International Journal of Students' Research In Technology & Management
Vol 5, No 1, January 2017, pp 32-37, ISSN 2321-2543,
https://doi.org/10.18510/ijsrtm.2017.517

ANOMALY DETECTION OF EVENTS IN CROWDED ENVIRONMENT AND STUDY OF VARIOUS BACKGROUND SUBTRACTION METHODS

[1]Meenal Suryakant Vatsaraj, [2]Rajan Vishnu Parab, [3]Prof.D.S.Bade
[1,2]M.E. [Second Year] - Electronics and Telecommunication Engineering, AlamuriRatnamala Institute of Engineering and Technology (ARIET), Thane - Maharashtra, India
[3]Assistant Professor,Vidyalankar Institute of Technology,Wadala East, Mumbai – Maharashtra, India
Email: meenusv513@gmail.com

Article History: Received on 22nd Feb. 2017, Revised on 10th March. 2017, Published on 6th May 2017

Abstract - Anomalous behavior detection and localization in videos of the crowded area that is specific from a dominant pattern are obtained. Appearance and motion information are taken into account to robustly identify different kinds of an anomaly considering a wide range of scenes. Our concept based on histogram of oriented gradients and markov random field easily captures varying dynamic of the crowded environment. Histogram of oriented gradients along with well known markov random field will effectively recognize and characterizes each frame of each scene. Anomaly detection using artificial neural network consist both appearance and motion features which extract within spatio temporal domain of moving pixels that ensures robustness to local noise and thus increases accuracy in detection of a local anomaly with low computational cost. To extract a region of interest we have to subtract background. Background subtraction is done by various methods like Weighted moving mean, Gaussian mixture model, Kernel density estimation.

Keywords *- anomaly; artificial neural network (ANN); background subtraction; computational cost; Gaussian mixture model (GMM); histogram of oriented gradients (HOG); kernel density estimation (KDE); markov random field (MRF); region of interest (ROI); weighted moving mean (WMM).*

I. INTRODUCTION

Anomalous behavior detection is a specific model which gives specific behavior and therefore it is a critical function in the view of security observation system. eg.locationof movements into restricted place such as there shall not be any movement in the object, detection of wrong direction of motion where all objects should represent direction of motion in only one direction and if object is moving in opposite direction must be detected. E.g. given vehicle entered into no entry zone. Thus, surveillance scenario on roads, stations, airports and mall became more difficult as to analyzehuge amount of data keeping main aim of safety. Detecting frames automatically with anomalous behavior or interesting region long stream video has lead much work to analyze the problem in last few years. Numbers of methods are used to detect anomalous behavior. It is very difficult to analyze abnormal behaviorand motions in crowded area by traditional computerized vision method because of barriers, changing crowd density, different unpredictable nature of motions. To overcome computational cost is higher that must be kept within limit. Frames of the crowded scene in real time are as fast as possible to take quick action by security person if something observed which is not usual.

To detect anomalous event region of interest is considered. Region of interest is nothing but foreground can be obtained by background subtraction. Image foreground is required for further processing such as object recognition. Mostly Region of Interest in the images are objects like human faces, car, text which are considered as a foreground. Background subtraction is used to localize the object after the preprocessing like image denoising, enhancing the image and post processing like morphology. When a video is captured from static camera to detect moving objects background subtraction is widely used. The approach consist detecting moving object when compared and obtain different behavior reference frame and current frame which are called as background model or background image. Various methods for background subtraction are Weighted moving mean (WMM), Gaussian mixture model (GMM), Kernel density estimation (KDE).

Anomaly can be defined as a less probability of occurrences in video that is to be observed. Detection of abnormal behavior in crowded areas is spatio temporalchange in appearance and motion. Example is a person driving a bicycle through crowded region. The instantaneous change in velocity such as increase or decrease in speed or slow or fast movement of individual in crowded area which indicates unusual and something dangerous has occurred.

