International Journal of Advances in Applied Computational Intelligence

♦ ISSN: 2833-5600 Vol: 08, No. 01, 2026 ♦ PP. 01–04
Article DOI: https://doi.org/10.54216/IJAACI.080101

FULL LENGTH ARTICLE

Motion Vector–Guided Object Detection and Tracking for Smart Surveillance Systems

V. Vinothini[1,*] **N. Karthika Devi[1]** **R. Nazreen Roja[1]** **G. Mahendran[2]**

[1] Assistant Professor, Syed Ammal Engineering College, Tamil Nadu, India
[2] Professor, Syed Ammal Engineering College, Tamil Nadu, India

Emails: vinovijay028@gmail.com · karthijanu4@syedengg.ac.in · nazreenroja@syedengg.ac.in · ecehod@syedengg.ac.in

Received: January 03, 2026 **Revised:** February 03 2026 **Accepted:** March 01, 2026 ⋆ Corresponding author

ABSTRACT

Multiple moving object detection and tracking are challenging roles in many computer vision applications such as object navigation and human identification. Object tracking is one of the key challenges for securing against crime, supporting public safety, and enabling effective traffic management systems. In video surveillance applications, detection of multiple moving vehicles from video is the major task for tracking and understanding the behaviour of the detected objects. Performance of object detection algorithms is degraded by factors such as fog or haze, occlusion, dynamic background, poor illumination, and low resolution. Fog is one of the major bottlenecks of video surveillance applications. The proposed Dark Channel Prior algorithm using guided filter (GDCP) is adapted for fog removal. The Gaussian Mixture Model (GMM) is proposed for detecting multiple moving objects, and features are extracted from the detected objects using Motion Vector Estimation. The K-Nearest Neighbor algorithm is used for tracking the moving objects (vehicles) using the detected features. Efficiency is improved due to the adoption of the proposed fog-removal algorithm and feature extraction for effective tracking. There are wide varieties of applications in moving object detection and tracking.

Keywords: Gaussian Mixture Model (GMM) ▪ Motion Vector Estimation ▪ K-Nearest Neighbors (KNN)

1. INTRODUCTION

Video surveillance is an emerging topic which detects, tracks, and identifies multiple moving objects without human intervention. Applications of video surveillance systems include security, traffic surveillance, crowd statistics, human iden-tification, and detection of anomalous behaviour. Moving object detection is very difficult when videos are captured in challenging weather conditions such as winter, snow storms, snow on the ground, fog, and air turbulence [1].

Object detection and tracking have a wide variety of applications in computer vision, video surveillance, vision-based control, video compression, and man-machine interfaces.

A robust, accurate, and high-performance approach is still a great challenge today. The difficulty level of this problem de-pends strongly on how the object to be detected and tracked is defined [2]. In the proposed method, multiple moving objects are detected effectively by using the Gaussian Mixture Model method. One of the bottlenecks of video surveillance systems is weather condition, especially fog. Fog or haze is the cause of many road accidents. Since fog affects the visual quality of an image, it leads to misinterpretation of target objects. To improve the efficiency of video surveillance systems, fog should be eradicated using the Guided Dark Channel Prior technique. Motion Vector Estimation is used for effective tracking of multiple moving objects.

2. RELATED WORK

Rawan Younis and Nabil Bastaki [3] proposed an improved fog removal technique by estimating the transmission map using the Proposed Adaptive Filter (PAF) to recover scene depth from a foggy image. He Yi, Sang Nong, Gao Changxin, and Han Jun [4] proposed a Gaussian mixture model for object detection and tracking. Features such as speed, empty ratio, aspect ratio, and dispersion ratio were estimated using the Histogram of Gradient method. The human and vehicle labels were then given as input to SVM classifiers and classified as humans or vehicles.

Sriram K. V. and Havaldar R. H. [5] proposed moving object detection using the GMM method with morphological operations, while HOG was used for feature extraction. Kirubaraj Ragaland and Tharics P. [6] surveyed object detection techniques such as frame differencing, optical flow, and background subtraction. Ding Ma and Zhezhou Yu [7] proposed a new real-time algorithm based on compressive sensing. This algorithm uses the behaviour of compressive sensing, with the major advantage of reducing time consumption during the tracking process. KNN classifier is then used to provide differences between the target and the background.

