International Journal of Engineering Research and Technology.
ISSN 0974-3154 Volume 11, Number 5 (2018), pp. 771-789

http://www.irphouse.com

An Interactive Review of Object Motion Detection, Classification and Tracking Algorithms Related to Video Analysis in Computer Vision

Hussam Saleh Abu Karaki

School of Computer Science,
Universiti Sains Malaysia, Malaysia

Saleh Ali Alomari

Faculty of Science and Information Technology,
Jadara University, Jordan

Abstract

During the past three decades, a number of moving objects' detection, classification and tracking methods have been reported in the literature. This chapter presents a review of object motion detection, classification and tracking algorithms related to video analysis in computer vision. It also presents the studies pertaining to vehicle motion detection, classification, and tracking. Furthermore, the major strengths and weaknesses of these systems are reviewed. In general, detection, classification, and tracking are considered distinct areas of research. Nonetheless, for moving objects' tracking purposes, the detection of movement and classification of candidate objects are essential in most cases.

Keywords: Object Motion Detection, Object Tracking, Object Classification, Feature Classification

I. INTRODUCTION

The past decade has been attracted much attention on the three key tasks for image understanding, which are the object motion detection, classification, and tracking. The automat moving object detection is a critically low-level task for many video analysis and tracking applications. A largely when practically applied for a large scale aerial imagery. In [1] [2] [3] automat Moving pixels processed for various purposes, which is including urban traffic monitoring, the object classification [4], registration and tracking [5] [6] [7]. Nowadays, in the field of computer vision, the multiple objects tracking on an image plays a very important role. The object tracking has many applications in computer vision, such as video surveillance, visual interface, medicine, augmented reality man, computer management, robotics, and video compression. The fast action movements of objects will be able to change shapes directly, unstructured structures of objects, scenes and cameras have problem tracking objects. Moreover, the tracking is frequently performed in the context of the main program that needs the location and shape of the object in each frame. In [8] [9], they proposed a new robust moving object detection method used for both fixed and moving camera captured video sequences, which used by the Markov Random Field (MRF), that to get the label of the field fusion. In [10], they proposed a work using a Markov Random Field model that used to detect moving vehicles in diverse weather condition, however, the proposed work is limited, because it is only applicable to grayscale videos, to handle the spatial ambiguities of gray values. In [11], the authors proposed a new region matching used a motion estimation approach, which is used to achieve moving object detection and tracking from video captured by a moving camera. Furthermore, the fuzzy edge strength of every pixel location is included in the Markov Random Field (MRF) modeling, which preserves the object boundary for segmentation. However, the spatial segmentation problem is solved using the MAP estimation principle. To find the moving objects in subsequent frames they are a used region based estimation, where v2 test based local histogram corresponding is used to detect the moving object, in order to reduce search complexity. The computational cost of this method is too high, limiting it to specific offline applications. Furthermore, this method does not give a good result especially when shadows or occlusion are present. In [12] proposed a new method for the detection of moving objects for video captured by a moving camera. In this method, feature points are first detected using a Harris corner detector, after that the method will optical flow is used for feature matching for two consecutive frames. Moreover, these feature points are classified as belonging to background or foreground features with the support of multiple view geometry. Furthermore, the foregrounds regions are getting based on the foreground feature points, and the image variation is calculated using affine transformation depend on the background feature points. Moving object regions are gained using the merging the foreground regions and image difference. Finally, the moving objects are detected based on the history of the motion and the continuous motion contours and refinement schemes. This method is useful for real-time applications and requires no additional sensors. However, this method works very well only for the detection of slow motion moving objects. The current study, they propose a method for moving object detection and tracking from video captured through a moving camera without extra sensors. However, the authors in this paper present a new modified and extended version of the above work [12], to greatly increase the performance. The proposed method can be useful for real-time applications and works well for the detection of fast moving objects. As well as, the feature points in the frames are detected using the modified Harris corner detector proposed by [13], after that the further classified as belonging to the foreground or background features with feature matching based on optical flow and the assistance of multiple-view geometry [12].

