# Unsupervised Video Surveillance for Anomaly Detection of Street Traffic

Muhammad Umer Farooq, Najeed Ahmed Khan
Computer Science & IT department
NED University of Engineering & Technology
Karachi, Pakistan

Mir Shabbar Ali
Urban and Infrastructure Engineering department
NED University of Engineering & Technology
Karachi, Pakistan

*Abstract*—**Intelligent transportation systems enables the analysis of large multidimensional street traffic data to detect pattern and anomaly, which otherwise is a difficult task. Advancement in computer vision makes great contribution in the progress of video based traffic surveillance system. But still there are some challenges which need to be solved like objects occlusion, behavior of objects. This paper developed a novel framework which explores multidimensional data of road traffic to analyze different patterns of traffic and anomaly detection. This framework is implemented on road traffic dataset collected from different areas of the city.**

*Keywords*—*Kalman filter; Gaussian mixture model; DBSCAN clustering; similarity matrix; occlusion; computer vision; traffic surveillance; Intelligent Transport Systems (ITS)*

## I. INTRODUCTION

Major problem arises with the rapid development and expansion of urbanization includes traffic congestion, accidents, traffic violations, etc. These factors have great influence on human life. Over last three decades Intelligent Transport Systems (ITS) is playing a vital role in optimization of traffic flow, avoid traffic congestion and other issues of road traffic. ITS collect and analyze data to extract information which can used to detect and prevent accidents and in this way damages to properties can be minimized and human life can be saved.

Although great progress has been made on video-based traffic surveillance but still the detection and prediction of accidents is a very complicated task and researchers are still facing various difficulties and challenges for practical ITS applications like vehicle all day surveillance, occlusion, pose variation, different resolutions and vehicle behavior understanding [1].

The accurate detection of vehicles can fulfill the strive to have better traffic data and a reliable traffic surveillance. Image processing of traffic videos is an effective approach to track vehicles and as a consequence various factors like velocity, change of lane, vehicle trajectories can be efficiently analyzed. Measurement of accurate density is possible if trajectories of vehicles are observed over a road till certain distance. Monitoring of vehicle trajectories is one of the key points to predict any incident. Detecting anomaly behavior is an important aspect of video surveillance for several applications. In the last few years various researchers focused on this issue [1], [2].

Patterns of trajectories and abnormal behaviors recorded in videos can decipher anomalies. With the development in visual surveillance, accuracy in tracking of moving objects becomes high which provides reliable source of detection of incident. Two approaches, parametric and non-parametric are used to identify abnormal behaviors on road [3]-[6]. The detection of abnormal events in parametric method is performed on the basis of feature extraction from the observed data [3]. The tracking algorithms can be used to extract visual features for example appearance, position, speed etc. Statistics of the data provides the real picture of normal and abnormal patterns of the vehicle activities [4]-[6].

Based on the challenge faced by present traffic surveillance system, a novel framework has been developed here which explores multidimensional data of road traffic to analyze different pattern of traffic and anomaly detection. This framework is implemented on real road traffic data set collected from different areas of the city.

This paper is arranged as follows. Section 2 presents background and related work being performed in this area. Section 3 deals with proposed framework for detection and tracking of objects, extraction of features, preprocessing, similarity based trajectory clustering and anomaly detection. Section 4 discusses the results and experimental analysis. Finally the conclusion of paper is mentioned in Section 5.

## II. LITERATURE REVIEW

In unsupervised learning, the model of data is hidden and it's up to the system to learn it. The data is differentiated into different classes using their statistical property. The results of these approaches are dependent on the data it's being applied on. The advantage of unsupervised approaches is the non-essentiality of any sort of special training. Clustering of trajectories of moving objects is needed to detect abnormal activities. Identification of outlier by spatial measurements was done to study vehicle behaviors from trajectory data [7].

In [8], Fu et al. propose a hierarchical clustering framework to cluster trajectories using pairwise similarities. After trajectory preprocessing the similarity matrix is constructed based on the pairwise distance of sample trajectories. Spectral clustering is then applied on the similarity matrices. A second layer of clustering is done on the already clustered trajectories to achieve finer classification. Abnormal trajectories are determined using point-to-point distance between the trajectory in question and the template

trajectory. Spectral clustering only slightly improves the results obtained from the popular K-means approach but requires more work.

