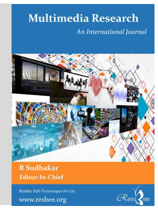
Vol.7 No.2 April 2024

Multimedia Research

Received 19 October, Revised 8 November, Accepted 13 March



Multi-Object Tracking Using Improved Lion-Based Clustering with Reference to LVP Pattern

Marumalla Meher Aditya
Department of CSE
Neil Gogte Institute of Technology, Hyderabad, India
aadhimarumalla@gmail.com

Abstract: MOT in video surveillance is still a challenging aspect due to various complexities like complex occlusion, different poses, small sizes of objects, and so on. Henceforth more research works are under process that concentrate on solving object tracking problems by considering both spatial and visual information. Under these circumstances, this paper aims to propose a new MOT model that tracks the objects precisely. At first, the visibility model of tracking is proposed based on the second derivative model, which considers the second derivative function to predict the objects. Secondly, a spatial tracking model is developed using a nonlinear function. Meanwhile, the objects are tracked by subjecting a new optimization algorithm namely Female Update enabled Cub Updating in LA (FU-CU-LA) that effectively tracks the objects even if moved to the subsequent frames. Subsequently, spatial tracking and optimization-based tracking are integrated, and finally, the resultant center point is integrated with the visual tracking model. The proposed FU-CU-LA-based tacking system is compared with other models to certain measures and proves the superiority of the proposed work.

Keywords: Multi-Object Tracking; Video Surveillance; Optimization; LA; Spatial Tracking Model

Nomenclature

Abbreviation	Expansion
MOT	Multi-Object Tracking
HDAT	Hierarchical data association tracking
NN	Neural Network
CNN	Convolutional Neural Network
MHT	Multiple Hypotheses Tracking
RAFMPL	Relational Appearance Features and Motion Patterns Learning
HOG	Histogram of Oriented Gradient
SDVM	Second Derivative Visual Model
NFSM	Nonlinear function based spatial model
TWSM	Tangential Weighted Spatial Model
MOTP	Multiple Object Tracking Precision

1. Introduction

Real time multi object tracking based on visual observations is an important process for any high-level computer vision application, such as activity recognition, event categorization, visual geo navigation, virtual reality, etc. Object tracking based on visual and temporal cues is defined as the problem of estimating the spatial locations of an object in the image plane as it moves using the information extracted from the video [9]. Real time MOT [10] [11] [12] is a challenging problem, that remains at the heart of recent research in computer vision. The complexity of multi object tracking is caused by many situations: (1) when multiple objects move close to each other or present occlusion, identity management becomes a very complex operation; (2) complex occlusion with static objects in the scene may occur. Solving the hijacking and data association problems in such situations is challenging, especially in complete occlusion situations.

So far, many algorithms have been proposed for this object-tracking process, and some of them are the expectation maximization model, HDAT framework, etc. Moreover, some of the deep learning techniques like NN and CNN have been used for this purpose. Further, in some cases, multiple distinctive trackers are used to track the objects when there is no interaction among the observations.

https://doi.org/10.46253/j.mr.v7i2.a2

More researchers have involved their research work under many meta-heuristic approaches that are based on swarm optimization. The interaction between objects is modeled as species competition and repulsion. Game theory is used for MOT [13] [14] [15] [16], in which the tracking problem is solved by finding the Nash equilibrium of a game between multiple independent trackers. Distributed MOT approaches for multi-object [23] [24] [25] tracking is proven to be more adapted for real-time applications [17] [18] [19] [20] because the complexity of the multi-object configuration grows linearly in terms of the number of tracked objects.

Nevertheless, the major problem of the MOT [21] [22] is the complexity of modeling the interaction among independent trackers, especially when objects present complex occlusions. The performance of state-of-art object detectors is far from perfect. Missed detections, false positives, and inaccurate responses occur frequently in the detection process especially when using rapid detectors. Data association in multi-object tracking, such as MHT and other methods meets the challenges caused by error of detection, occlusion, and similar appearance among multiple objects. To overcome these challenges we have proposed a new one that tracks objects precisely by considering both spatial and visual information. It is a visibility model for tracking based on the second deviation model to predict the objects. The spatial model is used for non-linear functions to track objects. A new optimization algorithm called Female Update enabled Cub Updating in LA (FU-CU-LA) is used for effectively tracking objects even if they move to subsequent frames. Then the model is compared with other models to find its superiority and achieve better tracking performance on multiple objects compared to conventional models.

In this paper, section 2 describes the literature review of models. A summary of the proposed multiple object tracking: a new nonlinear function-based spatial tracking model approach is represented in section 3. Section 4 portrays the visual and spatial tracking model: a new workflow for multi-object tracking. Section 5 depicts the inter-frame step size optimization: introducing improved LA and section 6 portrays results and discussions. Section 7 explains the advantages and disadvantages and Section 8 concludes the paper.

2. Literature Review

2.1 Related Works

In 2024, Ruoying Liu et al., [1] presented PDT-YOLO algorithm and intra-scale feature interaction module (AIFI) to enhance detection accuracy. A lightweight convolution module (GSConv) was introduced to construct a multi-scale efficient layer aggregation network module (ETG) for improved feature extraction. Multi-attention mechanisms were integrated to optimize feature expression for occluded targets. The Wise-IoU with a dynamic non-monotonic focusing mechanism was utilized to enhance accuracy and generalization. PDT-YOLO algorithm improved mAP50 and mAP50:95 by 4.6% and 12.8% on the DAIR-V2X-C dataset and by 15.7% and 11.1% on the IVODC dataset. The algorithm achieved a detection frame rate of 90 FPS in an actual traffic environment, meeting the needs of roadside object detection in complex traffic scenes.

In 2017, Gwak [2] implemented a new data association model called RAFMPL that would facilitate MOT. Here, the developed relational feature-based models were entirely more varied than the traditional models, where they generated the tracks that were based on relational data by choosing a single reference object and utilizing the feature differences among reference and other objects. Additionally, the motion model has enabled motion pattern learning (linear and nonlinear patterns) that was concerned with data association. Further, the developed model rectified all the complexities in MOT by improving the robustness of occlusion issues that occurred in real-time scenarios. Finally, the results have proved the performance of work over other models.

In 2018, Li and Cheng [3] implemented an online learned Hough forest approach that was based on enhanced fusion matching of multi-feature. At first, the samples in terms of positive and negative were chosen as per the low-level association between detection responses. Then the feature models were constructed using color histogram-like HOG along with optical flow data. After this, they generated the longer trajectory associations based on the online learned Hough forest model. At last, they developed a trajectory-matching algorithm that was based on multi-feature fusion. The probability matrix was generated along the weighting factor with the introduction of similarity measures and feature point matching. Finally, the model was evaluated over other traditional models for certain data sets.

