ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 5, May 2016

Event Detection in Video Using Saliency Value and Histogram of Optical Flow

Chetana D Patil, Bharathi V K

M. Tech Student, Dept. of Digital Electronics and Communication Engineering, PESIT, Bangalore, India

Assistant Professor, Dept. of Digital Electronics and Communication Engineering, PESIT, Bangalore, India

ABSTRACT: The range of applications is huge in the field of object/event detection, so abnormal event detection has gained marvelous research attention. In video analysis abnormal event detection is a challenging task. In this paper, we proposes an architecture for abnormal event detection in video using saliency value and Histogram of optical flow, for many video surveillance application event detection is most important. An event might be characterized as an event that deviates from the normal. The main difficulty in detecting such events is their unpredictability. The problems such as diversity of monitoring scene, different crowd density and mutual occlusion among crowds etc result in a low recognition rate for abnormal crowd behavior detection. In order to solve these problems, this project uses salience value for each frame, to describe the normal patterns on the scene, we are using a Histogram of optical flow orientation and magnitude (HOFM) feature descriptor based on optical flow information and KNN classifier is used for classification.

KEYWORDS: Histogram of optical flow, Histogram of optical flow orientation and magnitude (HOFM), salience value.

I. Introduction

Abnormal event detection in crowded scenes is an important task in intelligent surveillance video systems. Discovery of suspicious or anomalous events from video streams is an interesting yet challenging task for many video surveillance applications. By automatically finding suspicious events, it significantly reduces the cost to label and annotate the video streams of hundreds of thousands of hours. Abnormal crowd behavior detection is an advanced topic researched in fields of computer vision and digital image processing. The problems such as diversity of monitoring scene, different crowd density and mutual occlusion among crowds etc result in a low recognition rate for abnormal crowd behavior detection. However, it is extremely difficult to design activity recognition approaches without specific knowledge of the scene and the target activities. Therefore, researchers have developed approaches to locate and recognize anomalous events and possibly hazardous human motions using only the knowledge regarding the normal behavior at a given location, without requiring an extensive knowledge of the scene. The main difficulty in detecting such events is their unpredictability. The problems such as diversity of monitoring scene, different crowd density and mutual occlusion among crowds etc result in a low recognition rate for abnormal crowd behavior detection. In order to solve these problems, this project uses salience value for each frame, to describe the normal patterns on the scene; we are using a Histogram of optical flow orientation and magnitude (HOFM) feature descriptor based on optical flow information. In this approach, the HOFM is extracted from cuboids sampled over several frames from non-overlapping special regions. During the learning stage, where, only videos containing normal events are presented, we extract and store the HOFM feature vectors for each spatial region, generating a set of normal patterns. Then, during the testing stage, after extracting HOFM, a nearest neighbor search is performed considering only that particular region and, according to the distance to the best matching pattern, the event taking place at the particular location and time might be classified as anomalous. According to experimental results, the proposed descriptor combined with a simple nearest neighbor search is able to detect anomalous events accurately.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 5, May 2016

II. LITERATURE SURVEY

Yang Cong et al. [1] has proposed to detection of abnormal events via a sparse reconstruction over the normal bases. In this paper they introduced the sparse reconstruction cost (SRC) over the normal dictionary to measure the normalness of the testing sample. By introducing the prior weight of each basis during sparse reconstruction, the proposed SRC is more robust compared to other outlier detection criteria. Their method provides a unified solution to detect both local abnormal events (LAE) and global abnormal events (GAE). We further extend it to support online abnormal event detection by updating the dictionary incrementally. Experiments on three benchmark datasets and the comparison to the state-of-the-art methods validate the advantages of our algorithm. Ernesto L Et al. [2] proposed a method to presents an automatic technique for detection of abnormal events in crowds. Crowd behavior is difficult to predict and might not be easily semantically translated. Moreover it is difficult to track individuals in the crowd using state of the art tracking algorithms. Therefore they characterize crowd behavior by observing the crowd optical flow and use unsupervised feature extraction to encode normal crowd behavior. The unsupervised feature extraction applies spectral clustering to find the optimal number of models to represent normal motion patterns. The motion models are HMMs to cope with the variable number of motion samples that might be present in each observation window. The results on simulated crowds demonstrate the effectiveness of the approach for detecting crowd emergency scenarios. Vijay Mahadevan et al. [3] proposed model for normal crowd behavior is based on mixtures of dynamic textures and outliers under this model are labeled as anomalies. Temporal anomalies are equated to events of low-probability, while spatial anomalies are handled using discriminate saliency. An experimental evaluation is conducted with a new dataset of crowded scenes, composed of 100 video sequences and five well defined abnormality categories. The proposed representation is shown to outperform various state of the art anomaly detection techniques. Simon Hartmann et al. [4] proposed architecture on automatic abnormality detection in video sequences has recently gained an increasing attention within the research community. Although progress has been seen, there are still some limitations in current research. While most systems are designed at detecting specific abnormality, others which are capable of detecting more than two types of abnormalities rely on heavy computation. Therefore, they provide a framework for detecting abnormalities in video surveillance by using multiple features and cascade classifiers, yet achieve above real-time processing speed. Experimental results on two datasets show that the proposed framework can reliably detect abnormalities in the video sequence, outperforming the current state-of-the-art methods.