In this approach we propose a method which detects and localize anomaly by means of appearance and motion information. Here we introduce a descriptor which is having histograms of gradients (HOG) which captures appearance in videos. Local binary pattern is usedfortexture classification. When local binary pattern is combined with histogram of gradients it is easy to improve the performance of detection. Thus, we can easily obtain the features successfully by support vector machine and artificial neural network. Anomalies appearing in part of a frame can be captured by using algorithm only on region of interest. Spatial information is used to increase accuracy of detection, results of thevalidation of algorithm are observed on datasets of crowded people, results can be observed on traffic. Thus, method proposed can be generalized to varying traffic conditions.

II. Literature Survey

Anomaly detection technique has growing demand for security purpose.Andrei Zaharescu, Richard Wildes[1], Here they consider three methods.

1) Using spatio temporal oriented energy
2) Inclusion of subset in histogram comparison
3) Event driven processing

First method is used to model behavior by using energy distribution of spatio temporal orientation. This method easily captures full range of natural occurrences of visual space time patterns and been used for anomaly detection. Second method uses thecomparison of new observations with anautomatically obtained model for normal behavior. The third method, in event driven processing portions of the video stream which are having much deviation from the expectedare marked and concentrated as a region of interest and used for computation.

B. E. Moore, M. Shah and S. Wu[2], says in first category dynamics in objects which exist certain pattern is used to build a model which is capable of moving and localizing with different patterns. This method consists three major accepts. First, it defines uniquely particle trajectories to crowded model scenes. Here also propose new coefficient trajectories to model arbitrarily complicated crowds flowing in some directions. Second is a bunch of disorder not changing features that are calculated to use to detect anomalous activity.Finally, the structure having aprobabilistic model which consist Gaussian mixture model to describe probability density function of the normality is used with a certain probability threshold value to identify whether it is normal or abnormal motion and also localize it within the captured frame. This method works for highly dense video in which patterns exist with global motion. It is unable to find local abnormalities which take place in small regions of frames.

According to Y. Cong, J. Yuan, and J. Liu [3], by using dictionary learning method sparse reconstruction cost is calculated to measure normality of the testing sample. This method provides asolutionfor detecting both local and global abnormal events by updating the dictionary incrementally.K. Nishino, L. Kratz[4], to identify similar 3D volumes in the local area KL divergence are used and new prototypes are produced by using 3D Gaussian distribution of spatio temporal gradients. J. Shi, C. Lu, and J. Jia[5], sparse combination method is used instead of class scarcity based methods which greatly decreases the computational cost, obtains very high speed performance at the cost of its accuracy. In this method abnormal events are detected using sparse combination learning method.

According to A. Briassouli, V. Kaltsa, and I. Kompatsiaris, and M. G. Strintzis[7], Histogram of oriented gradients, histogram of oriented swarms with markov random field is used for recognizing frames of images. Here, motion modeling is done using ahistogram of oriented swarm descriptor. Histogram of theoriented gradient is used for dividing images into small frames. The main concept to analyze motions in the idea of the video consist swarm agents (flying over) is use to monitor the movements in the crowded scene. They are used to form histograms of swarm which is used to detect anomaly event with anunderline motion by means of Region of Interest (ROI) analysis. To obtain extremely high complex and stochastic motion information swarm modeling is used to analyze videos of crowd in swarm model agents. Prey that efficiently track the motion (as shown in Figure 1) is used to map the information regarding motion into more informative space. Prey which are tracked by the swarm consist optical flow magnitude values of apixel which are inside Region of Interest Prey does not consist luminance value and their quantity varies in frame equal to the number of Region of Interest. Region of Interest consists a rectangular area around aninterested point which consists fix number of neighboring pixels (n) and m is a number of theframe. Each pixel position in Region of Interest of frame j is equal to the magnitude of O_{ij} where $1 \leq i \leq n$ and $1 \leq j \leq m$. Thus prey position X_p is denoted as

$X_p(t) = O_{ij}$

Where t is time instant.