Soumya Varma and Sreeraj M. [8] proposed an efficient background subtraction algorithm known as Gaussian mixture model, which provides faster object detection. A set of simple and effective features is extracted by the support vector matrix. The performance of the algorithm is analysed using various kernels of SVM and also K-Nearest Neighbor with various parameters; using the SVM classifier, the system proves to be effective and efficient.

3. MATHEMATICAL EQUATIONS, SUBSECTIONS, TABLES, AND FIGURES

GMM is one of the popular techniques for estimating the background or foreground for moving object segmentation. It assigns a number of Gaussian distributions to each pixel to obtain the foreground [9]. If there are no variations in the pixel values, all Gaussian distributions approximate the same value. Gaussian Mixture Model (GMM) is a parametric probability density function represented as a weighted sum of Gaussian component densities. GMMs are commonly used as parametric models of probability distributions of continuous measurements or features in automation systems such as biometric systems. Using the iterative Expectation Maximization algorithm or Maximum a Posteriori (MAP) estimation from a well-trained prior model, GMM parameters are calculated as:

$$\omega_i = \frac{1}{T}\sum_{t=1}^{T} P_r(i/X_t,\lambda) \tag{1}$$

Here, X_t is the parametric probability of each pixel. Mean is calculated by:

$$\mu_i = \frac{\sum_{t=1}^{T} P_r(i/X_t,\lambda)X_t}{\sum_{t=1}^{T} P_r(i/X_t,\lambda)} \tag{2}$$

From the weight, mean variance is given by:

$$\sigma_i^2 = \frac{\sum_{t=1}^{T} P_r(i/X_t,\lambda)x_t^2}{\sum_{t=1}^{T} P_r(i/X_t,\lambda)} - \mu_i^2 \tag{3}$$

3.1 Preprocessing

The proposed edge-preserving smoothing technique uses guided filtering [10]; therefore, it can accurately remove artifacts present in the input image. Guided Dark Channel Prior is similar to the Dark Channel Prior (DCP). Atmospheric light is estimated based on the imaging law of very dense foggy regions. In that method, refinement of the transmission map is estimated using a guided filter so that halo effects are removed effectively. The main advantages of this method are accurate transmission-map refinement and more accurate atmospheric light estimation. Therefore, the fog-free image in Figure 2 does not look dim, and halo artifacts are removed effectively. Transmission-map refinement is performed using a guided filter, giving more accurate results for subsequent steps.

3.2 Multiple Object Detection

GMM is used to estimate the background or foreground for moving object segmentation. It assigns multiple Gaussian distributions to each pixel to obtain the foreground [9]. If there are no variations in the pixel values, all Gaussian distributions approximate the same value. GMM parameters are calculated using iterative Expectation Maximization or Maximum a Posteriori estimation from a well-trained prior model, as shown in Figure 3.

3.3 Feature Extraction

In machine learning, pattern recognition, and image processing, the first step in feature extraction begins with an initial set of derived data and constructs extracted features [1] that are informative, accurate in the region to be detected and tracked, and non-redundant. These features support better human understanding. Feature extraction is the process of dimensionality reduction. In the proposed system, features are extracted using Motion Vector Estimation, as shown in Figure 5. The major problem during feature extraction is limited frame variation in video sequences; therefore, important regions of frames remain constant in many frames. This redundancy can be overcome by the proposed method. Motion Vector Estimation uses two-dimensional vectors of the frame. These vectors are used for motion prediction by comparing two-dimensional vectors between two successive frames. The technique operates on macro blocks of the reference frame. The frame is divided into macro blocks, and the area of interest (target frame) is compared with the reference area (reference frame).

3.4 Tracking

Multiple moving vehicle tracking consists of calculating vehicle trajectories in traffic surveillance. For man-machine interface applications, tracking and automatic detection of moving vehicles are important tasks [11]. K-Nearest Neighbor (KNN) is a traditional machine-learning algorithm. The working principle of the algorithm is that if most adjacent samples of a sample in feature space belong to a category, then the sample belongs to that category. On the other hand, KNN depends on limited adjacent samples for classification. Applying KNN to object tracking is efficient and has been demonstrated in past works. The reasons this paper uses KNN as the tracking algorithm are that it does not require a training set for training and has lower time complexity compared with other approaches.