This paper presents the significant proposed algorithms related to the current study and compares it in terms of its capabilities, limitations, advantages, and disadvantages. Section II, discusses the motion detection algorithms, Section III, explains the classification algorithms and Section IV, and presents the object tracking algorithms. Finally, at the end of the paper will provide the interesting findings of the review.

II. MOTION DETECTION

The object detection is the key technique in the field of intelligent transportation, a rich body of scholarly work exists on the subject. In this field, the targets frequently include cars or traffic signs; such targets are generally accompanied by rich prior information, which can be utilized to enhance the accuracy of target tracking (for example., car shape and group behavior are used to distinguish and predict the cars), and spatial and scale prior is used to improve the detection performance of traffic signs [14] [15].

Motion detection using optical flow refers to obtaining flow field velocities for two sequential images to indicate the differential brightness changes between the two frames. Optical flow is considered in this research to be most suitable for detecting motion in aerial videos in comparison to other detection techniques such as background subtracting based on its accuracy for estimating motion in outdoor non-fixed cameras [16]. Moreover, the Horn and Schunck in [17] method was indicated to have the ability to produce better results than Lucas and Kanade in [18] method for the purpose of analyzing aerial videos and estimating motion in their frames. For the rest of this section, a literature review of different approached founded initially on Horn and Schunck and/or Lucas and Kanade methods will attempt to establish such hypotheses through surveying different possible applications, exploitations, and enhancements on optical flow motion detection methods in general and on the Horn and Schunck method in specific. Several algorithms have been proposed in optical flow field to detect motion in different scenarios [19] [20] [21] [22] [23]. In [21] optical flow was used with indoor stationary cameras to detect moving objects such as vehicle and human from video stream depending on the observed brightness and nearby points in the image. This technique was implemented in motion detection software that can determine the motion region, count motion level, and count object numbers. However, there are limitations in highlighting specific regions of moving objects when there are changes in velocity smoothness and brightness.

For fixed cameras, a combined motion segmentation and optical flow algorithm for moving object tracking were found in [19]. Nonetheless, optical flow was only calculated on pixel level using motion segmentation to ensure that no comparisons were made between background-foreground regions. In [24], the optical flow is calculated in silhouette regions using 2-way ANOVA and object segmentation is used to minimize the effect of brightness change. In [25] proposed detecting abnormal motions in crowd monitoring scenes using Horn and Schunck method in video streams. A detection and tracking method for outdoor scenes by applying line computed by gradient-based optical flow and edge detector was proposed in [22]. The edge-based feature has robust performance and is insensitive to illumination changes. For free-moving cameras, motion clustering and classification were used to detect the moving objects in [20]. Fusion Horn and Schunck method in [19] were proposed for aerial colored images using least squares to estimate the flow field of each color plane and then fuse all distinct fields into one field. Lucas and Kanade optical flow method were used in [26] and were combined with stereo camera in [27] [28] for UAVs to navigate urban environments in a reactive way by performing a control oriented fusion. In the reviewed literature, Lucas and Kanade method was more suitable and provided effective results when the flow was presumed to be constant in local neighborhood of the pixel under consideration in [29] [30], where the method's equation is solved for all pixels in the local neighborhood by least squares criterion [27] [31]. While Horn and Schunck method was more effective for scenarios with the smooth flow over the entire frame, i.e. motion of objects is not restricted to a certain neighborhood [32] [33]. Furthermore, many researchers have attempted to apply hybrid approaches for motion detection, which uses a combination of two or more different methods including optical flow to solve the motion detection problems. Nine optical flow computation techniques were studied and evaluated in [27], and eight of the classical optical flow algorithms and were compared and their performance on complex scenes was objectively evaluated in [33]. A hybrid algorithm of the optical flow field and the temporal difference method was proposed by [34] to detect the motion object area. The method uses the temporal difference to calculate the difference between two or three consecutive frames, low pass filter and edge detection of moving objects were used to filter the differential image. However, the optical flow method implemented the Horn and Schunck algorithm which is used to compute the velocity from a spatiotemporal derivative of image intensity. Though the combination of temporal difference and optical flow in [34] was indicated to be suitable and effective for objects moving within stationary camera video, in moving cameras the combination does not produce effective results.