International Journal of Students' Research In Technology & Management
Vol 5, No 1, January 2017, pp 32-37, ISSN 2321-2543,
https://doi.org/10.18510/ijsrtm.2017.517

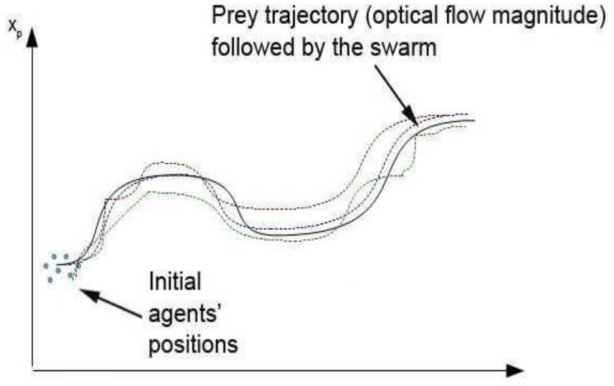


Fig.1.A swarm following prey

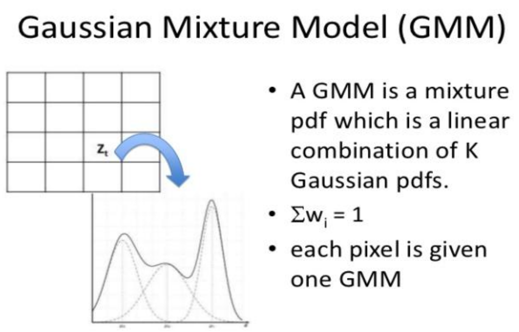
To construct the HOS descriptor, we required agent position which is obtained by prey motion patterns. Agent's position is continuously updated and their new values are taken for spatio temporal location t. Thus triplets over each time window are used to include information such as temporal to obtain final motion descriptor. Resulting feature vector is obtained by averaging triplets of thehistogram of oriented gradients and histogram of theswarm in atime window of m frames.

Background subtraction can be done by using following methods-

1) Weighted moving mean- Weighted moving mean is a type of moving mean which is also called as rolling mean or rolling average in which finite impulse response filter are use to analyze a set of points which are part of a series of observations and then average of different subsets is obtained.

2) Kernel density estimation- While considering no particular assumption to create parameter means the pdf of a random variable in the core is called Kernel density estimation. It is the technique in which estimation ofunknown probability distribution of random variable which is based on sample points taken from distribution.

3) Gaussian mixture model- It isapdf which is calculatedas sum of weighted gaussianpartial density. Gaussian mixture model are mainly used as a probability distribution of continuous measurement in biometric system.



Training data is used to estimate the Gaussian mixture model parameter. The Gaussian mixture model is aprobabilistic model. Data points in Gaussian mixture model are generated from amixture of a finite number of Gaussian distribution with unknown parameter.

III. Methods Of Background Subtraction

1) WEIGHTED MOVING MEAN

A weighted mean is any mean having different multiplying factors which tend to different weights to data corresponding to different positions in the sample window. Thus weighted moving mean is the convolution of the set of points with fix weighting function. Thus we can remove pixels with same features from the background to obtain region of interest in a image frame. The weighted moving mean is denoted as below:

$$\text{WMA}_M = \frac{np_M + (n-1)p_{M-1} + \cdots + 2p_{(M-n+2)} + p_{(M-n+1)}}{n + (n-1) + \cdots + 2 + 1}$$

Where,latest sample is n and theprevious sample is n-1.

As shown in Figure 2 denominator is a triangular number $=\frac{n(n+1)}{2}$. Denominator is sum of individual weights. According to graph weights are decreasing from highest value to zero. Highest weight is for datum point and then it is decreasing to zero.