Framework based on temporal study was given to detect outliers on the basis of changes in trajectories [9]. Pan et al. in [10] develop a traffic surveillance system for detecting vehicle flows. They use a background subtraction method for separating the vehicle from the background. Using an edge detection technique a region of interest is created around the vehicle. They also propose a lane-dividing algorithm for counting the number of passing vehicles. In [11], Basharat et al. propose a scheme for identifying vehicular anomalies. Pixel level pdfs of motion and sizes of objects in static camera environments are modeled from the tracks. Each pdf is modeled as a multivariate Gaussian mixture model of motion and size of the object. The authors propose an unsupervised Expectation Maximization based algorithm for learning the GMMs. Their training phase generates a scene model consisting of all the GMM parameters of all the pixels with sufficient training data. This scene model is used to detect anomalous patterns which differ from the patterns observed in the training data.

Piciarelli [12] suggested a trajectory based clustering method. Anomalies in trajectories were extracted through a tree formed from a series of clusters. Anomalous trajectories are those having less probability when compared with the tree. Vehicles moving in a wrong direction can easily be identified by this method but this approach fails to estimate speed of vehicle. Correspondingly, movement through unusual path can be determined by path frequency and time used [13]. Although variations in direction and speed were estimated but small changes in motion cannot be detected.

The main goal is automatically determination of suitable number of trajectory clusters especially when there are large quantities of trajectories. Features of local trajectories were examined by vector quantization [14]. Wang et al. [15] used spectral clustering to perform trajectory clustering for detection and prediction of anomalies. Expectation-maximization algorithm was proposed by Vasquez and Fraichard [16] for computation of trajectory similarities. The authors [17] compile a survey of activity definitions based on trajectory data. Their compilation covers diverse sets of real time applications such as behavioral analysis, anomaly detection, object interaction etc. Using the motion information of the objects, topological scene description is built. This scene topology is automatically learnt and distinguished by points of interest and motion characterized by activity paths. In most of the methods mentioned above, the anomalies are detected by first categorizing the normal events which are far more common compared to the anomalous events. This categorization is done by extracting certain features of the events. These features can be in terms of image frames containing the activity being modeled as GMMs or in terms or features containing spatiotemporal information. These extracted features are categorized using one of the many clustering algorithms or machine learning algorithms. In [1] the challenges in traffic video surveillance are discussed in detail. The framework developed here deals with problems like occlusion, behavior of objects.

## III. METHODOLOGY

Vehicle pattern and anomaly detection system is shown in Fig. 1. The proposed frame work is based on unsupervised techniques i.e. there is no requirement for learning stage. Approach based on real time trajectory has been established for the recognition of pattern and anomalous behavior in video sequences. An automated method is suggested for detection of vehicle pattern and anomaly detection. Vehicle trajectories are used for detection of pattern in the field of a camera. Surveillance videos are recorded from defense Khayaban-e-Ittehad and university road near FUUAST, Karachi.

The video series are recorded with various resolutions, different time frame, and under varying conditions of weather.

The developed framework includes the following steps:

