Engineering and Technology Journal
Journal homepage: engtechjournal.org

Automatic Objects Detection and Tracking Using FPCP, Blob Analysis and Kalman Filter

Hadeel N. Abdullah [a*], **Nuha H. Abdulghafoor** [b]

[a] Electrical Engineering Dept. University of Technology, Baghdad, Iraq. 30002@uotechnology.edu.iq

[b] Electrical Engineering Dept., University of Technology, Baghdad, Iraq. 30216@uotechnology.edu.iq

*Corresponding author.

Submitted: 21/05/2019 Accepted: 31/08/2019 Published: 25/02/2020

KEYWORDS

Object Detection, Object Tracking, Fast Principle Component Purist (FPCP), Blob Analysis, Kalman Filter

ABSTRACT

Object detection and tracking are key mission in computer visibility applications, including civil or military surveillance systems. However, there are major challenges that have an effective role in the accuracy of detection and tracking such as the ability of the system to track the target and the response speed of the system in different environments as well as the presence of noise in the captured video sequence. In this proposed work, a new algorithm to detect moving objects from video data is designed by the Fast Principle Component Purist (FPCP). Then, we used an ideal filter that performs well to reduce noise through the morphological filter. The Blob analysis is used to add smoothness to the spatial identification of objects and their areas, and finally, the detected object is tracked by Kalman Filter. The applied examples demonstrated the efficiency and capability of the proposed system for noise removal, detection accuracy and tracking.

How to cite this article: H. N. Abdullah and N. H. Abdulghafoor "Automatic objects detection and tracking using FPCP, Blob analysis and kalman filter," Engineering and Technology Journal, Vol. 38, No. 02, pp. 246-254, 2020.

DOI: https://doi.org/10.30684/etj.v38i2A.314

1. Introduction

The process of tracking is one of the main tasks in computer visibility [1]. It has an important turn in many areas of research, such as movement guessing, recognition and analysis of human and nonhuman Vitality, 3D representation, mobility in vehicles, and others. Object tracking is the most common attribute in automated monitoring applications because the individual human employer cannot manage the controlled area, especially when the number of cameras rises. In addition, in the medicinal application, the operator cannot sometimes analyze the video taken by the device; especially in crucial cases, the detection and tracking system is more efficient than the human. It is also used in anti-theft systems, traffic management systems, and others. The tracking system can track single or multiple animation objects in different environments. In general, the object detection and tracking system include the different stages like background subtraction, object detection, and object tracking, as shown in the block diagram in Figure 1.

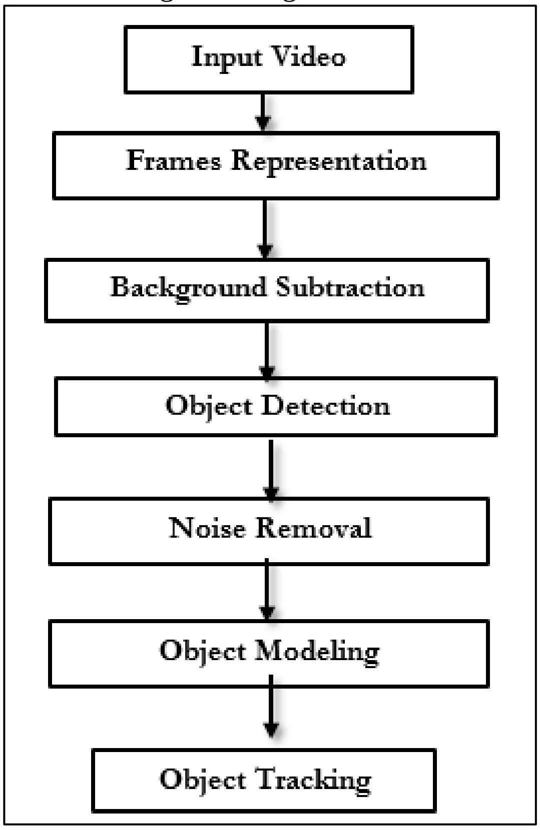

Figure 1: The Block Diagram of Object and Tracking System

The first basic step in many fields of image processing and computer vision is the background subtraction, also known as Foreground Detection, which extracts the foreground of the image for subsequent processing (such as object selection and identification). These are the most important areas of the picture, which are called objects such as humans, cars, texts, etc. This stage maybe after the pre-processing phase of the image, which may include noise reduction of images, and before the subsequent treatment phase such as morphology, etc. A more common way to detect moving objects in videos is background subtraction. The basic principle is to detect moving objects from the difference between the current frame and the reference frame, often called a "background-image" or "background model". Some common methods in this area include the use of frame differentiation, optical flow, analysis of principal components, and background mixture models [2].