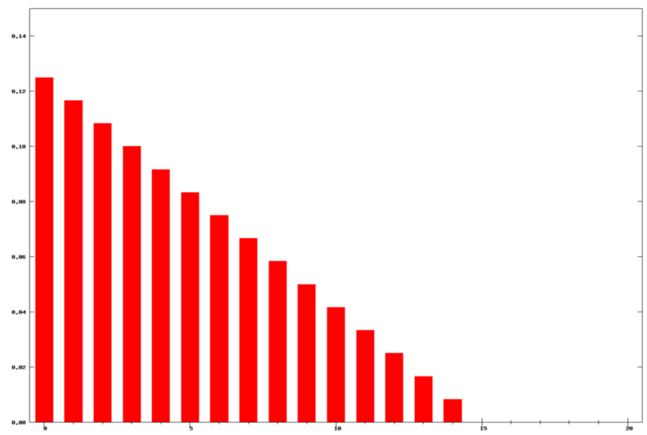


Fig.2.Weighted moving mean

2) KERNEL DENSITY ESTIMATION

Applying an adaptive threshold method based on varying image frames thus effectively subtract the varying background. Kernel density estimation applied for background subtraction when considered no background movement and keeping track of the intensity values of pixels of foreground over the time. The intensity of the pixels can be modeled with a Gaussian kernel with the image noise over the same time is modeled by Gaussian distribution with zero means. Kernel density estimation model gives most recent information related to animage sequence and updates continuously to obtain fast changes in the background (example, flying birds and we can also detect traffic by subtracting waving tree as shown in Figure 3).

Fig.3.Detection for a traffic sequence consisting waving tree using kernel density estimation

By using most recent information, we can estimate this density function at any time for quickly changing intensity distribution of the pixels. Considerc_1, c_2 . . . c_n are the recent samples of the intensity value of the pixels probability density function of a particular pixel at time t, c_t can be estimated with the kernel estimator k as

P(c_t) = $\frac{1}{n}\sum_{i=1}^{n} k(c_t - c_i)$ where, n is number of frames.

To avoid the drawback of manually selecting parameters for adifferent environment non-parametric approach that is Bayesian rule with kernel density estimation is used to decrease the computational cost. Kernel density estimation is modern technique which saves the memory effectively because it requires fewer images to initiate expected backgroundrepresentation.In this technique, thenormalization step is obtained by reducing the few occurrence of less important background which does not seen over a larger time with addition of certain probability.

3) GAUSSIAN MIXTURE MODEL

In non parametric method input space divided into small parts and then probability of function p(x) is estimated to calculate each small part. Thus, dividing the data space to the curse of dimensionality in thestatistical structure we need flexible model which is semi parametric model called as gaussian mixture model.

Gaussian mixture model is demoted as

p(x) = $\sum_{m}^{M} p(x/m)p(m)$

Gaussian mixture model can able to model class condition densities p($x/c\pm$) with higher flexibility with preserving a comprehensive of statistical properties of data in terms of mean, variance etc. gaussian mixture model uses the same model of mixture of gaussian in which each cluster now described with non gaussian random variable s_m .

$p(x/m)$ = δ[x-($a_m s_m + b_m$)]

$s_m$is considered as hidden source which are responsible to generate observation x's given a_m and b_m and consider noiseless condition. s_m's are generated by assuming to be zero mean and unit variance.

IV. Background Subtraction And Hog Feature After Reading Input Frame

Algorithm consist data obtain from automatically extracted Region of Interest instead of considering entire frame. Therefore in our process we concentrate on pixels which contain information relevant to events appearing in Region of Interest. Therefore computational cost is lower.

We show Region of Interest as a rectangular area around each interesting point with a fix size and also defining point of interest as a dense grid on the foreground of the region. Foreground grid is updated continuously. The interest points and resulting Region of Interest are to be considered informative only when Region of Interest at least 60% of the motion feature. Otherwise, this Region of Interest points assumesnoisy and discarded. If Region of Interest consist fewer moving pixels but not reaching limit of 60% then they are considered as motionless, noisy data and ignored.

