A Multi-object Tracking Method Based on Fuzzy-logic and Top-Down Detector

Dr. Matheswari Rajamanickam
Assistant Professor, Department of Computer Science,
M.E.S College of Arts, Commerce and Science, Bangalore, Karnataka, India.

Abstract - Object detection and tracking aims at detecting and tracking several objects over a sequence of images. The research works focus on an integrated approach for detection and tracking of moving objects. An object classifier in spatial domain with Kalman filter based people tracing is developed. The research work segments and classifies objects into multiple classes using Spatial Classification of Multi-Class Objects method. Initially, the Markov Random Field (MRF) principle is used to estimate the shape of objects. The maximum a posteriori probability (MAP) estimation principle is performed for spatial classification of objects. The research developed an improved Kalman Filter method of pattern matching and tracking moving objects in a closed environment is presented in the paper. The top-down approach based on Kalman Filter (KF) is performed to detect the chromatic shadows of objects. Next, Kernel pattern segment function detects the moving object’s pattern. Finally, object tracking is performed using the proposed method to track the moving objects in a particular region using the minimum bounding-box method. Experimental results reveal that the proposed method achieves better performance in terms of classification time, classification accuracy, pattern matching time, pattern matching accuracy, true detection rate and object tracking accuracy with respect to the number of video frames per second.

Keywords: Spatial classification, Markov Random field, Kalman Filter, Top-down approach, Object Detection, object tracking.

I. INTRODUCTION

Moving object detection and tracking are considered as a significant and demanding task in the applications of human-computer interfaces, robotics, video surveillance system etc. The higher-level computer vision tasks like 3D reconstruction and object representation receives input from these applications. The detection and tracking of moving objects from video sequence helps in better understanding the behavior of objects. Therefore, analyzing the behavior of objects in the video frames and obtaining visual records of objects play an important role in the video surveillance systems. The identification of dynamic shadows, occlusion detection and tracking of multiple moving objects in crowded places are still a challenging task. Another significant issue is the detection of chromatic shadows of objects appearing at any given time and differentiating shadow objects from the real objects is still a promising task. Therefore, pattern matching is an essential step for detecting and tracking of single or multiple moving objects.

An intelligent system using Batch incremental SVMs (BISVM) classifier was designed in [1] for classifying four classes of moving objects using a new feature descriptor that considers both object form and moving information to improve the classification accuracy. The feature extraction and classification approach needs to be integrated with other tracking algorithms for applications in dense crowded regions. A multi-sensor fusion framework designed in [2] for Moving Object Detection and Tracking and also reduced the number of false detections. However, the visual descriptors of the image are not generated for detecting the moving objects.

In [3], Scale Invariant Feature Transform technique (SIFT) method was introduced for point extraction and matching. Though the method effectively locates regions of the image, the method failed to consider more color information for a robust pattern matching strategy. In [4], an Enhanced Rao-Blackwellized Particle Filter (E-RBPF) is introduced for multiple target detection and tracking. However, this method failed to track large number of targets in crowded areas.

Active contour models (ACMs) was designed in [5] for detecting and tracking an object system. However, the presence of dynamic shadows was not addressed. An advanced fuzzy aggregation-based background subtraction (AFABS) was designed in [6] for moving object detection in dynamic background conditions. However, the quality of the video was compromised.

A normalized self-adaptive optical flow for discovering moving object region was designed in [7]. Exact identification of moving object areas with different object size was achieved despite complex background condition. However, the method is suitable for video sequences captured using single static camera. An efficient background subtraction algorithm for object-tracking task under static and dynamic background conditions was proposed in [8].
An adaptive Kalman Filter method was used to extract foreground moving objects. The method needs to be implemented in better localization of moving or stationary object than other background subtraction schemes for tracking.

The detection of moving objects in uneven imaging environment was designed in [9] using various categories of cameras and unmanned aerial vehicle (UAV) cameras. An efficient and accurate method for segmentation of moving objects under challenging conditions like uncovered background, temporary poses and global motion of background (GMOB) was designed in [10]. However, pixel error still needs to be reduced.