Object modelling represents the object of interest in a scene. To represent an object, features are extracted that uniquely defines an object. These features or the descriptors of an object are then used to track the object. A feature is an image pattern that differentiates an object from its neighbourhood. The features of an object are converted into descriptors, also referred to as appearance features, using some operations around the features [1]. The commonly used object representations for tracking are centroid, multiple points, rectangular patch, and complete object contour, etc. while the descriptors of an object such as probability densities of object appearance (Histogram), template, blob analysis, etc.

The object tracking is to select and give individual paths to each object in the video sequence. Objects can be humans on the street, cars on the road, players on the pitch, or from a group of animals. The object is tracked to extract the object, identify the object and track it, and the decisions related to their activities. Trace objects can be classified as points tracking, kernel tracking, and trace shadow images. The general techniques of tracking such as Kalman Filter, Particle Filter, Mean Shift Method, etc.

The main contribution of this work is as follows:

1- The proposed algorithm uses the FPCP technique to extract the motion areas from different backgrounds of the captured video frame without the need for further input. As a result, so the outputs have good speed and accuracy.

2- Using the method of analyzing the blob spatially simultaneously to select the effective pixels in the motion zones and to determine the area of the object at the same time. In addition to using efficient tracking technique, Kalman's Filter is thus an efficient and integrated way to track multiple objects in the same captured video frame.

This paper is ordered as follows: Section 2 explains the related works, Section 3 explains methodologies (Mathematical Background), Section 4 explains the proposed algorithm, Section 5 explains the results and discussion. Finally the conclusions in Section 6.

2. Related Works

Since the last few decades, many researchers have proven algorithms for detecting and tracking objects. In this section, we demonstrated some of these algorithms related to the proposed system. According to [3], motion brim is extracted in polar-log coordinate; then the gradient operator is employed to compute the optical flow directly in every motion regions. Finally, the object is tracked. In the proposed work in [4], the active background is reconstructed and the object size is determined as a preliminary task, to extract and track the object in the foreground. The method in [5] is the object detection is done by Gaussian Mixture Model (GMM), and Kalman Filter does the tracking. In this method, Object detection is determined based on the size of the foreground. Therefore Errors will occur in determining the object such as the object and its shadow are merged as an object or representing two adjacent compounds as a single object. The paper [6] developed; the algorithm includes optical flow and the motion vector estimation for object detection and tracking. The detection and tracking system in [7] is sophisticated depend on optical flow for detection; the object tracking is done by blob analysis.

In Prabhakar et al. [8], a moving object tracking system using morphological processing and blob analysis, which able to distinguish between car and pedestrian in the same video. In the paper [9], the foreground is extracted from the background using multiple-view architecture. After that, the forward movement date and editing schemes are used to detect the animated objects. Finally, by detecting the center of gravity of the moving object, it is used to trace the object based on the Kalman Filter.

In the method [10], animated objects are represented as groups of spatial and temporal points using the Gabor 3D filter, which works on the spatial and temporal analysis of the sequential video and is then joint by using the Minimum Spanning Tree. The proposed technique described in [11], split into three stages; Foreground segmentation stage by using Mixture of Adaptive Gaussian model, tracking stage by using the blob detection and evaluation stage which includes the classification according to the feature extraction.

After exploring some of the published research on the detection and tracking of the object, it was found that the discovery and tracking of the object is a complex task because of many elements of dynamic tracking such as determining the type of camera moving or static, the random. Change of the speed of the object, the intensity of light and darkness, etc.