In 2018, Li *et al.* [4] proposed a fuzzy logic data association model for MOT. Initially, the authors designed the fuzzy interface that incorporated expert experience into the data association, which was to enhance the MOT performance. With the consideration of errors and alterations in shape, error of motion, and appearance approaches, all the rules were utilized for determining the degrees of fuzzy membership that substitute the probabilities of association among objects and the measurements. Next, to deal with the fragmented trajectories, they developed a track-to-track association model, which was based on a fuzzy synthetic function. Further, the model needs no statistical model assumptions in terms of measurement noise. The experimentation was carried out under different datasets and has proven the effectiveness of the model for the reduction of the count of fragment tracks.

In 2018, Liye*et al.* [5] developed a new enhanced algorithm namely the local excitatory global suppression oscillation network model that was utilized in target tracking and image preprocessing. The correction army has split and has combined both minimum and maximum values of neighboring pixels. The preprocessing time has been improved using a grating scanning model, iterative integration, and segmentation process. To minimize the segmentation time, the previous iteration pixels were not considered in the current iteration. Finally, the developed model was executed in Xilinx ISE 14.2 with ISIM-based simulation. When compared to the conventional models, the model has attained minimal segmentation time.

In 2024, Ranyang Zhao *et al.* [6] implemented a robust and efficient two-stage joint model based on R-FCN for online multi-object tracking, utilizing adaptive sparse anchoring, feature aggregation, feature disentanglement, and position-sensitivity to improve efficiency and robustness against occlusion. The model leverages position-sensitivity as an effective tool for addressing both interclass and intra-class occlusion in natural scenes, using a two-stage approach for detection and ReID feature extraction simultaneously. The Re ID feature vector of the proposal was calculated based on specific equations involving position vectors and feature matrices. PSMOT model achieved competitive performance in multi-object tracking while maintaining time efficiency. The model effectively addressed occlusion by exploiting position-sensitivity, leading to robustness against occlusion in tracking scenarios.

In 2024, Yu-Hsiang Wang*et al.* [7] presented SMILE track integrates an object detector with a Siamese network-based Similarity Learning Module (SLM) to address challenges in Multiple Object Tracking (MOT). The SLM calculates appearance similarity between objects using a Patch Self-Attention (PSA) block and a Similarity Matching Cascade (SMC) module with a GATE function for robust object matching across frames. SMILE track achieved an improved trade-off between running speed and tracking accuracy over existing benchmarks, outperforming BYTE Track on MOT17 and MOT20 datasets. Different patch layouts were evaluated to improve SLM performance, with type-E patch layout showing the best results.

In 2018, Ju et al. [8] developed a data association model, which has efficiently exploited structural constraints in the presence of huge camera motion. Additionally, to reduce the incorrect associations along with false positives and misdetections, they have developed a new event aggregation model that has integrated the evaluated assignment costs via structural constraints. They have also used structural constraints for tracking the missing objects while redetection. With this process, the missing object identities could be continuously retained. Finally, the validation of the model over other traditional models was continued in terms of certain measures.

2.2 Review

Table I shows the features and challenges of state-of-the-art MOT models. Here, the Stop and Go model [1] is robust and it could minimize the track-matching errors. However, the ability of convergence is less and the model doesn't work in multi-camera systems. RAFMPL model [2] maximizes the tracking accuracy rate but it doesn't deal with more objects or classes. Moreover, the model doesn't work in real-time applications and it requires long-term changes. Hough forest model [3] resolves the problems in object deformation and it is more robust too. It also works in complex scenarios. However, additional improvements are needed to attain accurate results. Fuzzy logic [4] is quite understandable and also has an easy reasoning process. Here, the interpolation is very complex and the model fails if there are restrictions in memory and time. The weak clustering model [5] has attained a high tracking rate and the model is efficient in the static background with a high tracking rate. However, the real-time application is impossible and more extension work is needed for handling dynamic background. HSV-local binary pattern histogram-based appearance model [6] can reduce the identity switches and also have high prospects. However, the model requires weakly supervised learning models and the model doesn't obtain the sparse globally detected results. The authors of [7] have used the Gaussian mixture model that could attain a high tracking rate. The model also can handle complicated scenarios. However, the knowledge collection is complex. The data association model [8] could efficiently recover the missing objects and the model could handle a huge count of datasets. However, improvement is needed to attain error-free object tracking. Thus, some advanced and enhanced MOT models rectify the above-mentioned drawbacks.

Author [citation]	Methodology	Features	Challenges
Ruoying Liu et al., [1]	PDT-YOLO algorithm and intra-scale feature interaction module (AIFI)	• The algorithm enhanced the detection accuracy of multi-scale targets, optimized feature expression for occluded targets, and improved model sensing accuracy and generalization ability.	• The algorithm may still face challenges in scenarios with high occlusion levels, potentially leading to false detections.
Gwak [2]	RAFMPL model	• Increases the accuracy rate.	• Dealing with more classes (objects) is difficult. • Real-time application was not possible. • Needs long-term changes.
Li and Cheng [3]	Hough forest model	• Solves the issues of object deformation. • More robust. • Work on complex scenarios.	• Further enhancement was needed to attain the accuracy rate of object tracking.
Li *et al.* [4]	Fuzzy logic	• Not sensitive to altering environments and erroneous rules. • Quite Understandable. • Has a simple reasoning process.	• Interpolation was quite complex. • Fails ineffective process if there exist restrictions in time and memory.
Liye *et al.* [5]	Weak clustering model	• Efficient in static background. • The tracking rate is high.	• Real-time application was impossible. • Extension was needed under a dynamic background.
Ranyang Zhao et al. [6]	R-FCN	• The method boosted both detection and Re ID performance. • Enhanced robustness against occlusion, leading to outstanding performance and time efficiency.	• Currently implemented only for tracking pedestrians, limiting its applicability to other scenarios.
Yu-Hsiang Wang et al. [7]	SMILE track	• Improved the trade-off between running speed and tracking accuracy. • Robust object matching across frames, leads to enhanced MOT performance.	• Struggle with rapidly moving objects, potentially leading to increased identity switches.
Ju *et al.* [8]	Data association model	• Can effectively recover the missing objects. • Can use a large number of datasets.	• Extension work was needed for error-free object tracking.

3. Proposed Multiple Object Tracking: A New Nonlinear Function-Based Spatial Tracking Model

The proposed multiple object tracking model is diagrammatically illustrated in Fig 1. The representation shows the structure of the working principle, where there are several steps:

Step i: Getting the input video is the first step, from which the object area is identified using certain steps in terms of both neighbor and texture pixels.

Step ii: Subsequently, apply the second-order derivative-based visible tracking model in the frames that are extracted.

Step iii: Meanwhile, apply the proposed non-linear function-based spatial tracking model in the extracted frames.

Step iv: Finally, combine both the visibility model and spatial tracking model and attain the tracked output.