- Detection and tracking of objects
- Extraction of feature
- Preprocessing
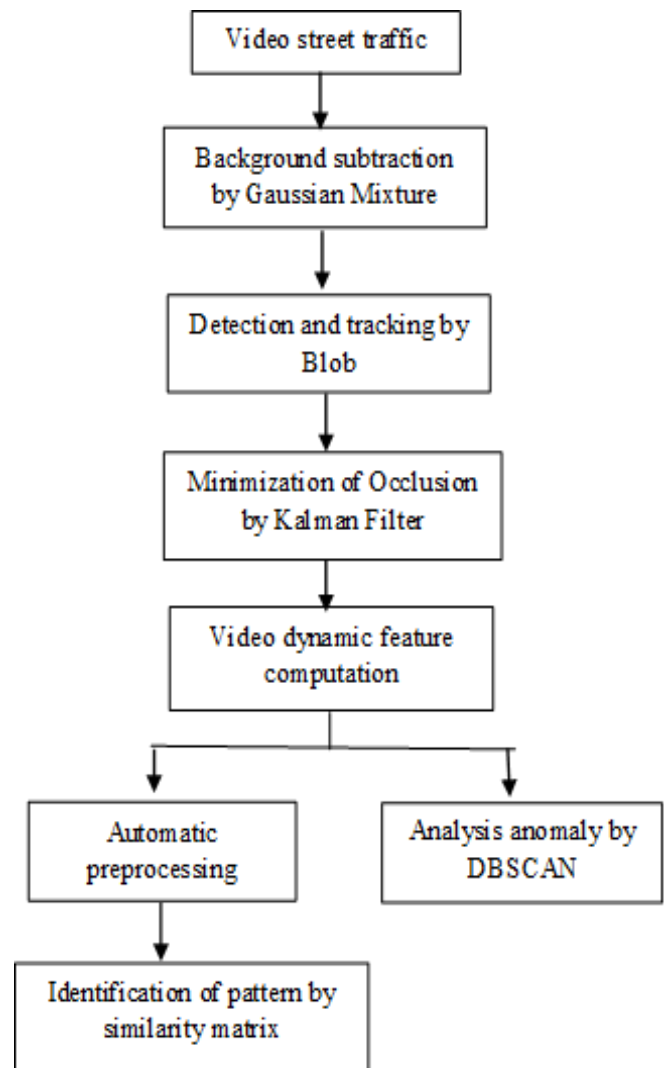- Similarity based trajectory clustering
- Anomaly detection.



Fig. 1.   Flow diagram of developed framework.

### A. Detection and Tracking of Objects

Background subtraction technique is implemented in the proposed system. It is essential to efficiently detect foreground because the consequent process like tracking of objects, extraction of features, etc. are mainly dependent on the results of foreground detection technique. Gaussian Mixture Model (GMM) is employed for background subtraction and detection of foreground objects [23] is performed by using Expected Maximization (EM).

Effect of alterations in the illuminations is minimized by applying various distributions for modeling the pixel values. The probability of pixel having the value of Xt is found to be:

$$P(X_t) = \sum_{i=1}^{k} w_{i,j} \eta \left( X_t, \mu_{i,t}, \Sigma_{i,t} \right) \qquad (1)$$

Where, $w_{i,j}$ is weight, $\mu_{i,t}$ represents mean and $\sum_{i,t}$ symbolizes Co-variance of the ith object [18]. Probability Density Function (PDF) [19] can be formulated as:

$$\eta(X, \mu, \Sigma) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} e^{-\frac{1}{2}(X-\mu)^T \Sigma^{-1}(X-\mu)} \qquad (2)$$

Kalman filter is applied for the indication of each track. American scientist R.E. Kalman established Kalman filtering algorithm for discrete random systems based on Wiener filtering, also can be used for linear continuous time systems [20]. General discrete systems can be expressed as in formula [21].

$$X_{(i+1)} = Q_{(i+1,i)} X_{(i)} + J_{(i+1,i)} s_{(i)} \qquad (3)$$

$$R_{(i+1)} = H_{(i+1)} X_{(i+1)} + c_{(i+1)} \qquad (4)$$

In (3) and (4), $X_{(i)}$ is the p-dimensional state vector, $R_{(i)}$ is the v-dimensional observation vector, s(i) is the y-dimensional state vector, $c_{(i+1)}$ is the z-dimensional noise vector measurement. $Q_{(i+1,i)}$ is the state transition matrix from the time i to time (i+1), $J_{(i+1,i)}$ is the incentive transfer matrix from time i to time (i+1), $H_{(i+1)}$ predictive output matrix for ( i+1) moments.

This filter is applied to locate track in every frame, it also decided the probability of each detection which is being allocated to each track. Maintenance of track is an essential point. In selected frames, some of the detections are assigned to tracks, whereas others remained unassigned. The updating of assigned tracks is performed by using corresponding detections.

### B. Extraction of Features

The features used in proposed framework include objects position, trajectory, orientation, visible age in the scene and invisible count. Table I show these features in terms of x-axis, y-axis, width, height, age, visible and invisible count. Unassigned tracks are considered invisible and a new track is originates by unassigned detection. Each of the tracks retains sum of the number of successive frames, where it kept unassigned. If the count surpasses a definite threshold, it is assumed that the object move away from the view of camera and the track is deleted.