3. Methodologies (Mathematical Background)

I. Fast principal component pursuit

FPCP was recently suggested [12,13] as a powerful alternative to Principal Components Analysis (PCA). This method will be used in various applications, including foreground/background modelling, data analysis, whether in text or video format and image processing. The PCA was formulated initially [12]:

$$\arg min_{L,S} \ \left||L|\right|_{o} + \lambda \, ||S||_1 \ \text{ s.t. } \ D = L + S \tag{1}$$

Where $D \in R^{m\times n}$ is the observed matrix, $||L||_o$ is the nuclear norm of matrix L (i.e. $\sum_k |\sigma_k (L)|$) and $||S||_1$ is the l^1 norm of matrix S. Numerous changes have been made to eq. (1) by changing the restrictions on sanctions and vice versa. So that the eq. (1) became:

$$\arg min_{L,S} \ \frac{1}{2}\left||L + S - D|\right|_F + \lambda \, ||S||_1 \ \text{ s.t., } \| L \|_o < t, \tag{2}$$

The constraint $\|L\|_o < t$ is active, represents a constraint of equality, so it is suggested that the algorithm ranks the same rather than relax the nuclear base, so the function is as follows:

$$\arg min_{L,S} \ \frac{1}{2}\left||L + S - D|\right|_F + \lambda \, ||S||_1 \quad \text{ s.t. rank } (L) \approx t \tag{3}$$

This adjustment ignored the initial selection of the parameter λ, the background-modelling compound L is often low, and in practice, there is no difficulty in selecting the appropriate value for t.

The normal process to solve eq. (3) by the substitutional minimization as follow:

$$L_{k+1} = \arg_L \min \ ||L + S_k - D||_F \quad \text{s.t.} \quad \text{rank}(L) \approx t \tag{4}$$

$$S_{k+1} = \arg_S \min \ ||L_{k+1} + S - D||_F + \lambda ||S||_1 \tag{5}$$

The eq. (3) can be solved by taking a partial Singular Value Decomposition SVD of (D - S_k) with respect to t. while the eq. 4 can be solved by element-wise shrinkage. The background of the videos is supposed to lie in a low-level sub-space, and the moving objects should be in the foreground as if they were gradually soft in the spatial and temporal direction. The proposed method integrates the Frobenius and l^1-norm base into a unified framework for simultaneous noise reduction and detection. The Frobenius base uses the low-level property in the background; the contrast is improved by the l^1 norm standard [13].

II. Noise filtering

Animated digital pictures often overlap with a set of noise based on prevailing conditions. Some of this noise is very disturbing when implicated in altering the intensity of video frames. It spoils pixels randomly and divides into two extreme levels: relatively low or relatively high, compared to adjacent pixels [8]. Subsequently, it is necessary to apply refinement technicalities that are able to handle various types of noise.

Morphological processes are performed to extract important features of useful images in the impersonation of shapes in the region and their description. We have used both the morphology of the closure and corrosion, respectively, to remove parts of the road and unwanted things. After the morphological closure process, provided that the appearance of the object was not destroyed, and that many small punctures and separate pixels has been filled in the form of one sizeable real object [8]. The following is the definition of morphological closure process and the applicable structural element B.

$$P * B = (P \oplus B) \oplus B \tag{6}$$

Where:

$$B = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \tag{7}$$

The matrix P, which includes moving object information, is obtained through the detection process. An integral part of the morphological expansion and erosion processes is a structural element of a flat shape. There is a binary flat structure element with a living value, either 2-D or multi-dimensional, in which the real pixels are included in the morphological calculation, and false pixels are not. The middle pixel of the structure element, called the parent, determines the pixel in the image being processed.

III. Blob analysis

Blob analysis is used to determine two-dimensional objects in an image. The detection depends on the spatial properties using assured standard. In many applications where the calculation is time-consuming, one can use point analysis to eliminate points that do not matter based on specific spatial properties and retain relevant points for further analysis [8]. The foreground object is adjusted to the blob region. The object corresponding to the point area is detected as a composite object and features as a bounded box. The detected object will be ignored as a foreground but does not correspond to the point area and is not marked with a bounding box.