III. METHODOLOGY

In our proposed work, we present an approach for anomaly detection, illustrated in Figure 1 There are two important phases in the architecture they are Testing and Training phase. During the training phase, our approach extracts a novel spatiotemporal feature descriptor, called Histograms of Optical Flow Orientation and Magnitude, to capture the moving patterns from non-overlapping regions in the video and to describe the normal patterns on the scene. Such descriptors are stored to be used as instances of normal events (the training video sequences contain only normal events). Saliency value each frame is calculated and stored in a database. During the testing phase, considered frames are done through frame differences and incoming patterns for each region are compared to the respective stored patterns using a k nearest neighbor search approach. Those patterns presenting a significant difference compared to the entire stored one are considered as anomalous. The experimental evaluation demonstrates that our approach is able to detect anomalous events with success, achieving better results.

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 5, May 2016

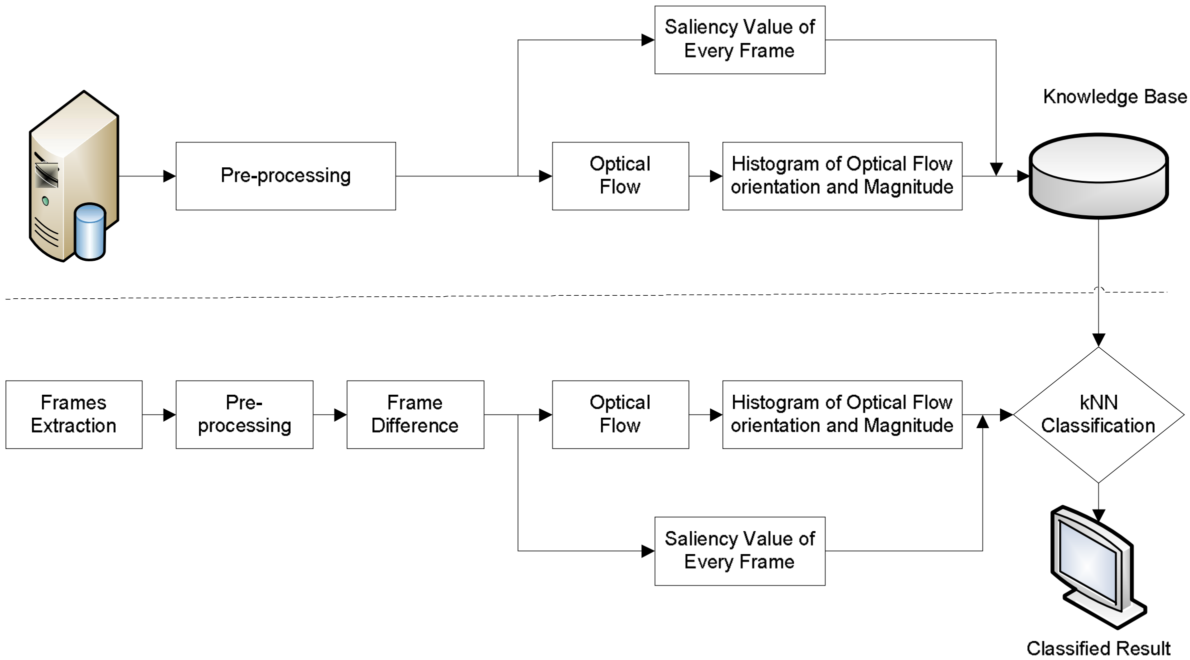

Fig. 1. Block Diagram of proposed system.

A. PRE PROCESSING

Pre-processing is used to convert the character image into gray scale image and to adjust the size of the image. Pre-processing is any form of signal processing for which the output is an image or video, the output can be either an image or a set of characteristics or parameters related to image or videos to improve or change some quality of the input. This process will help to improve the video or image such that it increases the chance for success of other processes. In this paper we considered image as input and these images are subjected to pre-processing this will resulting in gray scale conversion and resizing.