II. RELATED WORK

Object Tracking with Multi-View Support Vector Machines was presented in [11]. The method is robust and accurate despite occlusions and changing background conditions. A temporal-spatial variable scale algorithm for identifying multiple moving targets from complicated backgrounds was designed in [12]. An octree decomposition algorithm based on temporal-spatial domain was developed. However, the objects were identified with minimum incidence of false alarms and miss detection rate.

A computationally low cost and robust detection and tracking of moving objects system was presented in [13]. The system is trained to detect moving objects and keeps target of even when no detections are observed. An efficient tracking of moving objects was achieved by employing DATMO method. However, the system was imposing a trade-off between performance and initialization time.

An object tracking method using an adaptive template matching with Sum of Squared Difference (SSD) was designed in [14]. A robust object tracking algorithm called Simplified Codebook Masked Camshift algorithm (SCMC algorithm), was presented in the research work in [15]. However, object tracking in case of changing illumination is still a challenging issue. A robust video segment proposals with painless occlusion handling was developed in [16]. The method is robust to occlusions. Segment size interpolation is used to identify occlusions. However, the research work need to design better features or a more sophisticated system to track a specific object in one video and generalize it across videos.

A spatio-temporal-based approach for segmenting objects was presented in [17]. A tracking method using particle filter using a Rao-Blackwellized particle filter for identifying objects in a high dimensional state-space was presented in [18]. A fuzzy Adaptive Resonance Theory (ART) neural network for detecting the moving objects was designed in [19]. The background model combined with the proposed method is capable of learning from the new video sequences efficiently and more precisely. A forgetting procedure was applied to identify the neurons for discarding and reconstruction. Finally, the finding procedure is performed to estimate the extreme values more quickly [19]. A fast pattern matching approach for tracking objects guided by Haar Projection values (HPV) was designed in [20]. The Haar transform was performed for converting these image values to Haar Projection Values (HPV) for efficient pattern matching on sliding window of image scene.

III. RESEARCH METHODOLOGY

The research work proposes an integrated approach of detection and tracking of moving objects.

A Spatial Classification of Multi-Class Objects (SCMO) method and an Extended-Kalman Filter-based Tracking of Moving Objects (EKF-TMO) method is proposed in the study.

- First, the proposed method extracts input frames from the input video sequence.
- Objects are segmented and classified into multiple classes.
- Markov Random Field (MRF) and fuzzy logic is applied to obtain the shape of the moving or fixed object and preserve the objects’ boundary.
- Maximum a Posteriori (MAP) estimation and Bayesian algorithm is performed in the spatial domain.
- Chromatic shadows are identified and moving objects are detected and tracked by the Extended-Kalman Filter-based Tracking of Moving Objects method.

Object tracking involves the process of segmenting and classifying objects present in the extracted video frames, tracking its movement and position. Spatial classification is based on the spatial features like areas region, roads, and ponds or rivers. The SCMO method performs the classification of object based on their segmented texture features which results in improved object classification accuracy with minimum time. The SCMO method takes the video file as input as shown in figure 1. The method then extracts image frames.
3.1 Shape Estimation of Objects

The MRF principle is applied for fixed and moving object in order to provide a label field fusion. The MRF fusion approach groups the label fields collectively. A label field consists of dissimilar information. In SCMO method, the MRF principle considers two label fields such as a region plot (R) which is obtained after segmenting the input frames and a spatial region plot (‘S’) of the application label field. The spatial region plot ‘S’ is application specific and it contains the information about the occlusion labels (i.e. boundary labels), motion labels, other high-level information. These two label fields are combined together to reduce local minima.