IV. Kalman Filter

Object tracking is a way to find and create a path to the object that was discovered. In this search, the Kalman Filter method was used to track an object in sequence with captured video [14]. The Kalman filter is a linear approach that operates in two basic phases of prediction and correction (update). The prediction phase is accountable for the scoop of the next state and position of the present object. However, the correction phase provides the parameters with their instance; they combine the actual measurement with the previous estimate to improve the trajectory where the object information detected in the previous frame is used and provides an estimate of the object's new position. The

Kalman Filter has the ability to rating the tracking locations with minimal datum on the location of the object. Initially, the status S_t and measurement X_t paradigm are determined to predict the next site. The paradigm matrixes are defined as [5]:

Prediction

$$\ddot{X}_t = A\ddot{X}_{t-1} + Bu \tag{8}$$

$$S_t = AS_{t-1} A^T + Q \tag{9}$$

Where; A - state transition matrix, B - coverts control input and Q - process noise covariance.

Correction: The measurement update equations are given as:

$$K_t = S_{t-1}H^T (HS_{t-1}H^T + R)^{-1} \tag{10}$$

$$\ddot{X}_{t+1} = \ddot{X}_t + K_t(Y_t - H\ddot{X}_t) \tag{11}$$

$$S_{t+1} = (I - K_tH)S_t \tag{12}$$

Where K- Kalman gain, S- measurement matrix, R- measurement error covariance and H-model matrices. The prediction of the next state S_{t+1} is done by integrating the actual measurement with the pre-estimate of the situation S_{t-1}.

4. The Proposed System

In this pager, object detection and tracking algorithm, a collection of two famous computer visibility technologies, Fast Principle Component Purist (FPCP) and the Kalman filter, was introduced. FPCP is used in the object discovery phase. It provides quick and delicate object detection on other methods such as background subtraction.

FPCP does not provide the path of motion; instead, it supplies acquaintance about the orientation of the object and its motion in vector form. This system has many features, including the possibility of tracking more than one object and the speed of response to a change in speed and change in the scale. Previously any process, at first, the video is taken by the stationary camera. The video is only a series of cascading frames, so the object detection manner must first detect the moving object in these cascading frames. Then the algorithm converts the video into two-dimension matrices to facilitate handling in mathematical calculations, to reduce the time calculation and memory requirements.The proposed algorithm shows the detail stages of the system are followed to that deals with background separation, object and feature extraction. The processing includes removing the noise. Then the process status of all pixel is tested by blob analysis and clustered it to detect the object. Initialize the tracking stage and update the tracker in every new frame.

The proposed algorithm

Input: Movie file = a video of size m x n x nframe.

Output: Outputframe = frame salient with bounding boxes for each object;

Video to Matrix Conversion [X}

Object detection using FPCP algorithm.

Input [X] video matrix size m x n x nframe

Output [L]: Low-Rank Matrix, [S]: Sparse Matrix

Save foreground matrix [S]/

While **Read frame is done**,

Do Extract Frame

Let, Tc = trackers for a count of moving objects in present frame;

1. Create New Video Player as output.
2. Read the frame and its foreground [S_k.].
3. Remove any noise and holes in the foreground frame by morphological filter.
4. Discover the moving object by grouping pixels connected spatially and temporally using the Blob analysis method.
5. Build a function matrix to calculate the position, area and the dimension of border-box as vector of each the Object1; . . . ; Objectn.
6. Assign each detected Objects by their above vector in order to initialize the track.
7. Assign the initial position and status to Kalman Filter for each trackers.
8. The motion of each track was been estimated by Kalman Filter.
9. Initialize Kalman Filter prediction to update the position in next frame, and set the probability of each detection being set for each paths.
10. The assigned tracks are updated using the corresponding detections.

11. Update the current status in the present frame to assign the new detection tracks or remove any invisible or lost tracks.
12. Data association for the same detection object for the present frame.
13. Display the resulting frames.
14. End while.

5. Results and Discussion

The algorithm proposed in MATLAB (2018b) has been applied, and their experiments were performed on a Computer type MSI GV63 with Intel Core i7 8750H, NVIDIA RTX 2060 6G, 256GB SSD+1TB and 16 GB RAM. It has three stages are foreground detection, filtering and tracking. The proposed algorithm detects the movable objects accurately and keeps track of their appearance in the sequence video frames. Video data has been used in any format as an input to the proposed work, and good results have been obtained in various article conditions on this indoor, outdoor, light traffic and dense traffic. The efficiency of the proposed algorithm was evaluated; the experiential outcomes were as follows:

Figure 2(a) shows the original framework. In Figure 2(b) the foreground was extracted by the FPCP detection and showed holes and noise on the frame. In order to clarify and soften the frame, the morphology was done.