Appearance modeling histograms of oriented gradients are used to extract appearance features. It uses gray scale images invariant to illumination. Histograms of oriented gradients effectively distinguish local edges and gradient features that can predict variations in appearance in thelocal area of the image. Histograms of oriented gradients are used as it is direction invariant. Histogram of oriented gradients has direction invariant appearance feature. Histogram descriptor is applied to Region of Interest block which are tracked over the time. Therefore final histogram of gradient descriptor of each block consist temporal information process of computation as below:

Each block k is divided into 2x2 matrixes. Therefore spatial location information is divided into 4regions, and local noise is minimized. Example, if a noise present in Region of Interest is divided into 2x2 matrix, then noise is limited to only one cell instead of whole area. Therefore, it leads to less noisy appearance descriptor.

A weighted histogram of gradients is created for each by using 9 beans.

Histogram of oriented gradients of c^{th} cell (1 ≤ c ≤ 4) in block k of frame j which is represented as HOG_j^k(c) having dimension 1x9.

After that every histogram is normalized, and resulting cell histograms of 4 cells are averaged and thus forming 1x36 blocks descriptor.

Finally appearance descriptor is nothing but averaging of 3 frames for each cell c in block k (as shown in Figure 4).

$HOG_{j,j+2}^k$(c) = E[HOG_j^k(c),HOG_{j+1}^k(c),HOG_{j+2}^k(c)]

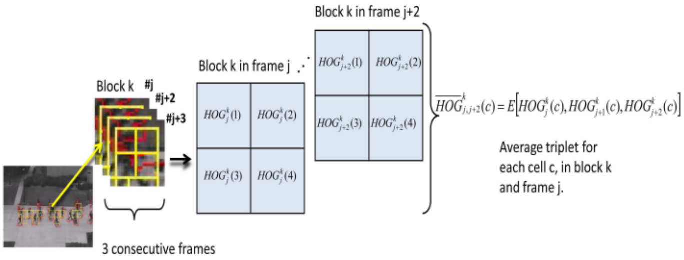


Fig.4.Extraction of HOG

V. Markov Random Field Model

Markov Random Field (MRF) consistaset of random variables which are represented as unidirectional graphical model with markov networks. Random field is considered as markov random field which is same as Bayesian network by means of representation of their dependency. In Bayesian networkdependencies are directed and acyclic while in markov networks these are undirected and cyclic.

VI. Anomaly Detection Using Artificial Neural Network

Artificial neural network is the best example of machine learning which consist simple mathematical or neural network. Set of learning algorithm consists evaluation function of many inputs which are unknown. A collection of interconnected neurons are called artificial neural network. It consists three parameters such as inter connection series between different neuron layers. Different weightage use to update interconnection learning process and finally activation process. There are many connected neurons which also get activated simultaneously.Finally, anomalies in videos of thecrowded area are detected by using theartificial neural network. The accuracy of detecting anomalies is increased due to theuse of theartificial neural network.

VII. Conclusion

In present survey paper, we illustrate various methods and techniques which are used to detect anomalous behavior in thecrowded scene. A novel framework for anomalous behavior detection in adifferent scene which is recorded from astatic surveillance camera is studied. Various methods are observed to exploit the extraction of robust motion, and appearance features characteristics which effectively describes each scene. Various methods of background subtraction are studied such as Weighted moving mean, Gaussian mixture model, Kernel density estimation. The Weighted moving mean is the convolution of the set of points with fixed weighing function. Therefore we can remove pixels with thesame feature from background to obtain theregion of interest in the image. Kernel density estimation is an adaptive threshold method based on varying image frames thus effectively subtract the varying background by considering no background movement and keeping the values of intensity of pixels of foreground over the range. The Gaussian mixture model is a non-parametric method in which input image is divided into small parts and probability of function is estimated for each part. Gaussian mixture model uses class condition densities with higher flexibility which preserves statistical properties of data regarding mean and variance. Anomaly detsssection using classifier support vector machine (SVM) is better, but efficiency can be improved by using classifier artificial neural network (ANN).