The input video frame is segmented in to region plot (R) and the spatial region plot (S). These two labels are combined to provide the resultant label field (X). Let us consider, R and S is realization of a pair of joint Markov random field (MRF) and it is estimated by the property of the Hammersley–Clifford theorem. The joint probability density of the MRF is represented as follows:-

\[ P(R_t, X_t, \varphi_t) = \frac{1}{Z} \exp(-E_t(R, X|\varphi_t)) \]  

From (1), where Z represents the normalization factor. \( E_t(R, X|\varphi_t) \) denotes the local energy function measuring how well R and S fit together around the location \( l \). \( \varphi_t \) is the local joint neighborhood function around the location in \( l \). Therefore, the energy function is measured as follows:-

\[ E_t(R_t, X_t|\varphi_t) = -\sum_{i \in \varphi_t} \delta(R_t, R_i) \delta(X_t, X_i) \]  

From (2), \( \delta(X_t, X_i) \) and \( \delta(R_t, R_i) \) is the Kronecker delta function. The resultant label field is measured as follows:-

\[ X = \arg \min \sum_{i \in L} E_t(R_t, X_t|\varphi_t) \]  

(3)

For each location, the resultant value \( X \) maximizes the local conditional probability density function and this considerably reduces the classification time. Therefore, The SCMO method uses MRF principle to show the shape of the moving or fixed object. The output of the shape extraction using fusion approach is shown in figure 2.

From figure 2, the shape of the objects are identified at particular time. Next the edge-strength of the objects needs to be preserved for estimating object’s boundary. The fuzzy “If- Then” rules are used to calculate the edge-strength. The two gradient operators \( G_x \) and \( G_y \) at a point \( (x, y) \) in the image are considered. The gradient of the objects are mathematically defined as,

\[ \nabla f = \begin{bmatrix} G_x \\ G_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \]  

(4)

From (4), \( \frac{\partial f}{\partial x} \) denotes the derivative with respect to gradient in x direction, \( \frac{\partial f}{\partial y} \) represents the derivative with respect to gradient in y direction.

The four if-then rules in fuzzy system are used to determine the edge strength. In Rule 1, if \( G_x AND G_y \) is low then the normalized edge-strength \( N(x, y) \) is low. Rule 2, if \( G_x \) is low \( AND G_y \) is high then the normalized edge-strength \( N(x, y) \) is medium. In rule 3, if \( G_y \) is high \( AND G_y \) is low then the normalized edge strength \( N(x, y) \) is medium. In rule 4, if \( G_x AND G_y \) is high then the normalized edge strength \( N(x, y) \) is high. Since, the two
input variables $G_x$ and $G_y$ are measures of the intensity of gradient, thereby differing only in their directions of measure along x and y axis. The input image with fuzzy edge strength measure is shown in figure 3.

![Figure 3. Fuzzy Normalized Edge-Strength for Frame 1](image)

The fuzzy IF-THEN rules in fuzzy system are applied to measure the edge-strength as low, medium and high. As a result, the higher edge-strength is obtained for preserving the boundary of the object.

### 3.2 Classification of Objects in Spatial Domain

This MAP estimation function classifies the objects from the video frames effectively based on different class labels. Likelihood of each pixel’s class label is derived from the classification model to produce the final classified results. This approach is used to perform the spatial classification based on the analysis of moving objects related to its spatial characteristics, such as areas region, roads, and ponds or rivers. Bayes' theorem is used for classifying the segmented objects in the estimation process for reducing the false positive rate.

$$P(X|D) = \frac{P(D|X)}{P(D)} P(X)$$

From (5), $P(X)$ is said to be a probability of prior information about the true classification and it’s ratio of the number of location and the different class labels. Here, $P(X|D)$ represents the posterior probability. In case of the location prediction a decision-based on Bayes’ rule is considered for a particular location. The Bayes inference rule produces the MAP estimation and is represented as follows:

$$\arg \max P(X|D) = \arg \max P(X|D) P(D)$$

The MAP finds maximum value of the probability density function. Based on the above formulation, the Maximum posterior probabilities for different class labels are obtained.