In this paper, the authors discussed the most of the motion detection algorithms. The simplest ones mostly use a thresholding operation on the intensity difference, for instance, the difference between consecutive video frames or the difference between the current and background frames. Depend on these basic algorithms often yield a poor performance in [35]. To develop the performance, other proposed methods employ probabilistic models in [36] [37] and statistical tests in [38] [39]. However, the statistical tests and probabilistic models are used in the model to extract the background. The performance of these detection algorithms would be largely influenced by the choice of threshold. Higher performance can hypothetically be obtained by adaptively modifying threshold value. In [35] a several threshold adaptation methods have been proposed. The most successful algorithms of detection are those which exploit frame differencing and modeling of change labels using Markov Random Field (MRF) in Bayesian framework [39]. In [40], they proposed an algorithm for the detection of moving objects, using the structure of adaptive noise annulment. The detection algorithm in the proposed work is incorporated with Bayesian Markov random field (MRF) algorithm, due to enhancing the performance in terms of the shape continuity of the detected objects. However, the main benefits of the remove the background and correlation pixels on the successive frames. What is left at the output would be an approximation of moving areas. The shape of moving objects is then improved using Bayesian algorithm [39]. The algorithm appears to be very well-organized and effective in eliminating noise, illumination variations, shadows, and repeated motions in the background.

In [41] the authors developed new tools to speed up the dramatic change of detection of low-end laptop webcam stuff with exact hardware for processing high-speed images. The 1st algorithm is fast and ensures the 2nd algorithm edge of the object found more clearly at the expense of slow image processing. The proposed method developed to enhance and reduce the waiting time of the objects more than 45.5% and with less memory usage by about 14% while keeping the same accuracy. In [100] the authors proposed an enhanced framework for detecting vehicles in aeronautical surveillance using the Dynamic Bayesian Network (DBN) and found that this method has flexibility and good generalization capabilities. In [42] the authors proposed a new visibility model depend of on the second derivative to spatial tracking model, predict the objects and tangential weighted function due to track several objects at the same time. Furthermore, the tracking of multiple objects in low-resolution videos is not possible. In [43] the author's discuss more deeply about the Speeded up robust features (SURF), which is the most significant optimization of Scale Invariant Feature Transform (SIFT). However, the computational time for SURF remains large in the actual testing process. In [15] the author's discussed the Oriented FAST and Rotated Binary Robust Independent Elementary Features (ORB) algorithm, the algorithm was below an outdoor environment for the feature extraction. Local Difference Binary (LDB) used for feature binary descriptors, and k-nearest neighbor (KNN) to match the image descriptors [44]. Furthermore, the ORB called as rBRIEF, which is based on the BRIEF algorithm by extracting local invariant features [45]. BRIEF computing is fast, other than is sensitive to noise and it has no rotational invariance. Moreover, ORB designed to solves these two major disadvantages above of the BRIEF algorithm. The ORB started to detects corners using the Harris method after that utilizes the intensity centroid to calculate the direction of rotation [46]. The authors of this paper proposed a system which can calculate the quantity and characteristics of traffic in the real-time based on three modules, segmentation, vehicle counting, and classification. The main idea of the proposed system is developing a feature based counting system for vehicle detection and recognition under the conditions, which shows a challenge in recent systems, for example, illumination conditions and occlusions [47].