Following the detection of foreground, object of foreground is tracked from frame to frame with the help of blob tracking. Centroid of two dimensional areas is estimated by average of bounding box points. Although occlusion can occur between vehicles, but centroid can efficiently distinguish them and the tracking continues productively. After tracking, the set of points are recorded as trajectories of moving objects or vehicles given in Table I.

### C. Preprocessing

There is a need of preprocessing to obtain fix length of trajectory data and the algorithm used for this purpose is as follows:

*Step1:* Count the number of coordinates of each trajectory.

*Step2:* Calculate average (number of coordinates) for each trajectory to convert it into fixed length.

Fixed Coordinates=mode function (the top five modes have been selected and there average was calculated, therefore some of them move towards higher value and other get decreased)

*Step3:* For 1 to n Trajectories

If (Trajectory (Coordinates) < Fixed Coordinates)
//if no. of coordinates are less

Increase it coordinates to Fixed Coordinates
Else
//if no. of coordinates are greater
Take Fixed Coordinates from the trajectory and delete the extra co ordinates
End for//

### D. Similarity based Trajectory Clustering

Euclidean distance is not only used in data mining but it also helps to analyze trajectory. Due to its simplicity it is important in data mining of moving objects. Addition to that large trajectory data can be easily handled due to its linear time complexity. The sampling points must have same dimension. However, in real situation, the sampling time of each point in trajectory may not be same. Trajectory data is fixed to same length by preprocessing algorithm (Section III.C). In the developed framework, Euclidean distance is used to calculate clusters in trajectory data set.

Similarity base clustering algorithm:

- Threshold value is the criteria for which the similarities of trajectories are being defined.
- Calculation of difference of every new trajectory with all the clusters and then select that cluster which has minimum difference with trajectory (i.e. greater similarity).

N=number of Clusters
n=no. of Trajectories

**Algorithm:**
*Step1:*
Initially start with two trajectories
Calculate point to point distant metric between two trajectories
If (Trajectories<threshold)
        //Similar
      Merge them and calculate mean of two trajectories
Else
        // Distinct Trajectories
      Number of Clusters (N) will be 2 + mean of Existing Clusters
//if else end
*Step2:*
For trajectory 3 to every new upcoming trajectory (n Trajectories)
      Calculate point to point distant metric between two trajectories
        If (Trajectories<threshold)
           //Similar
           Merge it into similar cluster and calculate the new mean of selected Cluster trajectories
        Else
           // Distinct Trajectories
           N=N+1
           New Trajectory will be the mean of new Cluster
        If else end //
for end//

*E. Anomaly Detection*

Detection of anomaly is the identification of those points in dataset that do not follow the normal pattern or other points in a dataset. Outliers indicate the abnormal behavior of vehicle. In this research DBSCAN clustering which is an unsupervised technique is being used to detect anomaly. The main goal of this technique is to separate regular and irregular patterns in the video surveillance. DBSCAN deals with the density of points to perform clustering, as a consequence regions of high and low density separate. This algorithm can handles clusters of various size and shape [22].

The algorithm can effectively handle noise and a large data set. The main emphasis of DBSCAN is to access density with two critical parameters, i.e. Eps and MinPts. Eps means neighbors of point p and MinPts is minimum number of points.

$$Eps(p) = \{ q \in D \,|\, dis\tan ce\,(p,q) \le Eps \} \qquad (5)$$

$$p \in Eps\,(q) \qquad (6)$$

$$|\,Eps\,(q)\,| \ge Minpts \qquad (7)$$

Point p is density accessible from q, as given in (5) and (6).

For finding a cluster, DBSCAN initiates with a random point p and accesses all density reachable points from p applying region queries for p as well as for p's direct and indirect neighbors if necessary. This technique results in a cluster if p is a core point. However, if p is not a core point, no points are supposedly density reachable from p and thus DBSCAN assigns p to noise and continues the same technique with next point. The algorithm ends when all pints are either assigned to a cluster or to noise.