References

[1] Andrei Zaharescu, Richard Wildes, "Anomalous Behaviour Detection Using Spatiotemporal Oriented Energies, Subset Inclusion Histogram Comparison and Event-Driven Processing," CCV 2010, Part I, LNCS 6311, pp- 563–576, 2010.c Springer-Verlag Berlin Heidelberg 2010.

[2] B. E. Moore, M. Shah and S. Wu, " Chaoticinvariants of lagrangian particle trajectories for anomaly detection in crowded scenes," in Proc. IEEE Conference. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2010, pp- 2054–2060.

[3] Y. Cong, J. Yuan, and J. Liu, " Sparse reconstruction cost for abnormal event detection," in Proc. IEEE Conference Computer Vis. PatternRecognit. (CVPR), Jun. 2011, pp. 3449–3456.

[4] K. Nishino,L.Kratz , "Anomaly detection in extremely crowded scenes using spatio-temporal motion pattern models," in Proc.IEEE Conference. Computer Vis. Pattern Recognit. (CVPR), Jun. 2009, pp. 1446–1453.

[5] J. Shi, C. Lu, and J. Jia, "Abnormal event detection at 150 FPS in MATLAB," in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Dec. 2013, pp. 2720–2727.

[6] Jeisung Lee and Mignon Park, "An Adaptive Background Subtraction Method Based on Kernel Density Estimation," Sensors 2012, 12, 12279-12300; https://doi.org/10.3390/s120912279

[7] A. Briassouli, V. Kaltsa, and I. Kompatsiaris, and M. G. Strintzis, "Swarm based motion features for anomaly detection in crowds," in Proc. IEEEInt. Conference Image Process. (ICIP), Oct. 2014, pp. 2353–2357.

[8] A. Del Bue,M.Cristani, R. Raghavendra, and V. Murino, " Optimizing interaction force for global anomaly detection in crowded scenes," in Proc. IEEE Int. Conference

Computer Vis. Workshops (ICCVW), Nov. 2011, pp. 136–143.

[9] M. Shah, A. Gritai ,andA. Basharat, ," Learning object motion patterns for anomaly detection and improved object detection," in Proc. IEEEConference Computer Vis. Pattern Recognit. (CVPR), Jun. 2008, pp. 1–8.

[10] N. Conci ,H.Ullah, " Crowd motion segmentation and anomaly detection via multi-label optimization," in Proc. IEEE Int. Conference. PatternRecognit. Workshop (ICPRW), Nov. 2012.

[11] C. Fookes, D. Ryan,S.Sridharan and S. Denman, "Textures of optical flow for real-time anomaly detection in crowds," in Proc. 8th IEEEInt. Conference. Adv. Video Signal-Based Surveill. (AVSS), Aug./Sep. 2011, pp. 230–235.

[12] T.Thenmozhi, Dr.A.M.Kalpana, "Detection and tracking of moving object based on background subtraction," International Conference "Recent Innovation in Science, Technology and Management" (ICRISTM-16)at Indian Federation of United Nations Associations, New Delhi, India on 12 June 2016ISBN: 978-81-932712-3-0.

[13] R. Mehran, A. Oyama, and M. Shah, "Abnormal crowd behavior detection using social force model," in Proc. IEEE Conference ComputerVis.PatternRecognit. (CVPR), Jun. 2009, pp. 935–942.https://doi.org/10.1109/cvpr.2009.5206641

[14] V. Reddy, C. Sanderson, and B. C. Lovell, "Improved anomaly detection in crowded scenes via cell-based analysis of foreground speed, size and texture," in Proc. IEEE Conf. Comput. Vis. Pattern Recognition.Workshops (CVPRW), June. 2011, pp- 55–61.https://doi.org/10.1109/cvprw.2011.5981799

[15] V. Mahadevan, W. Li, V. Bhalodia, and N. Vasconcelos, "Anomaly detection in crowded scenes," in Proc. IEEE Conference Computer Vis. PatternRecognit. (CVPR), June. 2010, pp- 1975–1981.https://doi.org/10.1109/cvpr.2010.5539872