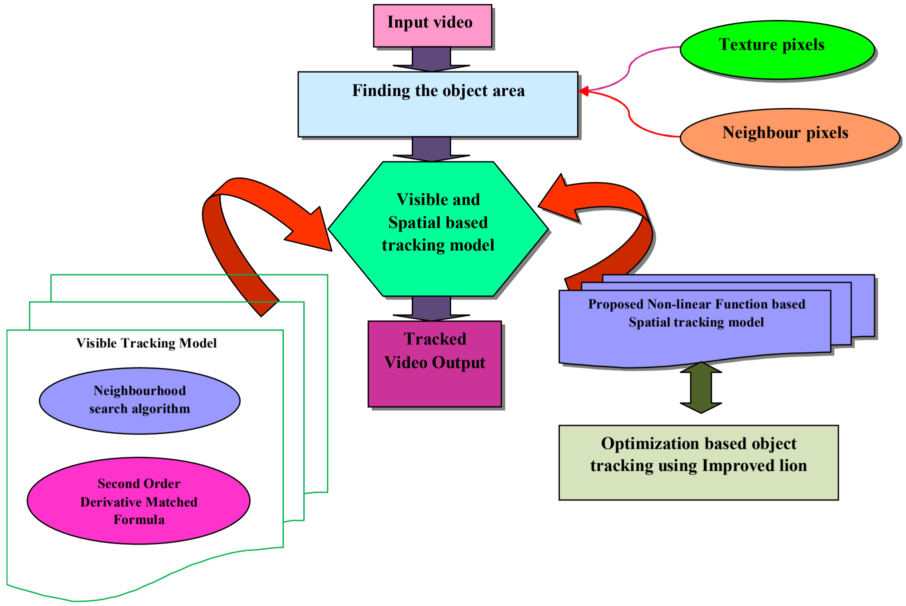


Fig. 1. *Block Diagram of Proposed Algorithm*

3.1 Object Area Identification

This is the first or foremost step of the proposed object-tracking framework. Initially, the input video (V) is read and extracts the frame V_i that determines the i-th frame of the video given in Eq. (1), where, j is a column of pixels that can be varied from 1 to N, and k is a row of pixels varying from 1 to M.

$$V = \{V_i; \quad 1 \le i \le n\} \tag{1}$$

The object area is identified with the consideration of pixels (texture and neighbor pixels) by the k-means clustering algorithm.

k-Means clustering: The k-means clustering algorithm [23] groups the objects into multi-groups k number of groups. where K is any of the positive integers. The grouping is based on the minimization of the sum of squares of distance among corresponding objects as well as the selected cluster center. The overall steps are given below:

(i) Randomly choose the cluster centres $R_{i,j}$ based on the range of needed clusters C as defined in Eq. (2).

$$R = \{R_{i,j}; \qquad 1 \le i \le C; 1 \le j \le 2\} \tag{2}$$

(ii) Repeat

(iii) Evaluate the distance among each object o_i and $R_{i,j}$, Assign the object o_i to the nearest cluster.

$$Euclidian\ Dis\tan ce(o_i, R) = \sqrt{\sum_{i=1}^{n} (o_i - R)^2}; \qquad 0 \le k \le m \tag{3}$$

(iv) For each cluster k, recalculate until no change in the cluster center. Subsequently, the count of objects chosen based on cluster value C is given as in Eq. (4).

$$R_1, R_2, \ldots R_C \tag{4}$$

4. Visual And Spatial Tracking Model: A New Workflow for Multi-Object Tracking

4.1 Visual Tracking using Second Order Derivative Model and Neighbourhood Search

In general, visual tracking [30] is a basic model to analyze the processing of video motion, visual surveillance, data mining, vision-based control, human-machine interactions, and so on. The Object appearance model of the visual tracking model is based on kernel density estimates, region color histograms, gaussian mixture model, etc. Nevertheless, the visual tracking of multi-object has been an important research topic due to some of the factors including the interaction of one another that poses inter-object occlusion, low probability of detection, and object confusion. To overcome all the mentioned drawbacks, a second derivative model and neighborhood search-based visual tracking algorithm are introduced, which is termed SDVM.

Neighborhood search algorithm: The object tracking by the Neighbourhood search algorithm [31] is processed with a few steps:

Step 1: Make a fix of the fixed reference point R_i^t in the key frame by template function as in Eq. (5).

$$R_i^t \in f(x,y) \tag{5}$$

Step 2: Reference point extraction in a new frame that is based on key frame reference point as defined in Eq. (6), where R_i^{t+1} is the reference point that is extracted and $f\left(r^k\right)$ indicates the function for sub-image generation.

$$R_i^{t+1}(x,y) = f\left(r^k\right) \tag{6}$$

Step 3: Identifying the Euclidean distance between the key frame and the frame pixel that is extracted as per Eq. (7).

$$ED(R_i^t, R_i^{t+1}) = \sqrt{\sum_{i=1}^{n}\left(R_i^t - R_i^{t+1}(x,y)\right)^2} \; ; \; 0 \le K \le m \tag{7}$$

Step 4: Find the new position by increasing or decreasing the reference point via unity, which is given in Eq. (8), which (i,j) indicates the reference point decreasing or increasing.

$$r_i^{t+1}(x,y) = \{x \pm i, y \pm j\} \tag{8}$$

Step 5: From the given input image, extract the new region and identify the Euclidean distance between the extracted region and the given input image.

Step 6: Repeat steps 4 and 5 till it attains the predefined search threshold. In the completion of step 5, the reference point having the least distance is considered as the detected images' final region, and it is defined in Eq. (9).

$$R_i^{t+1}(x,y) = \arg\left(\underset{K=1}{Min}\, P_K\right) \tag{9}$$

Since the abovementioned algorithm is not so successful in tracking the image (in case, there is interaction among multi-objects and frequent occultation), a second derivative model along the neighborhood search algorithm is developed [28]. The next spatial position of the object by second derivative-based visible pixels is defined in Eq. (10), (11) and (12), where R_i^t refers to the key frame's reference point, $R_i^{t+1}(x,y)$ indicates the next frame's reference point, $N_{xx}(t)$, $N_{xx}(t+1)$ indicates the second derivative of key frame and next frame based on x direction, $N_{yy}(t)$, $N_{yy}(t+1)$ are the second derivative of key frame and next frame based on y direction, ED refers to the Euclidean distance among key frame as well as next frame of reference point.

$$P_K = \alpha \times ED\left(R_i^t, R_i^{t+1}(x,y)\right) + (1-\alpha) \times ED\left(\frac{\partial^2 R_i^t}{\partial x^2}, \frac{\partial^2 R_i^{t+1}(x,y)}{\partial x^2}\right) + (2-\alpha) ED\left(\frac{\partial^2 R_i^t}{\partial y^2}, \frac{\partial^2 R_i^{t+1}(x,y)}{\partial y^2}\right) \tag{10}$$

$$\frac{\partial^2 R_i^t}{\partial x^2} = \frac{\partial^2 f(x,y)}{\partial x^2} = N_{xx}(t) \;\; ; \;\; \frac{\partial^2 R_i^t}{\partial y^2} = \frac{\partial^2 f(x,y)}{\partial y^2} = N_{yy}(t) \tag{11}$$

$$\frac{\partial^2 R_i^{t+1}(x,y)}{\partial x^2} = \frac{\partial^2 f(x,y)}{\partial x^2} = N_{xx}(t+1) \; ; \; \frac{\partial^2 R_i^{t+1}(x,y)}{\partial y^2} = \frac{\partial^2 f(x,y)}{\partial y^2} = N_{yy}(t+1) \tag{12}$$