Let us consider, a set of segmented objects $O_i = o_1, o_2, o_3, \ldots, o_n$, and the corresponding class labels $c_i \in C$. By using MAP estimation ($c_{MAP}$), the objects are classified by using Bayes classifier, and is expressed as follows:

$$c_{MAP} = \arg \max_{c_i \in C} P(c_i|O_i = o_1, o_2, o_3, \ldots, o_n)$$

From (7) the classifier predicts the probability of the class which has most likelihood value. Therefore, MAP returns class value where the probability is highest in the given frames. Spatial classification (SC) of input image is shown in figure 4.

![Figure 4. SC - Level 1, Level 2 and Level 3](image)

As shown in figure 4, the classification model produces the final classified results of input moving object.

### 3.3 Shadow Detection Using Top-Down Approach

Once after the classification of objects in to classes, the next step is the identification of chromatic shadows of objects. The detected objects are tracked using first order Kalman filter. The images obtained are shown in figure 5.

![Figure 5. (a) Input frames (b) Shadows of object at time “t” (c) Foreground objects at time “t”](image)
The general first order Kalman filter can be expressed as follows,
\[ x_k = S_k x_{k-1} + b_k c_k + n_k \]  
(8)
From (8), “\(x_k\)” denotes a “\(k^{th}\)” frame, “\(S_k\)” represents the state transition model which is applied to the prior state “\(x_{k-1}\)” (i.e. \((k-1)^{th}\) frame). Here, “\(b_k\)” denotes a control input model which is applied to control vector “\(c_k\)”. Here, “\(n_k\)” is the process noise which is assumed to be zero with covariance.

The initial state and noise at each step \(\{x_0, n_1, ..., n_k, u_1, ..., u_k\}\) are considered to be mutually independent. Each track is correlated with these parameters and Kalman filter (KF) is used to predict the object’s location using first order motion model. The Kalman filter includes two step process: prediction and update. Using the prior state, the prediction phase performs state estimation for obtaining current state as follows:-

Prediction state estimation= \(x_k = S_k x_{k-1} + b_k c_k\)  
(9)
Prediction error covariance = \(p_k = S_k p_{k-1} + S_k + n_k\)  
(10)

From (9) and (10), the prediction state estimation, is also called as a prior state estimation because doesn’t contain observation information about the current state.

Hence, the updated a posteriori state estimation is further derived as follows:-

\[ X_k(t) = X_k(t) + L_k(t) r_k(t) \]  
(11)
The Updated (a posteriori) covariance estimate is derived as follows:-

\[ p_k(t) = (1 - L_k(t)) p_k(t) \]  
(12)
From the above equation (11) and (12), the element “\(X_k(t)\)” denotes the updated state estimation of the “\(k^{th}\)” frame at discrete time ‘\(t\)’ and “\(L_k(t)\)” is the information obtained at discrete time ‘\(t\)’.

The information about the new associations between the foreground and shadow blobs, and their respective Kalman filters are updated. Top-down approach improves the detection of chromatic shadows and updates KF values of the foreground and shadow images correctly at the tracker end. The updated shadow images are shown in figure 6.

![Figure 6. (a) Original frame (b) Associations between foreground and shadow blobs](image)

3.4 Kernel Pattern Segment Function for Moving Object Detection

The next step is to categorize pixels belonging to foreground moving objects and background. Therefore, the moving objects are detected by using kernel pattern function. The kernel pattern function can be defined in terms of input mapping and be can be formulated as follows:-

\[ K_m(x, y) = \varnothing(x), \varnothing(y) \]  
(13)
From (13), “\(K_m\)” is the kernel pattern segment function and “\(\varnothing(x), \varnothing(y)\)” denotes an inner product between two seed points in the frame. The kernel function uses frequent patterns mined from a set of objects in a frame. The kernel with two seed points \((x, y)\) can be represented using feature vectors of a frame and can be formulated as follows:-

\[ K_m(x, y) = \exp\left(\frac{\|x - y\|^2}{2\sigma^2}\right) \]  
(14)
From (14), \(\|x - y\|\) denotes a squared Euclidean distance between the two seed points and “\(\sigma\)”is a free parameter.