The foreground frame is shown after the morphological process in Figure 2(c) and the final output frame as in Figure 2(d), which includes movable and tracked objects.

Several-sampled video is used in the various environment in order to test the performance of the proposed algorithm. The experiential outcomes are given in Figure 3. The first column (Figure 3(a)) shows the sampled frames of the video then the second column (Figure 3(b)) shows a clean foreground extracted frames by FPCP detection are given. The third column (Figure 3(c)) includes the detected and tracked objects by marking it with a circumferential box.

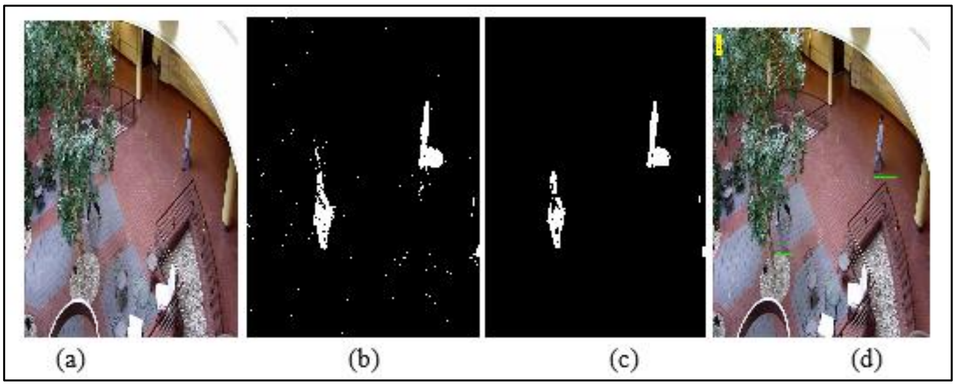

Figure 2: A sampled result (a) original frame (b) Foreground before and (c) after applying the morphological filter and (d) The final result

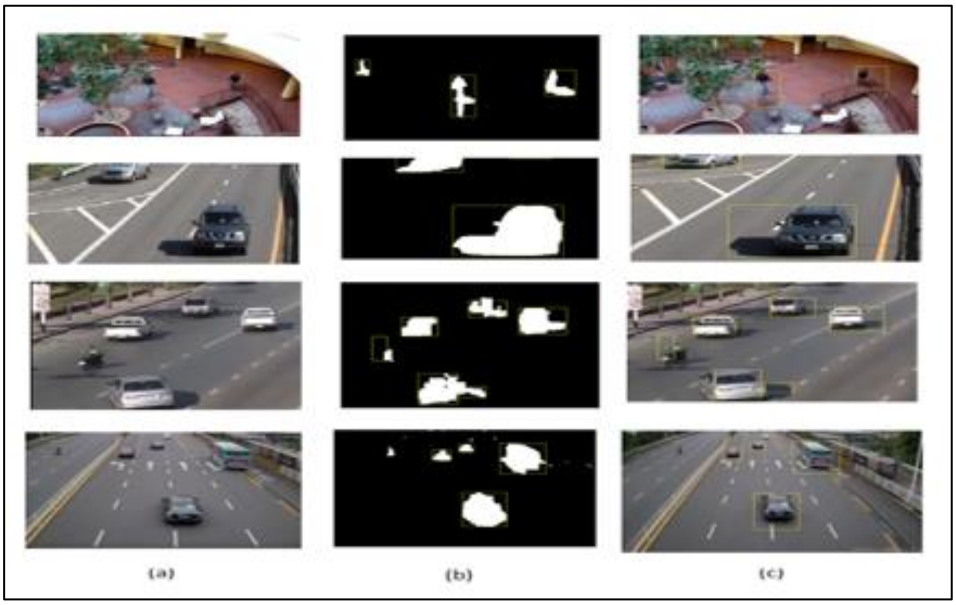

Figure 3: The proposed algorithm results on several-captured video, (a) Input Video, (b) the filtered foreground and (c) the resultant video