## IV. RESULTS AND DISCUSSION

In order to check the reliability and accuracy of a system, there is a need of good annotated dataset. A verified annotated dataset can assumed to be a ground truth to find the consistency of results of the system under consideration. As this data is not available publicly, therefore, observations of a survey person (video observer) have been set to be a ground truth. Traffic Surveillance videos are recorded from defense Khayaban-e-Ittehad and university road near FUUAST, Karachi. The results after applying Gaussian Mixture Model and Kalman filter on these video are shown in Fig. 2(a)-(d). The dataset contains 10 surveillance videos having 698 total trajectories. The examples of few objects in the dataset are given in Table I. From the table features of interest of the objects are computed for the anomaly classification.

Accuracy of the system is not affected by duration of videos. Abnormal events are represented by DBSCAN clustering using features age and angle, results are shown in Fig. 3(a), (b). The graphs (Fig. 3 (a), (b)) demonstrate that objects having age less than 10 and objects turning with angle greater than 56º are anomalous in the objects data set since the focus is on sharp turn and less age anomalous events in this context.

Different patterns of trajectories are calculated for similarity matching algorithm using fix length trajectory data set. In total three patterns were observed, most of the vehicles go straight on the road, some moves towards left and few took U-turn. The result of the proposed system is found to be 90% accurate when compared with the observations of the survey person (Fig. 4).

TABLE II.     EXAMPLES OF FEW OBJECTS IN THE DATASET

| Frame no. | Object no. | x-axis | y-axis | Width | Height | Visible Age | Visible count | Invisible count | Angle (rad) |
|---|---|---|---|---|---|---|---|---|---|
| 4 | 1 | 281 | 221 | 28 | 33 | 1 | 1 | 0 | 39 |
| 5 | 1 | 279 | 209 | 35 | 69 | 2 | 2 | 0 | 33 |
| 5 | 2 | 229 | 227 | 33 | 25 | 1 | 1 | 0 | 44 |
| 6 | 1 | 278 | 209 | 36 | 76 | 3 | 3 | 0 | 0 |
| 6 | 2 | 229 | 228 | 33 | 26 | 2 | 2 | 0 | 46 |
| 7 | 1 | 275 | 211 | 39 | 78 | 4 | 4 | 0 | 0 |
| 7 | 2 | 225 | 232 | 36 | 27 | 3 | 3 | 0 | 31 |
| 7 | 3 | 161 | 375 | 38 | 34 | 1 | 1 | 0 | 65 |
| 7 | 4 | 270 | 165 | 19 | 46 | 1 | 1 | 0 | 34 |
| 8 | 2 | 223 | 234 | 37 | 27 | 4 | 4 | 0 | 4 |
| 9 | 2 | 220 | 233 | 38 | 33 | 5 | 5 | 0 | 30 |
| 9 | 3 | 148 | 394 | 41 | 41 | 3 | 3 | 0 | 43 |
| 10 | 1 | 265 | 195 | 47 | 102 | 7 | 7 | 0 | 5 |
| 10 | 2 | 218 | 231 | 40 | 37 | 6 | 6 | 0 | 34 |
| 10 | 3 | 141 | 356 | 45 | 88 | 4 | 4 | 0 | 29 |
| 12 | 2 | 213 | 257 | 42 | 18 | 8 | 8 | 0 | 37 |
| 12 | 3 | 123 | 380 | 52 | 101 | 6 | 6 | 0 | 18 |
| 13 | 1 | 252 | 247 | 50 | 62 | 10 | 10 | 0 | 34 |
| 13 | 2 | 210 | 236 | 42 | 45 | 9 | 9 | 0 | 22 |
| 13 | 5 | 262 | 189 | 22 | 32 | 1 | 1 | 0 | 37 |
| 14 | 2 | 209 | 236 | 42 | 47 | 10 | 10 | 0 | 22 |
| 14 | 5 | 269 | 248 | 31 | 29 | 2 | 2 | 0 | 41 |
| 15 | 1 | 248 | 251 | 62 | 65 | 12 | 12 | 0 | 6 |