The matching relations of the seed points in two successive video frames can be formulated as follows:-

\[ P_1 = AP_2 \]  
(15)
The equation (15) can further be described as follows:-

\[ \begin{bmatrix} r_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} a_0 & a_1 & a_2 \\ a_3 & a_4 & a_5 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ y_2 \end{bmatrix} \]  
(16)
The Euclidean distance between the two seed points are measured and the threshold value are assigned to classify the foreground and background region. If the Euclidean distance measure is greater than the threshold value “(\(T_h\))” then it is set as a foreground (FG) seed point, otherwise, it is set as a background (BG) seed point.

The foreground seed point of the “\(K^{th}\)” frame can be formulated as follows:-

\[ P_k = \{p(x, y, k) | \forall p(x, y, k) \in FG\} \]  
(17)
Similarly, the foreground seed point in the \((k-1)^{th}\) frame is defined as follows

\[ P_{k-1} = \{p(x, y, k-1) | \forall p(x, y, k-1) \in FG\} \]  
(18)
The foreground seed point of the “\(K^{th}\)” and “\((K-1)^{th}\)” frames.
The foreground seed point of the two frames are extracted using functions defined in equations (17) and (18).

![Figure 7(a) FG seed point of kth frame](image)

![Figure 7(b) FG seed point of (k-1)th frame](image)

As shown in figure 7 (a) and (b), the foreground seed points of the two video frames are obtained for pattern matching. Pattern matching uses the foreground seed points of the “(k-1)th” frame to find the equivalent foreground seed points of the kth frame.

Finally, pattern matching can be formulated as follows:

\[ f(P_{(x, y, k-1)}) = P_{(x + \Delta x, y + \Delta y, k)} \]  \hspace{1cm} (19)

\[ F(P_{k-1}) = \{ f(P_{(x, y, k-1)}) \} \forall P_{(x, y, k-1)} \in P_{k-1} \]  \hspace{1cm} (20)

From (19) and (20), the matching function “f” and “P_{(x,y,k-1)}” represents the foreground seed point of the (k-1)th video frame. Here, \( P_{(x+\Delta x,y+\Delta y,k)} \) denotes an updated foreground seed point of the kth video frame. Finally, the FG regions of the moving objects can be formulated as follows:

\[ P'_{k} = F(P_{k-1}) \cup P_{k} \]  \hspace{1cm} (21)

In (21), “\( P'_{k} \)” represents the FG video frame and “F” represents the updated matching function. The obtained foreground regions and the updated foreground seed points are combined to obtain the regions of moving objects.

![Figure 8. Pattern matching with updated FG feature points](image)

Figure 8 shows the updated foreground feature points from the kth frame and (k-1)th frame.

3.5. Extended-Kalman Filter-based Tracking of Moving Objects (EKF-TMO) method

After the detection of moving objects in a video frame, object tracking is carried out using EKF-TMO method. The proposed method performs quantization of signals to reduce the information loss and improves the tracking performance. The center of seed point affinity feature values of the moving object region after quantization are performed as follows:

\[ C = \frac{1}{N} \sum_{i=1}^{n} P(x_i, y_i) \]  \hspace{1cm} (22)

In (22), “C” represents the center of the seed point affinity, and “N” denotes the total number of seed point in the region of moving object. Here, “\( P(x_i,y_i) \)” is the seed points of the moving objects’ region. Next, the bounding box approach is applied to detect the moving objects. The minimum bounding boxes of the moving objects are obtained using two vertical segmentations and one horizontal segmentation. The moving object search region (SR) is measured to perform tracking of objects more accurately within the bounding box. The search region (SR) is measured as follows:

\[ SR = \rho W_{Box} + \tau H_{Box} \]  \hspace{1cm} (23)

From (23), “\( W_{Box} \)” denotes a width of the bounding box and “\( H_{Box} \)” is the height of the bounding box. Where, “\( \rho \)” and “\( \tau \)” are the constant parameters.