Based on the identified problems, the proposed work improves real-time tracking and is based on a compressive sensing algorithm [12]. This algorithm uses features from compressed sensing, which ensures real-time tracking, and the objects are differentiated from the background using compressive sensing. Thus, the results shown in Figure 6 demonstrate that the algorithm can improve the accuracy of tracking an object (vehicle).

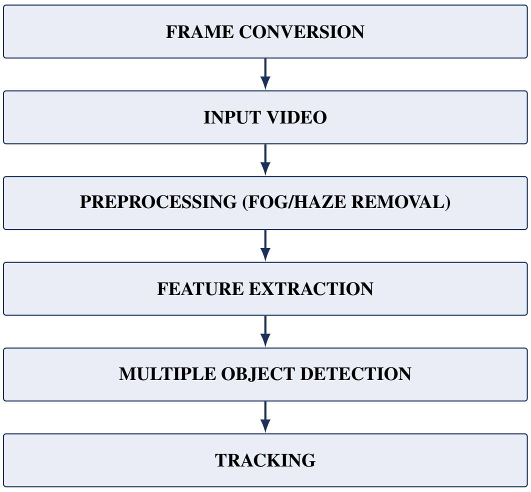


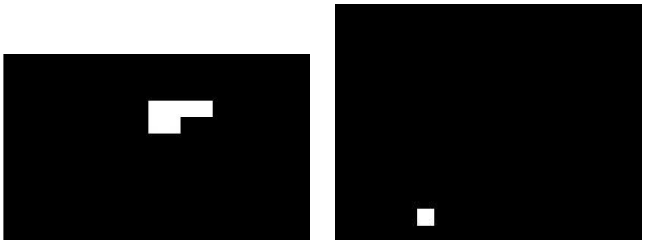
Figure 1. Proposed framework.

(a) Dataset-1 **(b)** Dataset-2

Figure 2. Dehazed output: (a) dataset-1, (b) dataset-2.

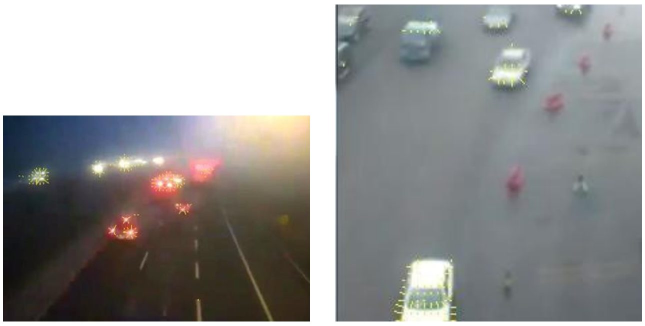
(a) Dataset-1 **(b)** Dataset-2

Figure 3. GMM method: (a) dataset-1, (b) dataset-2.



(a) Dataset-1 **(b)** Dataset-2

Figure 4. Saliency map: (a) dataset-1, (b) dataset-2.



(a) Dataset-1 **(b)** Dataset-2

Figure 5. Features: (a) dataset-1, (b) dataset-2.

(a) Dataset-1 **(b)** Dataset-2

Figure 6. Tracking: (a) dataset-1, (b) dataset-2.

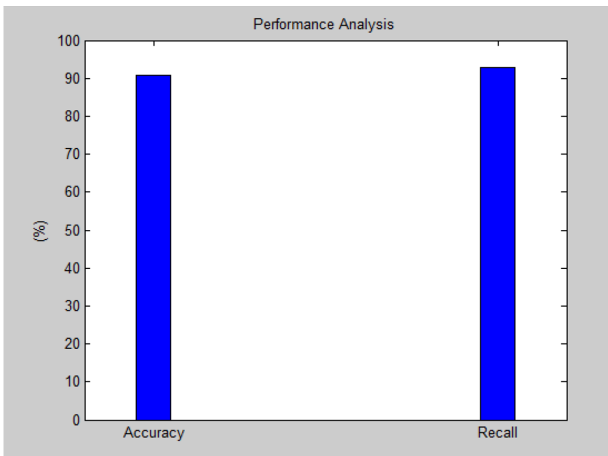


Figure 7. Performance analysis.