4.2 Spatial Tracking Using Proposed Non-linear function and Optimization Object tracking

This section explains the spatial tracking method [32] that can predict the subsequent spatial position of an object by a new NFSM**.** In this, there presentation of the object is in the form of x, y coordinate in the key frame. The respective object of the subsequent frame is predicted as per the direction of x, y coordinate change. Subsequently, the same object's position in the next frame is predicted as per the difference among object coordinates in the key frame with the subsequent frame. With the utilization of this spatial method, the frame's objects are tracked accurately. Here, the object's center point value is chosen for further evaluation, and the representation of a chosen center point in a cluster is as $r_i^t \ldots\ldots r_C^t$, which r_i^t indicates the i^{th} object in t^{th} frame. With the aid of the mentioned center point of r_i^t objects, the spatial tracking model can be utilized as per the formula given in Eq. (13), were Tr_i^{t+1} indicates the i^{th} object in $t+1$ frame. The object r_i^t belongs to (i,j)

$$Tr_i^{t+1} = \alpha r_i^t + (1-\alpha) Tr_i^t \tag{13}$$

α is the smoothing factor that is defined with the function of exponential. To enhance or improve the tracking performance (by avoiding several problems), this paper aims to introduce a non-linear function-based formulation, thereby the prediction of the next location by the proposed function is defined in Eq. (14), where τ is the time constant and CO indicates the weighted constant.

$$\alpha = 2\left(\exp(\tau) - 1 + CO\right) \tag{14}$$

Optimization-based Object tracking: After this, the optimization concept is also used to track the object more accurately and optimally to determine the movement of the object (frame to frame). Here, the respective object of the subsequent frame is predicted as per the direction of x, y coordinate change. Subsequently, the same object's position in the next frame is predicted as per the mean difference among LVP [34] patterns of the object in the key frame with subsequent frames $t+1$. With the aid of this, the frame's objects are accurately tracked. Hence the optimization-based tracking is indicated as Ob_i^{t+1}. Finally, the integration of both Ob_i^{t+1} (the object's center point in the optimization concept) and Tr_i^{t+1} (the object's center point in spatial tracking) is determined as per Eq. (15).

$$Ot_i^{t+1} = \left|\left\{Ob_i^{t+1}\right\} + \left\{Tr_i^{t+1}\right\}\right| \tag{15}$$

4.3 Integration of Visual and Spatial tracking model

Both the obtained visual and spatial tracking models are integrated; thereby obtaining better tracking performance. The combination of the second-order derivative function and nonlinear function of the spatial tracking model helps to overcome certain drawbacks of multi-object tracking including interactions among multiple objects, similar object appearance, frequent occlusions, and so on. The integration of both models is mathematically defined in Eq. (16), which is the final tracking TR_i^{t+1}. Here, R_i^{t+1} indicates the selected reference point, $c\left\{R_i^{t+1}\right\}$ indicates the object's center point that was selected from R_i^{t+1}, Ot_i^{t+1} indicates the obtained object's center point in integrated spatial tracking and optimization-based tracking.

$$TR_i^{t+1} = \left[c\left\{R_i^{t+1}\right\} + \left\{Ot_i^{t+1}\right\}\right] \tag{16}$$

5. Inter-Frame Stepsize Optimization: Introducing Improved La

5.1 Solution Encoding and Objective function

In optimization-based tracking, it is very important to select the inter-frame step size of i^{th} object movement. This optimal selection is done by a new improved algorithm FU-CU-LA. The solution given to the proposed algorithm is given in Fig 2, where $d_1, d_2, \ldots d_N$ indicates inter-frame step size (for example value among -4 to 4). The objective of the proposed optimization-based tracking process is defined in Eq. (17), where $\hat{L}$ indicates the mean difference among the LVP pattern of i^{th} object at $t+1$ frame and the LVP pattern of i^{th} object at t.

$$OB = \min\left(\hat{L}\right) \tag{17}$$

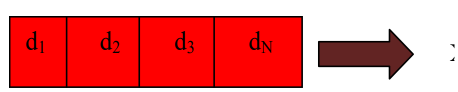


Fig. 2. *Solution Encoding*

5.2 Conventional Lion Algorithm

LA [26] comprises six processing phases namely, (a) Pride generation, (b) fertility evaluation, (c) mating, (d) territorial defense, (e) territorial takeover, and (f) termination.

Pride generation is the first process, which is the same as that of the initial process of many of the swarm-based optimization and evolution algorithms. New lions are produced from parent lions through the fertility evaluation, which is the second phase. The phase includes the most responsible process, which is called mating. Compared to other algorithms, Territorial defense, and territorial takeover are considered distinctive processes due to the precise inspiration of the social behavior of the Lion. The major role of the LA algorithm is purely based on two steps to find the optimal solution from a huge search space. The termination process proceeded at the end of the iteration/generation.

Problem Formulation: Let Eq. (18) be the objective function. From Eq. (18), the continuous uni-modal and multimodal function is specified as $f(\bullet)$ that has the solution space with size $\Re^m$, in which the real numbers are indicated by $\Re$, $z_i : i = 1,2,.....,m$ is concerned as the i^{th} solution variable and the solution vector dimension is specified as m, $z_i^{\max}$ and $z_i^{\min}$ are the maximum and minimum limits of i^{th} solution variable, correspondingly.

$$X^{opm} = \underset{z_i \in (z_i^{\min}, z_i^{\max})}{\arg\min} f(z_1, z_2, ..., z_m); \quad m \geq 1 \tag{18}$$

The optimal/target solution to be found from the provided optimization algorithm is specified as X^{opm} and is indicated by Eq. (20). The determination of solution space size of $f(\bullet)$ is defined in Eq. (19).

$$\Re^m = \prod_{i=1}^{m} (z_i^{\max} - z_i^{\min}) \tag{19}$$

$$X^{opm} = X : f(X) < f(X' \big| X' \neq X : z^r{}_i \in (z_i^{\max}, z_i^{\min})) \tag{20}$$

Here, the solution vector is indicated as X with the representation of $X = [z_1, z_2, ..., z_m]$. Eq (18) defines the objective function, which is the one to be solved as a minimization function.

Pride generation: According to the pride definition, it is initialized with the nomadic lion X^{nmad}, territorial lion X^{male}, and its lioness X^{female}. Pride does not have any nomadic lions; still, the generation is described in the process of pride generation. The lion representation is similar to that of the solution vector representation. While $m > 1$, the vector element of X^{male}, X^{female} and X^{nmad}, which means z_{lm}^{male}, z_{lm}^{female} and z_{lm}^{nmad} are concerned as the random integers having maximum as well as minimum limits, and in this, $lm = 1,2,....,Lm$. Further, the lion's length is specified as Lm, which is given in Eq. (21).

$$Lm = \begin{cases} m; & m > 1 \\ \overline{n} & otherwise \end{cases} \tag{21}$$

where as, the integers for deciding the lion's length are determined as m and $\overline{n}$. While $m = 1$, the searching is done on binary encoded lion and thus the vector element generation is made by 0 or 1, which is defined in Eq. (22) and (23).

$$gen(z_{lm}) \in (z_{lm}^{\min}, z_{lm}^{\max}) \tag{22}$$

$$\overline{n}\%2 = 0 \tag{23}$$

Where

$$gen(z_{lm}) = \sum_{lm=1}^{Lm} z_{lm} 2^{\left(\frac{Lm}{2} - lm\right)} \tag{24}$$

Here, Eq. (22) as well as (24) assures that the generated binary lion is in the solution space and Eq. (23) assures that the number of binary bits before and after the decimal point is equivalent.