The object with the bounding boxes and the final object tracking obtained is shown in figure 9.
Figure 9 clearly shows tracking the moving object with bounding box. The moving object tracking algorithm is described as follows:

1. **Step 1:** For each detected moving object, calculate the center of seed point affinity with Kalman Filter using (15).
2. **Step 2:** The tracked object is searched within the search region SR using (16).
3. **Step 3:** Obtain tracked objects to find the location of every moving object in a video frame.

**Algorithm 1. Extended-Kalman Filter-based Tracking of Moving Objects method**

The algorithm designed for moving object tracking involves identifying images from video frames and tracking its movement and position. For each detected object, the center of the seed point affinity value is calculated. Next, the search region is measured to locate the prediction of all moving objects using Kalman filter with surrounding bounding boxes.

**IV. EXPERIMENTAL SETUP**

An experimental evaluation of the proposed methods is implemented in MATLAB. The research work is simulated using Actions as space-time shapes dataset. This dataset consists of video sequence of different sizes. Video frames ranging from 10 to 100 frames per second, are given as input to the proposed system.

In the first phase of classification, the proposed method Spatial Classification of Multi-Class Objects (SCMO) is compared with the existing BISVM classifier [1] and multi-sensor fusion framework [2] methods for classification of objects.

In the second phase, the proposed Extended-Kalman Filter-based Tracking of Moving Objects (EKF-TMO) method is compared with the existing Scale Invariant Feature Transform technique (SIFT) in [3] and Enhanced Rao-Blackwellized Particle Filter (E-RBPF) in [4] respectively for pattern matching of objects and tracking them.

The performance is carried out on the factors such as classification time, classification accuracy, pattern matching time, pattern matching accuracy, true detection rate and object tracking accuracy. Tables below describe the efficiency of these methods.

**V. RESULTS AND DISCUSSION**

5.1 **Classification Accuracy and Classification Time**

Classification accuracy (CA) as in equation (24) is defined as the ratio of the number of objects that are correctly classified to the total number of objects in video frames. The formula for classification accuracy is defined in percentage as follows:

\[
CA = \frac{\text{Correctly classified objects}}{\text{No. of objects in video frames}} \times 100
\]  

(24)

Classification time (CT) as in equation (25) is defined as the time taken to classify the object in video frames which is measured in terms of milliseconds (ms).

\[
\text{classification time} = \text{No. of objects in video frame} \times \text{time taken for classifying one object}
\]  

(25)

<table>
<thead>
<tr>
<th>No. of Video frames/sec</th>
<th>Classification Accuracy (%)</th>
<th>Classification Time (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SCM</td>
<td>BISVM</td>
</tr>
<tr>
<td>-------------------------</td>
<td>-----</td>
<td>-------</td>
</tr>
<tr>
<td>10</td>
<td>70.25</td>
<td>60.10</td>
</tr>
<tr>
<td>20</td>
<td>75.64</td>
<td>65.85</td>
</tr>
</tbody>
</table>
Table 1 shows that the classification accuracy is improved by the proposed Spatial Classification of Multi-Class Objects (SCMO) method by 11% and 24% compared to the existing BISVM in [1] and Multi-sensor fusion framework methods in [2] respectively. Also, the time taken for classification of objects is reduced by the proposed Spatial Classification of Multi-Class Objects (SCMO) method by 15% and 25% compared to the existing BISVM in [1] and Multi-sensor fusion framework methods in [2] respectively.

5.2 True Detection Rate and Pattern Matching Time

True detection rate is defined as the ratio of the number of shadow image of the objects being correctly detected to total number of objects in video frames.

\[
TDR = \frac{\text{No. of shadow objects correctly detected}}{\text{No. of objects in video frame}} \times 100
\]  

(26)

In (26), “TDR” represents the true detection rate, which is measured in terms of percentage (%).