III. NON PROBABILISTIC CLASSIFICATION METHODS

The majority of object classification researches for vehicle classification found in the literature related to the category of non-probabilistic techniques based on historical reasons where such methods have always provided an easier implementation with decent estimations [48] [49]. Furthermore, non-probabilistic methods do support the objectives to be achieved in this dissertation. Hence, only non-probabilistic classifiers will be described in this section and the reference classifier for the rest of this research will point only to non-probabilistic methods. The major classifiers in computer vision include Fisher's Linear Discriminant, Quadratic Discriminant Analysis, Nearest Neighbor, K-Nearest Neighbor, Support Vector Machines, Ensemble Methods, Neural Network, Decision Tree and Adaboost Classifiers [50]. At their core, these algorithms are not directly equivalent to each other where they optimize different objective functions. In general, classification in computer vision is the process of assigning corresponding levels to the examined data with regards to categories of uniform characteristics. The corresponding levels are called classes and the classification will be carried out based on spectrally defined features in the feature space and thereby segments the feature space into distinct classes founded on a decision rule technique. The classification process follows a set of standard iterations which enables the classifier to recognize the class, which an object belongs to and therefore the process is also known as part of pattern recognition methods. The first step is to clearly define the classes based on the characteristics and objectives of the frame's data. The following step is to distinctively select the features which belong to each class having that there is no overlapping between classes. The third step is determining the most fitting decision rule by sampling the training data. Thereafter, the classifier should be selected based on the training set data to select the best classification technique by comparing the techniques according to the training data, common techniques include minimum distance classifier, maximum likelihood classifier, multi-level slice classifier, fuzzy set theory and expert systems. The fifth step is to carry out the classification process by assigning pixels in the segmented area to classes depending on the decision rule. The last step is post-classification through results verification in order to maintain the classification reliability and accuracy [51]. The recognition and classification in this case, is categorized into two main methods. Namely, shape-based methods are extracted, classified, where the features of the detected objects (i.e. pixels' characteristics). Furthermore, the motion-based methods are pattern recognition, where the pattern recognition, in this case, extracts the patterns of the object's motion rather than the object itself [52]. Among popular classifiers is the artificial neural networks classifier known as the Multi-Layer Perceptron (MLP). MLP is a feed-forward model which correlates an input data to its relevant outputs to create a directed graph. The MLP graph has several layers. Each layer consists of a set of nodes and is fully linked to the next layer [53]. The Fisher's linear discriminant classifier is a derived quantity used to solve binary classification problems where the classification of any object or event is set to belong to one of two classes {-1,1}, the classes are distributed differently in the feature space through a class separation algorithm and it has a low computational complexity [54]. The Adaptive boosting (AdaBoost) is a method of enhancing the classification process.

When object instances are classified and weighted by a certain classifier; instances with the lowest weights are considered classified with low accuracy and hence as misclassified. The AdaBoost algorithm then generates a new classifier that updates the weights distribution by increasing the misclassified instances' weights and decreasing the weight of instances with high weights. This process of machine learning continues until obtaining an optimal classifier final model. Adaboost can be utilized with numerous learning algorithms to enhance classifications. Due to the AdaBoost processing mechanism; it is considered weak with outlier and noise data and also performs better with classifiers with low error rate to minimize the search for an optimal final model [55]. Nearest neighbor (1-nearest neighbor) is a non-parametric simple method that selects metric measures and uses all the training data for classification which develops a high computational complexity but may use some acceleration methods to enhance performance. The K-Nearest Neighbor, on the other hand, provides more robust classification and could provide good performance for arbitrary class-conditional probability density functions. However, it generally has high computational complexity and the K value must be set using validation methods [50]. Another classifier is based on the Support Vector Machine (SVM) which depends on the structured risk minimization principle to formulate a statistical learning model [56]. A basic SVM classifier predicts for an input data the appropriate class of two possible classes. The prediction is carried out by producing a kernelized classifier that depends sparsely on the data. SVM has been proved to be successful and efficient in several pattern recognition applications [16]. It has the advantage that its objective function is convex, whereas the objective function in relevance vector classification is non-convex and only guarantees to converge to a local minimum while in binary classification cases, SVM aims at locating a maximum margin hyperplane that best separates the instances [56] [57].