Fertility Evaluation: In the chronological process of LA, each territorial lion and lioness becomes infertile and aged too, and thus the laggard lion forms. While the fitness is saturated the X^{male} and X^{female}, either the global or local optima is reached from which the best solution is not given. This fertility evaluation helps in skipping this local optima solution. Here, X^{male} turns to laggard and Lg_r that is the laggardness rate is increased by one, when $f(X^{female}) > f^{reff}$, where f^{reff} indicates the fitness of reference. If Lg_r the value goes beyond the greatest limit $Lg_r^{\max}$, then territorial defense happens. In X^{female}, the fertility is assured using the sterility rate St_r and after crossover, it increased by one. The updating is done as per Eq. (25) for X^{female} when St_r going beyond the $St_r^{\max}$. The mating process is done when the updated female $X^{female+}$ is considered X^{femle} due to its enhancement. Contrast to this, till the female generation gen_c count attains $gen_c^{\max}$. Throughout the process of updating, if there is no $X^{female+}$ to replace X^{female}, it is assigned as the female is fertile for generating the cubs.

$$X_{lm}^{female+} = \begin{cases} z_h^{female+} & ; if\ lm = h \\ z_{lm}^{female} & ; otherwise \end{cases} \tag{25}$$

$$z_h^{female+} = \min\left[z_h^{\max}, \max(z_h^{\min}, \nabla_h) \right] \tag{26}$$

$$\nabla_h = \left[z_h^{female} + (0.1rd_2 - 0.05)(z_h^{male} - rd_1 z_h^{female}) \right] \tag{27}$$

Where, l^{th} and k^{th} vector elements of $X^{female+}$ are given as $z_{lm}^{female+}$ and $z_h^{female+}$ respectively, the random integer is indicated as h and is generated that ranges in $\left[l, Lm\right]$, the female update function is defined as ∇, the random vectors are specified as rd_1 and rd_2, which is produced between the range [0,1].

Mating: This LA does the mating process by two fundamental steps and one complementary step. The initial steps are the crossover and mutation process, and the supplementary step is the gender clustering. Various research works are argued in the existing literature on the operations of crossover and mutation and their requirement for evolutionary algorithms. X^{male} and X^{female} gives birth to cubs by using the crossover and mutation process, where the derivation of solution is done from the elements of X^{male} and X^{female}. The crossover process gives four cubs.

Lion operators: Territorial defense grants a broad search of solution space and also directs the algorithm for evading from the local optimal point that identifies the diverse solution with equivalent fitness. The ordering of territorial defense can be made in this as producing survival fight, nomad coalition, and then pride and update of nomad coalition. By applying a winner-take-all approach, the simplification of the nomad coalition process takes place for identifying X^{e-nmad}. Subsequently, when the criteria given in Eq. (28), (29) and (30) are met, X^{e-nmad} is chosen.

$$f(X^{e-nmad}) < f(X^{male}) \tag{28}$$

$$f(X^{e-nmad}) < f(X^{male\_cub}) \tag{29}$$

$$f(X^{e-nmad}) < f(X^{female\_cub}) \tag{30}$$

After conquering the position X^{male}, the pride gets updated. Whereas, once the position X^{e-nmad} is conquered, the nomad coalition updating is done. The algorithm is provoked by the territorial takeover to update the X^{male} and X^{female}, when X^{male_cub} and X^{female_cub} is matured. That is, the age of cubs that goes beyond the maximum age of cub maturity $AG_{\max}$.

Termination: The termination of this algorithm is done when one of the following criteria in Eq. (31) and (32) are met.

$$Num_{gen} > Num_{gen}^{\max} \tag{31}$$

$$\left| f(X^{male}) - f(X^{opm}) \right| \le er_{th} \tag{32}$$

where, Num_{gen} is the count of generations that is initially at zero and then maximized by one if the territorial takeover happens. Further, er_{th} and $Num_{gen}^{\max}$ are defined as the error threshold and maximum count of generations, respectively, and the absolute difference is given by $|\bullet|$. The pseudo-code of conventional LA is given in Algorithm 1.

Algorithm 1: Conventional Lion Algorithm
Step 1: Initialize X^{male}, X^{female} and X^{nmad}
Step 2: Calculate $f(X^{male})$, $f(X^{female})$ and X_1^{nmad}
Step 3: Assign $f^{reff} = f(X^{male})$ and $Num_{gen} = 0$
Step 4: Store X^{male} and $f(X^{male})$
Step 5: Perform the fertility function
Step 6: Execute mating and produce cubpool
Step 7: Compute gender clustering and gain X^{male_cub} and X^{female_cub}
Step 8: Initialize AG_{cub} as zero
Step 9: Evaluate cub growth generation
Step 10: Execute the territorial defense when defense results to zero , go to step 4
Step 11: When $AG_{cub} < AG_{\max}$, go to step 9
Step 12: Formulate territorial takeover and gain the updated X^{male} and X^{female}
Step 13: $Num_{gen} = Num_{gen} + 1$
Step 14: Go to step 4, if the termination criteria not met
Step 15: Terminate the process

5.3 Proposed Algorithm

Though the LA algorithm solves the Multi objective problem, there needs to be an enhancement or modification in the working procedure of the algorithm that must be needed to resolve all the convergence issues proclaimed by the LA algorithm. The improvement is made in the territorial defense, where there has the updating of X^{male_cub} and X^{female_cub} using the mutation process. In the proposed work, the updating of both X^{male_cub} and X^{female_cub} is performed by the female update equation given in Eq. (25). The pseudo-code of the proposed algorithm is given in Algorithm 2. The flow chart of the proposed work is given in Fig 3.