a) Moving object detection     b) Background subtraction

c) Moving object detection     d) Background subtraction
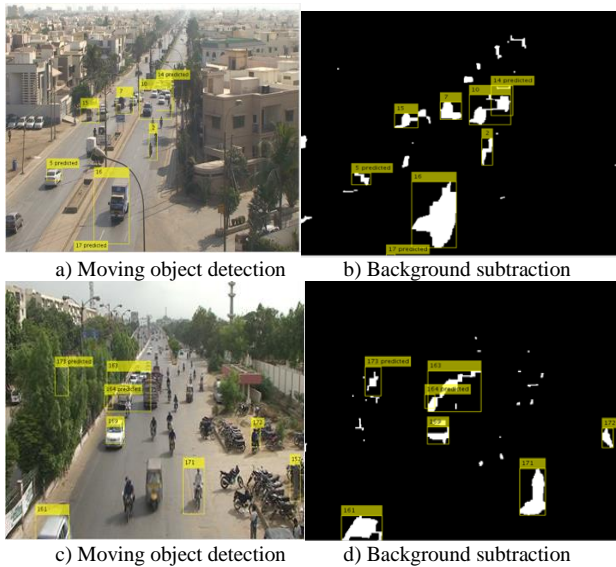
Fig. 2.     (a)-(d): Object detection and tracking of traffic at two different positions.
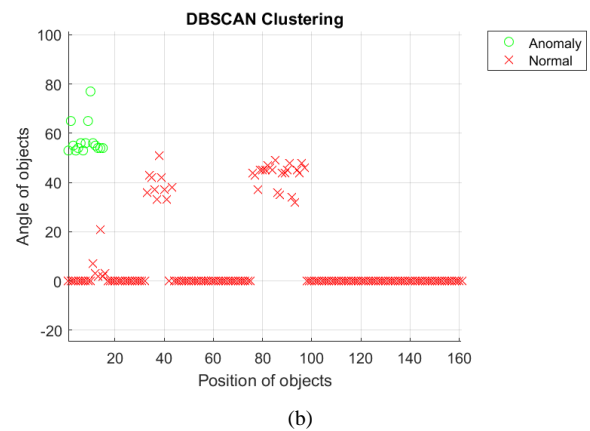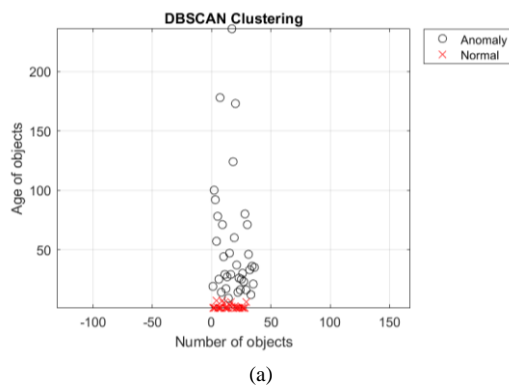


(a)



(b)

Fig. 3.     (a) Anomaly of objects (ageless the10); (b) Anomaly of objects due to speedy turn (angle more the 56º).
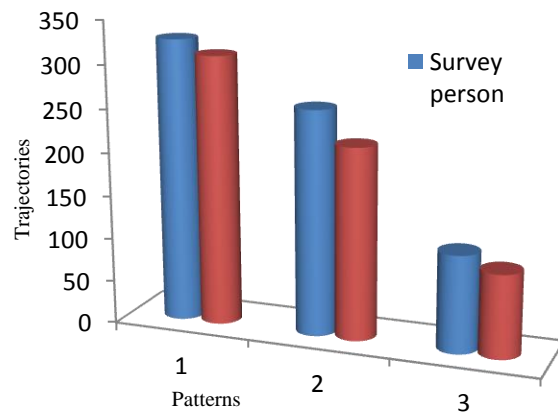


Fig. 4.     Comparison of results of survey person and algorithm.

## V. Conclusion

The video-based traffic surveillance system is playing efficient role in intelligent transport system. Well organized association is required between human and machine in order to decipher valuable information from big, multidimensional and diverse data set. These systems capture traffic scenes, evaluate the information obtained and finally differentiate between normal and abnormal activities. In this paper a novel frame work, for detection of pattern and anomaly, has been developed. The main purpose of this framework is to extract the dynamic feature of vehicle, investigate vehicle pattern and behavior on the road and detect anomaly by using single camera node. The framework is in such a sequence that capturing and understanding of vehicle behavior over the entire road becomes effective. Additional achievements in this research area will give more efficient ITS facilities for extensive applications.