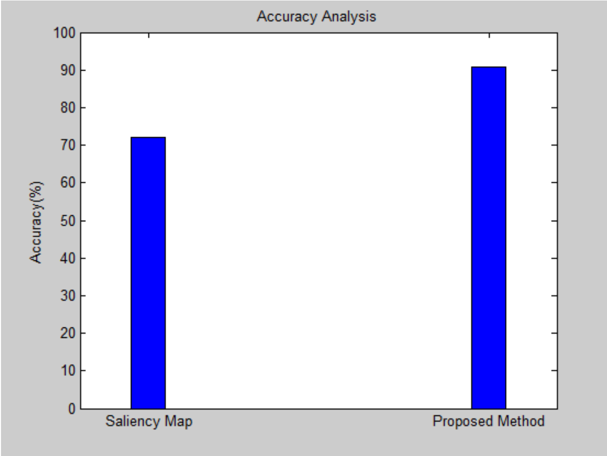


Figure 8. Comparative analysis.

4. CONCLUSION

The proposed algorithm of GMM for object detection and Motion Vector Estimation for feature extraction in surveillance systems detects and tracks multiple moving objects effectively using state-of-the-art detection algorithms and saliency map generation in the presence of fog or haze in video surveillance systems. It is observed that the proposed methods of fog removal by GDCP, GMM for multiple object detection, and KNN for tracking are executed, and the results show that the proposed method is better than existing algorithms. In future work, features will be extracted from detected moving objects using Optical Flow and Motion Vector Estimation detector techniques. The detected object will be classified as bike, car, or bus with the help of detected features. Moreover, deep learning algorithms such as CNN, DNN, and MFCN can be adapted for better accuracy.

REFERENCES

[1] N. Sabri, "Human detection in video surveillance using texture features," 2016, pp. 19–26.

[2] S. Kumar and J. S. Yadav, "Video object extraction and its tracking using background subtraction in complex environments," 2016, pp. 45–55.

[3] R. Younis and N. Bastaki, "Accelerated fog removal from real images for car detection," 2017, 9th IEEE-GCC Conference and Exhibition (GCCCE).

[4] H. Yi, S. Nong, G. Changxin, and H. Jun, "Online unsupervised learning classification of pedestrian and vehicle for video surveillance," 2017, vol. 26, pp. 145–151.

[5] K. V. Sriram and R. H. Havaldar, "Human motion detection and tracking for video surveillance," 2016, pp. 56–62.

[6] K. Ragaland and P. Tharics, "Survey on object tracking and detection methods," 2014, pp. 11–16.

[7] D. Ma and Z. Yu, "New video target tracking algorithm based on KNN," 2012, vol. 4, pp. 5497–5501.

[8] S. Varma and M. Sreeraj, "Video object extraction and its tracking using background subtraction in complex environments," 2016, pp. 45–55.

[9] M. Piccardi, "Background subtraction techniques: A review," 2016, vol. 4, pp. 3099–3104.

[10] J. Pang, O. C. Au, and Z. Guo, "Improved single image dehazing using guided filter," 2017, APSIPA ASC.

[11] N. Prabhakar, V. Vaithiyanathan, A. P. Sharma, A. Singh, and P. Singhal, "Object tracking using frame differencing and template matching," 2012, vol. 4, pp. 5497–5501.

[12] A. H. Mazinan and A. Amir-Latifi, "Applying mean shift, motion information and Kalman filtering approaches to object tracking," 2012, pp. 485–497.

[13] Cheng-Hsiung, M. N. Zhen, and Yu-Shung, "Single image haze removal with pixel based transmission map estimation," 2016, pp. 11–16.

[14] I. Kartika and S. S. Mohamed, "Frame differencing with post-processing techniques for moving object detection in outdoor environment," 2012, pp. 172–176.