Table 1 explains the mean execution times (sec) of the proposed system when implemented on 365 sampled video frames by Matlab version R2018b. The consuming time of evaluation reduces drastically. Accuracy is a measure of the performance efficiency of the object tracking system. The detection and tracking system precision can be calculated using the following formula:

$$\text{Accuracy} = \frac{\text{The total number of detected objects by system}}{\text{The total number of actual objects in video}} \quad (13)$$

The proposed algorithm for the different video input has been tested with different methods to evaluate its accuracy. The accuracy of the proposed tracker in different input scenes was compared and compared with other tracking systems, as shown in Table 2. It shows that the detection and tracking accuracy rate using the proposed algorithms is 100%. The results are optically acceptable except for an algorithm t.

The proposed algorithm for the different video input has been tested with different methods to evaluate its accuracy. The accuracy of the proposed tracker in different input scenes was compared and compared with other tracking systems, as shown in Table 2. It shows that the detection and tracking accuracy rate using the proposed algorithms is 100%. The results are optically acceptable except for an algorithm that proves that this multi-object tracking method is validated. It was concluded that the proposed algorithm was still competitive, although some results were closest to another, with little degradation in some because of the cost of some complicated calculations.

The detection precision of the suggested algorithm is compared with other known and present methods. The comparison shows the efficient performance of the proposed method on some of the selected frames shown in Figure 4. We compared the proposed algorithm with the most representative algorithms and for different frame sizes and settings for the tested videos; we used grey or chromatic video sequences. The results were comparable to the proposed algorithm with other algorithms.

To evaluate the visual performance of the proposed algorithm, we compared the proposed algorithm to 3 algorithms. The videos examined contain different background scenes, and multiple moving objects both outdoors and indoors (pedestrians, vehicles, etc.). We have chosen the following most methods to compare with our proposed method: (1) GMM method [5], (2) optical flow [7], (3) MODT [9].

Visual results are shown on the videos tested in Figure 4. Individual and group infantry, small dynamic background, and multiple traffic surveys as shown in Figure 4, the proposed algorithm is closest to Ground Truth (GT). Some of the results of the tested algorithms consider the foreground object as the background. The main reason is that the parts of the object remain static in the video and that the proposed algorithm has overcome this effect, and obviously the detection effect is better than other algorithms.

Table 1. Performance Execution Time

The Step	Mean Execution Time (Second)
Loading the video and conversion to matrix	5
FPCP foreground detection	10.87
Set Object System and frame reading	5.2
Noise Removal	1.15
The blob analysis System object	2.7
Tracking System	17.6
Total	42.52

Table 2. Percentage Accuracy%

Comparison Parameter	**GMM Method [5]**	**Optical Flow Method [7}**	**Motion Vector Method [6]**	**Proposed Tracking System**
Single Human	100	90	100	**100**
Speed Diversity	80	10	90	**85**
View Point Difference	90	90	90	**90**
Fixed Objects	80	70	90	**90**
Multiple Objects	85	20	90	**95**

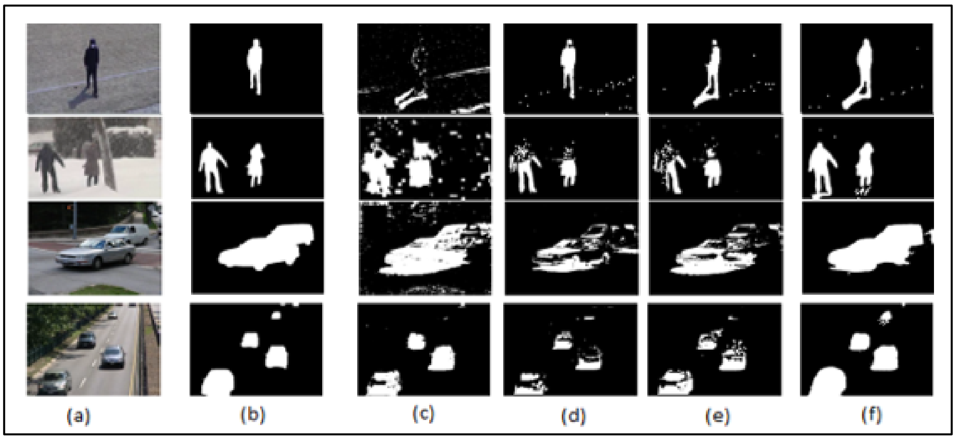

Figure 4: The comparison results for several experiment captured video: (a) Input Frame, (b) The Ground Truth, (c) GMM Method [5], (d) Optical Flow [7], (e) MODT [9], and (f) The Proposed Algorithm