Pattern matching time can be defined as the time taken by the proposed method to match patterns with the training patterns with respect to number of objects in video frames. Pattern matching can be formulated as follows: 

\[
PMT = \text{No. of objects in video frame} \times \text{time taken for matching}
\]  

(27)

From (27), “PMT” denotes pattern matching time, measured in terms of milliseconds (ms). A method is considered to be more efficient, if the time taken for pattern matching is minimum (ms).

Table 1 Tabulation for true detection rate

<table>
<thead>
<tr>
<th>No. of Video frames/sec</th>
<th>True detection rate (%)</th>
<th>Pattern Matching time (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EKF-TMO</td>
<td>SIFT</td>
</tr>
<tr>
<td>10</td>
<td>73.65</td>
<td>61.36</td>
</tr>
<tr>
<td>20</td>
<td>76.25</td>
<td>63.47</td>
</tr>
<tr>
<td>30</td>
<td>80.12</td>
<td>65.11</td>
</tr>
<tr>
<td>40</td>
<td>83.65</td>
<td>70.12</td>
</tr>
<tr>
<td>50</td>
<td>85.10</td>
<td>75.38</td>
</tr>
<tr>
<td>60</td>
<td>88.64</td>
<td>78.65</td>
</tr>
<tr>
<td>70</td>
<td>89.12</td>
<td>80.24</td>
</tr>
<tr>
<td>80</td>
<td>91.34</td>
<td>83.65</td>
</tr>
<tr>
<td>90</td>
<td>93.47</td>
<td>85.12</td>
</tr>
<tr>
<td>100</td>
<td>94.36</td>
<td>88.69</td>
</tr>
</tbody>
</table>

From Table 2, the TDR is considerably increased by the proposed Extended-Kalman Filter-based Tracking of Moving Objects (EKF-TMO) method by 14% and 32% compared to existing methods SIFT [1] and E-RBPF [2] methods respectively.
Table 2 illustrates PMT with respect to number of video frames. The proposed Extended-Kalman Filter-based Tracking of Moving Objects (EKF-TMO) method has reduced the matching time by 20% and 32% compared to the existing SIFT [1] and E-RBPF [2] methods respectively.

### 5.3 Impact of Moving Object Tracking Accuracy

Moving-object Tracking Accuracy (MTA) is defined as the ratio of number of objects being tracked to the total number of video frames per second. Moving object tracking accuracy is measured in terms of percentage (%) and is formulated as follows:

\[
\text{MOTA} = \frac{\text{No. of Objects being Tracked}}{\text{Number of objects in video frames}} \times 100
\]

Here, “MOTA” represents moving object tracking accuracy.

Table 3 illustrates that the MTA is improved by the proposed Extended-Kalman Filter-based Tracking of Moving Objects (EKF-TMO) method by 7% and 21% when compared to the existing SIFT [1] and E-RBPF [2] methods respectively. Finally from tables 1-3, it is clear that the proposed methods SCMO and EKF-TMO improves the overall performance of detection and tracking of moving objects.

### VI. CONCLUSION AND FUTURE WORK

An efficient Spatial Classification of Multi-Class Objects (SCMO) and Extended-Kalman Filter-based Tracking of Moving Objects (EKF-TMO) method were introduced to achieve efficient and robust moving object detection and tracking. The method makes use of Markov Random Field (MRF) principle to estimate the shape of moving objects. The fuzzy based approach is used to identify the edge-strength of each pixel location. This helps to preserve the object boundary. The spatial classification based on MAP estimation principle is performed and objects are classified into classes. The chromatic shadows of objects are detected using a top-down approach. Next, kernel pattern segment function is applied to formulate the seed point. Next, bounding box is applied to detect the moving object. Finally, the moving object tracking is achieved by using Extended-Kalman Filter-based Tracking of Moving Objects in a particular region with minimum bounding box. The experimental result reveals that the QKF-PM technique significantly improves the pattern matching accuracy with minimum time and improves the true detection rate. The QKF-PM technique also improves the moving object tracking accuracy compared to the existing state-of-the-art methods. The proposed methods can be combined with memory optimization algorithms for a more powerful and timely computational results.
REFERENCES