### References

[1] Tian B, Morris BT, Tang M, Liu Y, Yao Y, Gou C, Shen D, Tang S. Hierarchical and networked vehicle surveillance in ITS: a survey. IEEE transactions on intelligent transportation systems. 2015 Apr;16(2):557-80.

[2] Popoola, O.P. and K. Wang, Video-based abnormal human behavior recognition—A review. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 2012. 42(6): p. 865-878.

[3] Li, T., et al., Crowded scene analysis: A survey. IEEE transactions on circuits and systems for video technology, 2015. 25(3): p. 367-386.

[4] Jiang, F., et al., DETECTING ANOMALOUS TRAJECTORIES FROM HIGHWAY TRAFFIC DATA.

[5] Ivanov, Y.A. and A.F. Bobick, Recognition of visual activities and interactions by stochastic parsing. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000. 22(8): p. 852-872.

[6] Wren, C.R., et al., Pfinder: Real-time tracking of the human body. IEEE Transactions on pattern analysis and machine intelligence, 1997. 19(7): p. 780-785.

[7] Haritaoglu, I., D. Harwood, and L.S. Davis, W/sup 4: real-time surveillance of people and their activities. IEEE Transactions on pattern analysis and machine intelligence, 2000. 22(8): p. 809-830.

[8] Sun, P. and S. Chawla. On local spatial outliers. in Data Mining, 2004. ICDM'04. Fourth IEEE International Conference on. 2004. IEEE.

[9] Fu, Z., W. Hu, and T. Tan. Similarity based vehicle trajectory clustering and anomaly detection. in Image Processing, 2005. ICIP 2005. IEEE International Conference on. 2005. IEEE.

[10] Li, X., et al. Temporal outlier detection in vehicle traffic data. in Data Engineering, 2009. ICDE'09. IEEE 25th International Conference on. 2009. IEEE.

[11] Pan, X., Y. Guo, and A. Men. Traffic surveillance system for vehicle flow detection. in Computer Modeling and Simulation, 2010. ICCMS'10. Second International Conference on. 2010. IEEE.

[12] Basharat, A., A. Gritai, and M. Shah. Learning object motion patterns for anomaly detection and improved object detection. in Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on. 2008. IEEE.

[13] Piciarelli, C. and G.L. Foresti, On-line trajectory clustering for anomalous events detection. Pattern Recognition Letters, 2006. 27(15): p. 1835-1842.

[14] Patino, L., J. Ferryman, and C. Beleznai. Abnormal behaviour detection on queue analysis from stereo cameras. in Advanced Video and Signal Based Surveillance (AVSS), 2015 12th IEEE International Conference on. 2015. IEEE.

[15] Stauffer, C. and W.E.L. Grimson, Learning patterns of activity using real-time tracking. IEEE Transactions on pattern analysis and machine intelligence, 2000. 22(8): p. 747-757.

[16] Vasquez, D. and T. Fraichard. Motion prediction for moving objects: a statistical approach. in Robotics and Automation, 2004. Proceedings. ICRA'04. 2004 IEEE International Conference on. 2004. IEEE.

[17] Wang, X., K. Tieu, and E. Grimson, Learning semantic scene models by trajectory analysis. Computer Vision–ECCV 2006, 2006: p. 110-123.

[18] Morris, B.T. and M.M. Trivedi, A survey of vision-based trajectory learning and analysis for surveillance. IEEE transactions on circuits and systems for video technology, 2008. 18(8): p. 1114-1127.

[19] Turdu, D. and H. Erdogan. Improved post-processing for GMM based adaptive background modeling. in Computer and information sciences, 2007. iscis 2007. 22nd international symposium on. 2007. IEEE.

[20] Gong, X. and Z. Li. Efficient Foreground Segmentation Using an Image Matting Technology. in Computational and Information Sciences (ICCIS), 2013 Fifth International Conference on. 2013. IEEE.

[21] Dou, J., et al. Research on electric field localization algorithm based on Kalman Filter. in Chinese Automation Congress (CAC), 2015. 2015. IEEE.

[22] Andrews, A., Kalman filtering: theory and practice using MATLAB. 2001, JohnWiley&Sons.

[23] Najeed K., David C. H., Unsupervised learning of Appearance classes from Video, NED University Journal of Research – Applied Sciences, (ISSN: 1023-3873), Vol. 11, No.2, 2014, pp 15-25.