6. Conclusions

Object detection and tracking are the main and affront mission in many computer visibility implementations, such as monitoring, car salt works, routing, and automation. The proposed algorithm consists of three stages; the first stage foreground detection and filtering of various types of noise from images using FPCP technique, the second stage of the identification of animation objects and their region by blob analysis method and finally, the Kalman Filter is used for tracking the objects. This algorithm presents several benefits, such as multiple object detection and tracking in different environments. The disadvantages of this technique using one method will not produce perfect results because its accuracy is influenced by different operators such as the low resolution of captured video, change in weather, etc. In the future, we hope to expand our scope of detection and tracking of objects in overcrowded scenery or the appearance of severe contrast in lighting and real-time scenes.

References

[1] J. Cheng, J. Yang, Y. Zhou and Y. Cui, “Flexible background mixture models for foreground segmentation,” Image and Vision Computing, Vol. 24, No. 5, pp.473-482, 2006.

[2] Y. Wu, J. Lim and M.H. Yang, “Online object tracking: A benchmark,” Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2411-2418, 2013.

[3] H.Y. Zhang, “Multiple moving objects detection and tracking based on optical flow in a polar-log image,” International Conference on Machine Learning and Cybernetics, IEEE, Vol. 3, pp. 1577-1582, 2010.

[4] N.A. Mandellos, I. Keramitsoglou, and C.T. Kiranoudis, “A background subtraction algorithm for detecting and tracking vehicles,” Expert Systems with Applications, Vol. 38, No. 3, pp.1619-1631, 2011.

[5] R.Y. Bakti, I.S. Areni and A.A. Prayogi, “Vehicle detection and tracking using gaussian mixture model and kalman filter,” International Conference on Computational Intelligence and Cybernetics, IEEE, pp. 115-119, 2016.

[6] K., Kale, S. Pawar and P. Dhulekar, “Moving object tracking using optical flow and motion vector estimation.” 4th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO) (Trends and Future Directions), IEEE, pp. 1-6, 2015.

[7] S. Aslani and H. Mahdavi-Nasab, “Optical flow-based moving object detection and tracking for traffic surveillance,” International Journal of Electrical, Electronics, Communication, Energy Science and Engineering, Vol. 7, No. 9, pp.789-793, 2013.

[8] P. Telagarapu, M.N. Rao and G. Suresh, “A novel traffic-tracking system using morphological and Blob analysis”. International Conference on Computing, Communication and Applications, IEEE, pp. 1-4, 2012.

[9] W.C. Hu, C.H. Chen, T.Y. Chen, D.Y. Huang and Z.C. Wu, “Moving object detection and tracking from video captured by moving camera,” Journal of Visual Communication and Image Representation, 30, pp.164-180, 2015.

[10] K.S. Ray and S. Chakraborty, “Object detection by spatiotemporal analysis and tracking of the detected objects in a video with variable background,” Journal of Visual Communication and Image Representation, 58, pp.662-674, 2019.

[11] T. Mahalingam and M. Subramoniam, “A robust single and multiple moving object detection, tracking and classification”. Applied Computing and Informatics, 2018.

[12] T. Bouwmans and E.H. Zahzah, “Robust PCA via principal component pursuit: A review for comparative evaluation in video surveillance,” Computer Vision and Image Understanding, 122, pp.22-34, 2014

[13] P. Rodriguez and B. Wohlberg, “Fast principal component pursuit via alternating minimization,” IEEE International Conference on Image Processing, pp. 69-73, 2013.

[14] Q. Li, R. Li, K. Ji and W. Dai, “Kalman filter and its application,” 2015 8th International Conference on Intelligent Networks and Intelligent Systems (ICINIS), IEEE, pp. 74-77, 2015.