SVM does not assign certainty to its class predictions and is not so easily extended to the multi-class case. However, the multi-class classification task can be decomposed into a series of two class problems [58], wherein vehicle classification from surveillance imagery, only two classes originally exist (vehicle or non-vehicle) if all vehicle types and models were regarded only as "Vehicle". For classification certainty, due to the nature of the target objects which include a variety of models and shapes of vehicles; an exact match is hardly ever obtained and therefore certainty could never be a 100%. That is why the training set should emerge from an up-to-date local database holding as much as possible different views of probable vehicles to increase matching certainty. By employing the margin maximization theory [57] [58], training points residing on the margin between classes are exploited by SVM to maximize the margins between classes to create a sparse model that clearly classifies object instances. SVM is broadly used for the classification of objects in low-quality, grayscale and/or aerial videos due to the satisfactory results compared with its relatively low computational complexity [59] [25]. Hence, SVM will be utilized as the classification method for the current study. The following section provides reviews of SVM applications in similar studies.

In recent classification literature, SVM is regarded as a robust, accurate and effective pattern recognition and classification technique [58]. SVM has been significantly utilized in aerial motion detection and tracking systems [59] [60]. In [59] Dynamic Bayesian Networks are combined with SVM to detect and classify vehicles in colored aerial videos. Local-features are extracted based on the Canny edge detection and Harris corner detection methods and color-based features are extracted based on color model separating vehicle colors from non-vehicle colors by transforming the (RGB) color components into a vector. In [60] utilizes the HOG feature to classify vehicles in linear SVM method combined with an AdaBoost method for aerial images due to its insensitivity to illumination change and scene complexity. The approach boosts the HOG feature by reducing the feature vector dimensionality in order to decrease the computational complexity. In [25] employ a multi-class SVM classifier with compressed sensor measurements as features where each frame is expressed as a measurement vector and its dimensionality is reduced (compressed) for better performance. A hyperplane is used in [61] to classify objects in aerial images via SVM to separate shape-based features (geometric characteristics) through linear and non-linear equations to map the feature vectors to a higher dimensional Euclidean space. In general, SVM is capable of using several object features to classify an object [62] such as, edge-based features (gradient matching) [63], Local shape-based features (Eigenvalues) [52], Region-based features (intensity) [31] Texture-based features (local binary pattern). Suitable feature selection is an important task for generalization performance optimization, minimizing running time requirements, reducing the error rate and increasing confidence level and addressing the constraints and interpretational issues imposed by the recognition purpose itself [57] [58]. Each image could produce different features depending on its properties (color or grayscale) and depending on the object type required to be classified (e.g. vehicles, humans, etc.). Edge features represent image points with sharp brightness values and resemble a discontinuity measure separating an object area from the surrounding area. Major edge detector utilizes the Canny, Canny-Deriche, Differential, Sobel, Prewitt or Roberts-Cross methods for detecting this discontinuity [64]. Corner detection methods identify the intersection between two edges (two objects' boundaries) based on the local neighborhood of the point. Corner detection methods include the Harris operator, Shi and Tomasi, Level curve curvature, Smallest Univalve Segment Assimilating Nucleus (SUSAN) and the First corner based on Accelerated Segment Test (F-AST) methods [100]. Hough transform identifies positions of arbitrary shapes in an image like an elliptical. Finally, blob detection targets regions differing in their properties, such as brightness or color, from other areas surrounding those regions. Major blob detection methods are Laplacian of Gaussian (LoG), Difference of Gaussians (DoG), Determinant of Hessian (DoH), Maximally Stable External Regions (MSER) and Principal Curvature-Based Region (PCBR) methods. For any of the given features, the feature description is the format representation of the feature and the feature vector. The common feature descriptors in computer vision are Scale-invariant feature transform (SIFT) [65], Speeded Up Robust Features (SURF), Gradient Location and Orientation Histogram (GLOH) and Histogram of Oriented Gradients (HOG) [66]. The extracted features of an object may be used for classification as well as tracking purposes. The following section depicts the review of tracking methods in computer vision [101].

