not get result properly, in other paper they have told that the environment should be normal to get proper results, by DCT method the performance is low compared to DWT method, these are the some of the drawbacks we are overcoming in our proposed system. 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.