Algorithm 2: Proposed LA Algorithm
Step 1: Initialize X^{male}, X^{female} and X^{nmad}
Step 2: Calculate $f(X^{male})$, $f(X^{female})$ and X_1^{nmad}
Step 3: Assign $f^{reff} = f(X^{male})$ and $Num_{gen} = 0$
Step 4: Store X^{male} and $f(X^{male})$
Step 5: Perform the fertility function
Step 6: Execute mating and produce cubpool
Step 7: Compute gender clustering and gain X^{male_cub} and X^{female_cub}
Step 8: Initialize AG_{cub} as zero
Step 9: Evaluate cub growth generation
Step 10: Execute the territorial defense when defense results to zero, go to step 4. Here, X^{male_cub} **and** X^{female_cub} **are updated using female update as in Eq. (25).**
Step 11: When $AG_{cub} < AG_{\max}$, go to step 9
Step 12: Formulate territorial takeover and gain the updated X^{male} and X^{female}
Step 13: $Num_{gen} = Num_{gen} + 1$
Step 14: Go to step 4, if the termination criteria not met
Step 15: Terminate the process

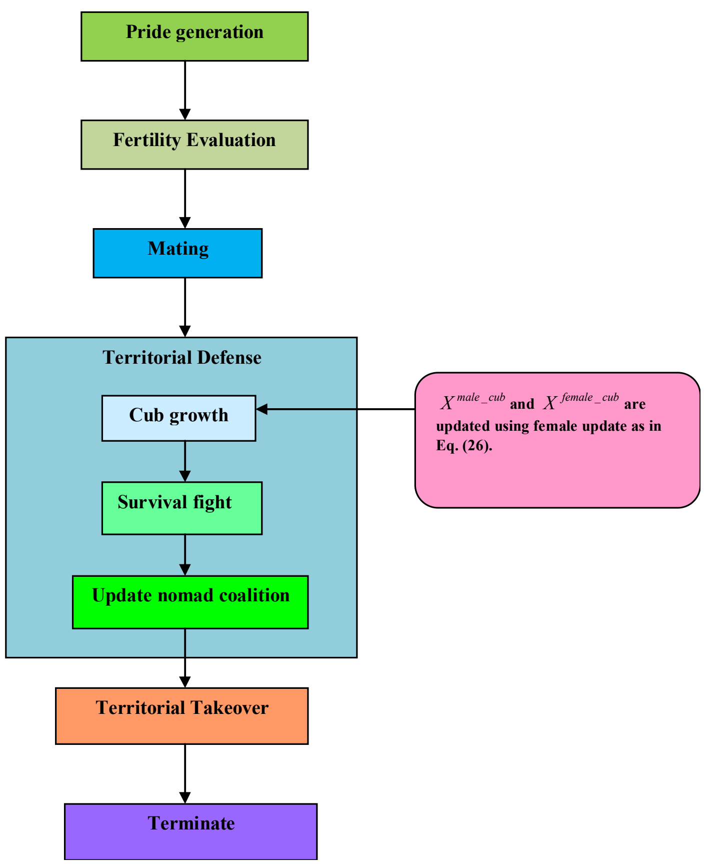


Fig. 3. *Flow chart of Proposed Model*

6. Result and Discussions

6.1 Simulation Setup

The proposed multi-object tracking model was implemented in MATLAB 2018 a. Four different videos were used from UCSD datasets [29]. The proposed model was compared with other conventional models like GM-PHD [27], SDVM, TWSM, Hybrid Model [28], and LA [26] respectively. The performance was measured under three criteria: Tracking number, Tracking distance, and MOTP. The tracking number is utilized to ensure that the proposed model is more accurate with the starting and ending object frames along with its presence on true frames. The other two measures are used for finding the divergence of 2D outcomes. Further, tracking distance is the Euclidean distance between the ground truth and tracked results. MOTP [33] is the tracking precision that indicates the correct or precise location of assessed persons.

6.2 Performance Analysis under Tracking Number Analysis

The performance Analysis under tracking number is illustrated in Fig 2. In this, 8 objects are concerned with analyzing the tracking number, where the black line indicates the accurate presence of the object in all the total frames of the given input videos. As per Fig 2 (a), every object should be present within 80 frames. However, it is observed that the LA algorithm seems to have attained the tracking number of 99, and this means all the objects are presented in 99 frames rather than 80 frames. Further, the proposed model attains the closest frames from the ground truth information, which means the proposed model shows the tracking number that obtained is closer to the ground truth. Similarly, all the graphical representations prove the betterment of the proposed model in getting closer to the ground truth information, whereas the other models attain less performance when compared to the proposed work.

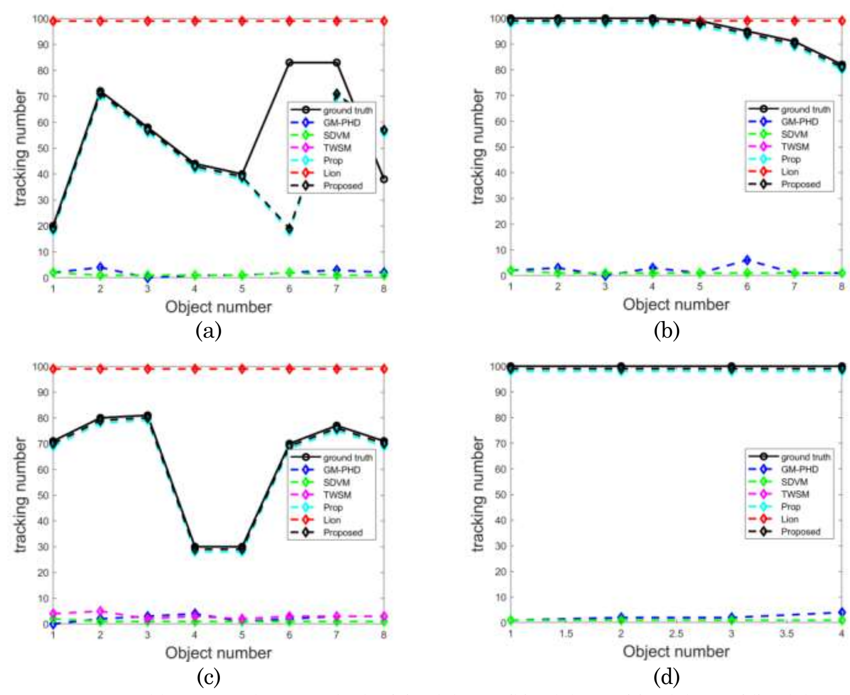


Fig. 4. *Tracking Number Analysis (a) Video 1 (b) Video 2 (c) Video 3 (d) Video 4*

6.3 Performance Analysis under Tracking Distance Analysis

Fig 3 shows the performance analysis on tracking distance. This is evaluated by identifying the distance between the original path and the object with tracked paths. The tracking distance with the least or minimum value of the respective approach is considered to be the better algorithm. The graphical representations given in Fig 3 show the performance analysis under different videos. Almost in all the cases, the proposed model seems to have better performance with minimal value. As per Fig 3 (a), the proposed model attains the least value that is in the range of 0.112 to 0.135 (approximately) for all the objects. However, the SDVM model is observed to have a tracking distance of 3.896, 8.524, and 12.342 for objects 1 to 3. All the graphs prove the betterment of the proposed work in attaining the least tracking distance when compared to others.