B. OPTICAL FLOW EXTRACTION

The proposed spatiotemporal feature descriptor uses as input the optical flow. To avoid computing optical flow for each pixel on the image, we first create a binary mask using image subtraction between the frame Ij and the frame$Ij + t$. Given a threshold d, if the resulting difference is less than d, the pixel is discarded; otherwise, this pixel p is set to its corresponding local cuboidsCi. Thus, each cuboids has a set of moving pixels. For each$p \in C_t^i$, we compute the optical flow. For that, we use Lucas- Kanade-Tomasi pyramidal implementation, where $p'$is the optical flow result for pixel p. The pixel p_ corresponds to pixel p inC_t^i.Optical flow represents apparent motion of the object relative to the observer. Note that we only use foreground pixels to represent the feature. This makes the estimations

$$\widehat{mot}_t(i,j) = \frac{1}{N_f}\sum_{n=0}^{N_f}\left\|[v_x^{(n)}, v_y^{(n)}]1\right\| \qquad (1)$$

Where for each foreground pixel n, $v_x^{(n)}$ and $v_y^{(n)}$are the optical flow in both spatial directions and N_f is the total number of foreground pixels. To further reduce the effect of noise, the motion feature for a region is averaged by the motion features in the same region of the neighbouring frames t-1and t+1

$$mot_t(i,j) = \frac{1}{3}\sum_{u=t-1}^{t+1} \widehat{mot}_u\,(i,j) \qquad (2)$$

Pyramid Lucas-Kanade is a spares method and the average motion will therefore not be normalized based on foreground pixels, but rather based on the number of flow vectors within the region:

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 5, May 2016

$$\widehat{mot}_t(i,j) = \frac{1}{N_{f_v}} \sum_{n=0}^{N_f} \| [v_x^{(n)}, v_y^{(n)}] 1 \| \quad (3)$$

Where for each optical flow vector n, $v_x^{(n)}$ and $v_y^{(n)}$ are the optical flow in both spatial directions and N_{fv} is the total number of flow vectors within the region.

C. HOMF (HISTOGRAM OF OPTICAL FLOW ORIENTATION AND MAGNITUDE)

Now, we present our proposed descriptor. As mentioned earlier, it uses optical flow information (orientation and magnitude) to build the feature vector for each cuboid. To do this, we define a matrix $F_{S\times B}$, where S is the number of orientation ranges and B the number of HOOF magnitude ranges. We are using the information of magnitude provided by the vector field resultant of optical flow. Thus, given pixels p(x, y, t) and p'(x, y, t) that belongs to cuboid c_i^t, the vector field $\vec{v}$ between p and p' is composed by magnitude m and orientation θ. In this way for each cuboid at time t, we compute the matrix feature F using Equation 2.

$$F(s,b) = \sum_{\vec{v}\to c_i^t} \begin{cases} 1 & \text{if } (s = \mathrm{mod}(m, M)) \text{and} \\ & (b = \mathrm{mod}(\theta, B)) \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

Where s ∈ {1, 2..S} and b ∈ {1, 2..B} denote orientation and magnitude ranges, respectively. The spatiotemporal descriptors are computed for each cuboid c_i^t. This situation does not modifies the main presented idea, because here we use information of the pixel in each optical flow result, i.e., that each pixel in the cuboid provides information for a determinate bin in the feature vector regarding the same cuboid.

D. KNN CLASSIFIER

Among several methods for abnormal event recognition, KNN (k Nearest Neighbors) classifier is one of the most commonly used method and has been applied in a variety of cases. KNN Classifier works as follows. First for each one of the training set elements a classification of it is performed based on various neighborhoods. The k value that maximizes the DC of each classification is found, therefore for each training set there corresponds a particular k value which is considered the best available. Afterwards, for each unknown element, the nearest neighbor is found and its k value is assumed (based on the "optimum" k array). Then, the KNN classifier is applied on that test element, using that k value.

The KNN algorithm is given below

1. Choose k number of samples from the training set to generate initial population (p1).
2. Calculate the distance between training samples in each chromosome and testing samples, as fitness value.
3. Choose the chromosome with highest fitness value store it as global maximum (Gmax).
   a. For i = 1 to L do
      i. Perform reproduction
      ii. Apply the crossover operator.
      iii. Perform mutation and get the new population.
      iv. Calculate the local maximum (Lmax).
      v. If Gmax < Lmax then a. Gmax = Lmax;

Output – the chromosome which obtains Gmax has the optimum K-neighbors and the corresponding labels are the classification results.

IV. EXPERIMENTAL RESULT

The experimental result for the above discussed methodology is discussed in this section. Figure 2 represents the overall experimental results. Figure 2 (a) represents input video frames this video is preprocessed and will get Figure 2(b) i.e. absolute difference binary image. Next will apply HOMF for feature extraction from that noise removed image we will get that is as shown in Figure 2(c), next SNN classifier will detect the frame which is having different activity among the frames as shown in Figure 2(d).