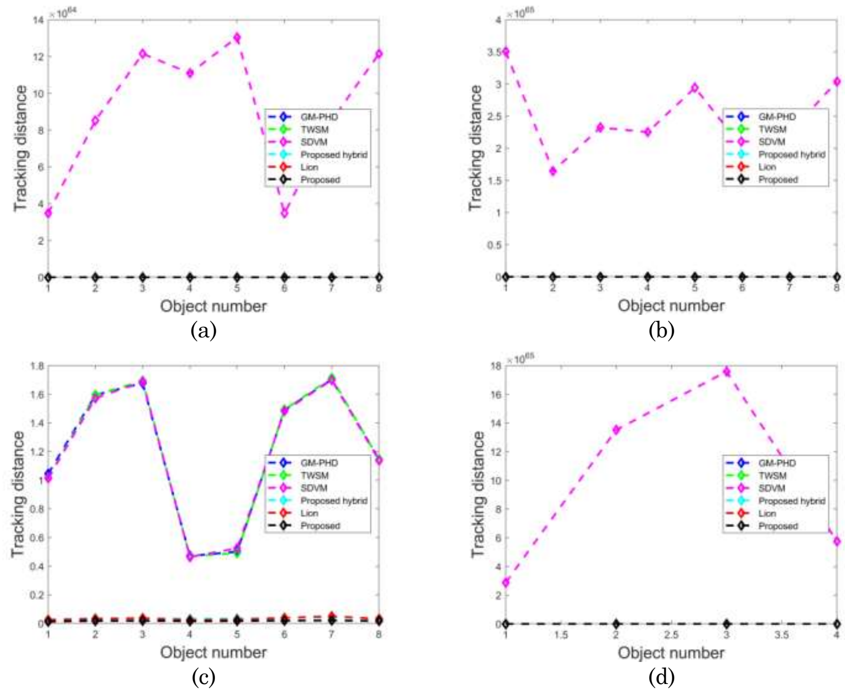


Fig. 5. *Tracking Distance Analysis (a) Video 1 (b) Video 2 (c) Video 3 (d) Video 4*

6.4 MOTP Analysis: Proposed Vs Conventional Models

Fig 4 shows the MOTP analysis of proposed over other models. From the graphs, it is observed that the proposed model attains better tracking performance on multiple objects. As per Fig 4 (a), the proposed model attains a high tracking performance of 99%, whereas the conventional SDVM model has attained performance of 81%, and conventional LA attains the lowest performance of 10%. Similarly, all the other graphical presentations reveal the performance analysis under all other videos, where it is observed that the proposed model defines a higher performance (above 90%) than others.

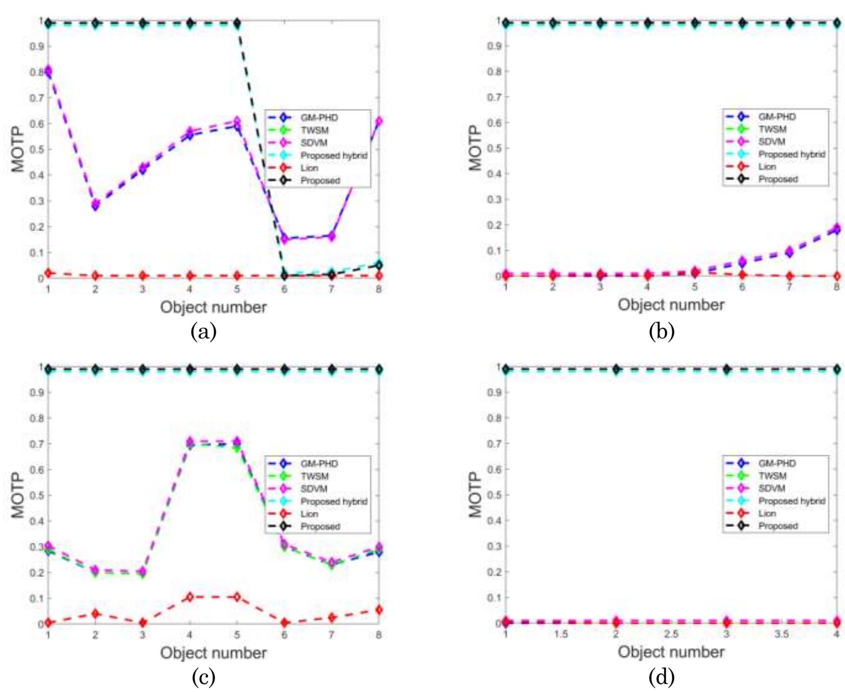


Fig. 6. *MOTP Analysis (a) Video 1 (b) Video 2 (c) Video 3 (d) Video 4*

7. Advantages and Disadvantages

Advantages

- Precise object tracking: The proposed model tracks objects precisely by considering both spatial and visual information, addressing challenges such as complex occlusion, different poses, and small object sizes.
- Improved visibility model: The model introduces a visibility model based on the second derivative function, which enhances the prediction of objects.
- Effective tracking in subsequent frames: The FU-CU-LA optimization algorithm enables effective tracking of objects even if they move to subsequent frames, improving the overall tracking performance.
- Superiority over other models: The proposed FU-CU-LA-based tracking system demonstrates superiority over other models in terms of certain measures, indicating its effectiveness in multi-object tracking.

Disadvantages

- Lack of information on specific numerical or qualitative results

8. Conclusion

This paper has proposed a novel MOT model, which has tracked the objects precisely. Initially, the visibility model of tracking was developed based on the second derivative model that has considered the second derivative function for predicting the objects. Next to this, a spatial tracking model was developed by a proposed nonlinear function. Meanwhile, the objects are tracked by subjecting a new optimization algorithm namely FU-CU-LA that could effectively track the objects even if they moved to the subsequent frames. Here, the inter-frame step size was optimally selected using the new FU-CU-LA algorithm that promises the preciseness of object tracking. Further, spatial tracking and optimization-based tracking were integrated and finally, the resultant center point was integrated with the visual tracking model. Finally, the proposed FU-CU-LA-based tracking model was compared with other existing models, and from the results, it was observed that the proposed model attains a high tracking performance of 99%, whereas the conventional SDVM model has attained the performance of 81%, and conventional LA attains the least performance of 10%. Hence the performance of the proposed model was proved over other models.

Compliance with Ethical Standards

Conflicts of interest: Authors declared that they have no conflict of interest.

Human participants: The conducted research follows the ethical standards and the authors ensured that they have not conducted any studies with human participants or animals.

References

[1] Liu, R., Huang, M., Wang, L., Bi, C. and Tao, Y., “PDT-YOLO: A Roadside Object-Detection Algorithm for Multiscale and Occluded Targets”, Sensors, Vol. 24(7), pp.2302, 2024.

[2] Jeonghwan Gwak, " Multi-object tracking through learning relational appearance features and motion patterns", Computer Vision and Image Understanding, vol. 162, pp. 103-115, September 2017.

[3] Wan Li and Cheng Wenzhi, " An online learned hough forest model based on improved multi-feature fusion matching for multi-object tracking", Multimedia Tools and Applications, pp. 1-14, 2018.

[4] Li Liang-qun, ZhanXi-yang, Liu Zong-xiang, and XieWei-xin, " Fuzzy logic approach to visual multi-object tracking", Neurocomputing, vol. 281, pp. 139-151, 2018.

[5] Liye Gui, Lisha He, Zhongyi Ni, and Tianlong Hong, " Visualized image segmentation for multi-object tracking by weak clustering technique", Multimedia Tools and Applications, pp. 1-9, 2018.