ISSN (Print) : 2320 – 3765
ISSN (Online): 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 5, May 2016

V. Conclusion

In this approach, we introduced a new method to detect anomalous events in videos. Moreover, we proposed abnormal event detection in video using saliency value and Histogram of optical flow, for much video surveillance application event detection is most important, so in this approach we used feature descriptor approach based on optical flow information estimated from the scene, called Histograms of Optical Flow Orientation and Magnitude (HOFM). Besides of measuring orientation based on temporal information, the proposed feature descriptor also extracts velocity information provided by the magnitude of the flow vectors. We experimentally compared the performance of the proposed descriptor to the classical histograms of oriented optical flow (HOOF) and achieved great improvements. The results demonstrated the suitability of the proposed HOFM to the anomaly detection problem. Such suitability becomes even more emphasized due to the fact that we are employing a simple nearest neighbor search to classify incoming patterns, as opposed to other approaches that employ very sophisticated classification and modeling techniques.

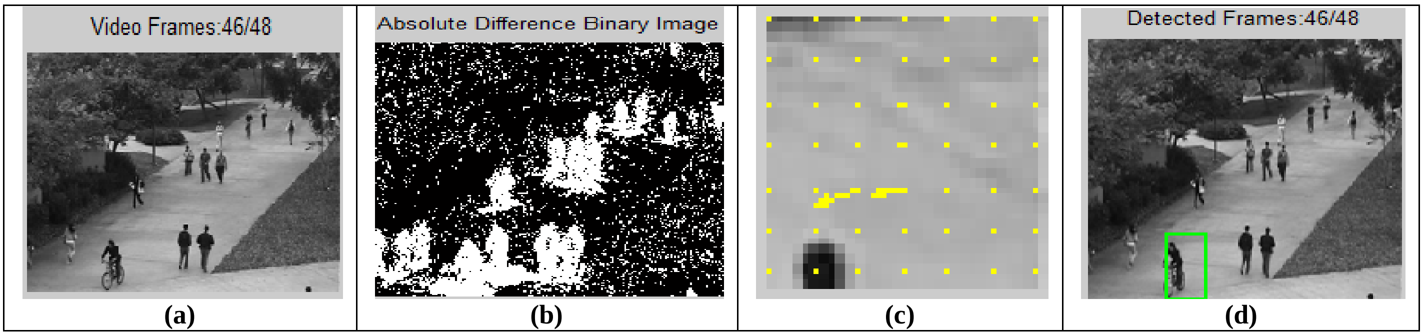

Figure 2: (a) Video Frames; (b) Absolute Difference Binary Image; (c) Activity Detected; (d) Detected Frames.

REFERENCES

[1] Yang Cong, Junsong Yuan and Ji Liu. Mahadik, "Sparse Reconstruction Cost for Abnormal Event Detection", 2011.

[2] Ernesto L. Andrade, Scott Blunsden and Robert B. Fisher, "Modelling Crowd Scenes for Event Detection", 2006.

[3] Vijay Mahadevan, Weixin Li, Viral Bhalodia and Nuno Vasconcelos, "Anomaly Detection in Crowded Scenes", 2010.

[4] Have, Simon Hartmann, Ren, Huamin, Moeslund and Thomas B, "Real-time Multiple Abnormality Detection in Video Data", 2016.

[5] G. Mariem, E. Ridha SSand Z. Mourad "Detection of Abnormal Movements of a Crowd in a Video Scene", Volume 8, Issue 5, 2016.

[6] Vikas Reddy, Conrad Sanderson and Brian C. Lovell, "Improved Anomaly Detection in Crowded Scenes via Cell-based Analysis of Foreground Speed, Size and Texture", pp. 55–61, 2011.

[7] Kristjan Greenewald and Alfred O. Hero, "Detection of Anomalous Crowd Behavior Using Spatio-Temporal Multiresolution Model and Kronecker Sum Decompositions ", 2014.

[8] Xiaoshuai Sun, Hongxun Yao, Rongrong Ji, Xianming Liu and Pengfei Xu, "Unsupervised Fast Anomaly Detection in Crowds ", 2011.

[9] Nahum Kiryati, Tammy Riklin Raviv, Yan Ivanchenko and Shay Rochel, "Real-time Abnormal Motion Detection in Surveillance Video", 2008.

[10] Rizwan Chaudhry, Avinash Ravichandran, Gregory Hager and Ren´e Vidal. Mahadik, "Histograms of Oriented Optical Flow and Binet-Cauchy Kernels on Nonlinear Dynamical Systems for the Recognition of Human Actions", 2009.