[6] Zhao, R., Zhang, X. and Zhang, J., “PSMOT: Online Occlusion-Aware Multi-Object Tracking Exploiting Position Sensitivity”, Sensors, Vol. 24(4), p.1199, 2024.

[7] Wang, Y.H., Hsieh, J.W., Chen, P.Y., Chang, M.C., So, H.H. and Li, X., “Smiletrack: Similarity learning for occlusion-aware multiple object tracking”, In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 38, No. 6, pp. 5740-5748), 2024, March.

[8] Ju Hong Yoon, Chang-Ryeol Lee, Ming-Hsuan Yang and Kuk-Jin Yoon, " Structural Constraint Data Association for Online Multi-object Tracking", International Journal of Computer Vision, pp. 1-21, 2018.

[9] S. C. Wong, V. Stamatescu, A. Gatt, D. Kearney, I. Lee, and M. D. McDonnell, "Track Everything: Limiting Prior Knowledge in Online Multi-Object Recognition," in IEEE Transactions on Image Processing, vol. 26, no. 10, pp. 4669-4683, Oct. 2017.

[10] W. Hu, W. Li, X. Zhang, and S. Maybank, "Single and Multiple Object Tracking Using a Multi-Feature Joint Sparse Representation," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 37, no. 4, pp. 816-833, 1 April 2015.

[11] A. Dehghan and M. Shah, "Binary Quadratic Programing for Online Tracking of Hundreds of People in Extremely Crowded Scenes," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 40, no. 3, pp. 568-581, 1 March 2018.

[12] S. Li, G. Battistelli, L. Chisci, W. Yi, B. Wang and L. Kong, "Computationally Efficient Multi-Agent Multi-Object Tracking With Labeled Random Finite Sets," in IEEE Transactions on Signal Processing, vol. 67, no. 1, pp. 260-275, 1 Jan.1, 2019.

[13] S. Duffner and J. Odobez, "Track Creation and Deletion Framework for Long-Term Online Multiface Tracking," in IEEE Transactions on Image Processing, vol. 22, no. 1, pp. 272-285, Jan. 2013.

[14] H. Yang, S. Qu, C. Chen and B. Yang, "Multiple Objects Tracking With Improved Sparse Representation and Rank Based Dynamic Estimation," in IEEE Access, vol. 6, pp. 42264-42278, 2018.

[15] W. Luo, B. Stenger, X. Zhao, and T. Kim, "Trajectories as Topics: Multi-Object Tracking by Topic Discovery," in IEEE Transactions on Image Processing, vol. 28, no. 1, pp. 240-252, Jan. 2019.

[16] X. Zhou, Y. Li, B. He and T. Bai, "GM-PHD-Based Multi-Target Visual Tracking Using Entropy Distribution and Game Theory," in IEEE Transactions on Industrial Informatics, vol. 10, no. 2, pp. 1064-1076, May 2014.

[17] P. Chen, Y. Dang, R. Liang, W. Zhu, and X. He, "Real-Time Object Tracking on a Drone With Multi-Inertial Sensing Data," in IEEE Transactions on Intelligent Transportation Systems, vol. 19, no. 1, pp. 131-139, Jan. 2018.

[18] Z. H. Khan, I. Y. Gu, and A. G. Backhouse, "Robust Visual Object Tracking Using Multi-Mode Anisotropic Mean Shift and Particle Filters," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 21, no. 1, pp. 74-87, Jan. 2011.

[19] Y. Lee, Z. Tang, and J. Hwang, "Online-Learning-Based Human Tracking Across Non-Overlapping Cameras," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 28, no. 10, pp. 2870-2883, Oct. 2018.

[20] K. Ahmadi and E. Salari, "Small dim object tracking using a multi objective particle swarm optimisation technique," in IET Image Processing, vol. 9, no. 9, pp. 820-826, 9 2015.

[21] W. Qu, D. Schonfeld, and M. Mohamed, "Real-Time Distributed Multi-Object Tracking Using Multiple Interactive Trackers and a Magnetic-Inertia Potential Model," in IEEE Transactions on Multimedia, vol. 9, no. 3, pp. 511-519, April 2007.

[22] Y. Yun, I. Y. Gu, and H. Aghajan, "Multi-View ML Object Tracking With Online Learning on Riemannian Manifolds by Combining Geometric Constraints," in IEEE Journal on Emerging and Selected Topics in Circuits and Systems, vol. 3, no. 2, pp. 185-197, June 2013.

[23] J. Liu, X. Cao, Y. Li, and B. Zhang, "Online Multi-Object Tracking Using Hierarchical Constraints for Complex Scenarios," in IEEE Transactions on Intelligent Transportation Systems, vol. 19, no. 1, pp. 151-161, Jan. 2018.

[24] J. Li, X. Zhou, S. Chan, and S. Chen, "Robust Object Tracking via Large Margin and Scale-Adaptive Correlation Filter," in IEEE Access, vol. 6, pp. 12642-12655, 2018.

[25] K. Granström, L. Svensson, S. Reuter, Y. Xia and M. Fatemi, "Likelihood-Based Data Association for Extended Object Tracking Using Sampling Methods," in IEEE Transactions on Intelligent Vehicles, vol. 3, no. 1, pp. 30-45, March 2018.

[26] Rajakumar Boothalingam, " Optimization using lion algorithm: a biological inspiration from lion's social behavior", Evolutionary Intelligence, vol.11, no. 1-2, pp.31–52, 2018.

[27] Xiaolong Zhou, Youfu Li, Bingwei He, and TianxiangBai, "GM-PHD-Based Multi-Target Visual Tracking Using Entropy Distribution and Game Theory", IEEE transactions on industrial informatics, Vol. 10, No. 2, pp. 1064-1076, May 2014.

[28] Felix M. Philip and Rajeswari Mukesh, " Hybrid tracking model for multiple object videos using second derivative based visibility model and tangential weighted spatial tracking model", International Journal of Computational Intelligence Systems, vol. 9, no. 5, 2016.

[29] UCSD datasets http://www.svcl.ucsd.edu/projects/anomaly/dataset.htm.

[30] Heng, W.J., Ngan, K.N., "The implementation of object-based shot boundary detection using edge tracing and tracking", Proceedings of the 1999 IEEE International Symposium on Circuits and Systems, Vol. 4, pp. 439-442, 1999.

[31] Kanungo, Tapas, Qigong Zheng, "Estimating degradation model parameters using neighborhood pattern distributions: an optimization approach", IEEE Transactions onPattern Analysis and Machine Intelligence, Vol. 26, no. 4, pp 520-524, 2004.

[32] Kratz L, Nishino K., "Tracking pedestrians using local spatio-temporal motion patterns in extremely crowded scenes", IEEE Trans Pattern Anal Mach Intell., vol. 34, no. 5, pp. 987-1002, 2012.

[33] Multiple objects tracking performance from http://www3.ntu.edu.sg/home/JSYUAN/index_files/papers/04270203_Yang_Yuan_Wu_CVPR07.pdf

[34] T. Y. Hung and K. C. Fan, "Local vector pattern in high-order derivative space for face recognition," IEEE International Conference on Image Processing (ICIP), pp. 239-243, 2014.