After locating the moving object, the next step is to identify the object. To do this first the feature extraction techniques are applied on to the localized object. Speeded Up Robust Features (SURF) is a local feature detector that used for tasks, such as, registration, object recognition, 3D reconstruction, and classification. It is partly inspired by the Scale Invariant Feature Transform (SIFT) descriptor. The typical version of Speeded-up Robust Features is numerous times faster than Scale Invariant Feature Transform and claimed by its authors to be more robust against diverse image transformations than SIFT. The SURF is used to locate and recognize objects, faces, to make 3D scenes, as well as to track objects and to extract points of interest [67]. The Histogram of Oriented Gradients (HOG) is a feature descriptor used in image processing and computer vision designed for the purpose of object detection, which is used to calculate the occurrences of gradient orientation in localized portions of the original image [68] [69]. Enhanced Local Vector Pattern (ELVP) is a novel vector representation developed to represent the structure information of local texture and 1D direction and the adjacent pixels with diverse distances from diverse directions. Depend on the vector representation, the Local Vector Pattern(LVP) is proposed to present different 2D spatial structures of micropatterns with different pairwise directions of the vector of its neighborhoods and the reference pixel.

Genetic Algorithm (GA) is used for finding optimal features from the ELVP, SURF, and HOG. The optimal solution can be developed using a population of strings, which is called the genotype of the genome, and by encoding candidate solutions, which is called phenotypes, due to an optimization problem. The optimal solutions are represented in binary as strings of 0s and 1s. The evolution regularly begins from a population of randomly generated individuals and happens in generations. In every generation, the fitness of every individual in the population is evaluated. Several individuals are selected from the current population, which is based on their fitness, and modified recombined and possibly randomly mutated, which is to form a new population. The new population is used in the next new iteration of the algorithm. The algorithm terminates when either a maximum number of generations has been produced or can say the satisfactory fitness level has been reached for the maximum population. Moreover, if the algorithm has terminated due to the greatest number of generations, a satisfactory solution may not have been achieving the optimal solution. Genetic Algorithm (GA) proceeds to initialize a population of solutions randomly after that improves it through repetitive application of the mutation, crossover, and inversion and selection operators [69] [70].

IV. OBJECT TRACKING

The human tracking is one of the main topics in intelligent video surveillance systems. By tracking human in the videos, it is possible to collect their trajectories for high-level analytics and applications, for example, human counting, people flow estimation, criminal tracking, and so on. The object tracking in computer vision can be defined as depicting the movement path for a moving object in a frame sequence [71]. Tracking follows the methods of detecting and recognizing moving objects in order to be aware of the tracked object [72]. Once a moving object is detected, tracking holds the responsibility of identifying the objects' path in the subsequent frames through path alignment or prediction techniques or by simply indicating its location and direction of movement in each slide. Advanced tracking techniques require ensuring that each object is correlated properly with the same object in the following frames and therefore calls for identifying each object by a set of characteristics for making certain tracking. Once an object is detected and classified as a vehicle, the following phase includes tracking this object. Tracking is the process of estimating the motion parameters and locations of the object starting from the first frame (initialization position) and for the subsequent frames [73]. A tracking method typically consists of three major components, namely Object Representation, Dynamic Model and Search Mechanism. Object's appearance is affected by several factors with respect to the environment model i.e. the dataset characteristics. Object representation determines the most suitable function to search for a target object in a frame e.g. representing the object as a region of connected pixels in the frame. Moreover, advanced adaptive representation schemes could be used based on generative or discriminative formulations. Since the representation could differ slightly between frames; a Dynamic Model should be provided by predefining the model or by allowing the system to learn it from a training data. The purpose of the Dynamic Model is to reduce the search space and computational load in each processed frame by predicting possible target states for the object [74]. Object tracking for a mobile robot traveling in crowded urban environments, building on the previously proposed Deep Tracking framework [75] [76].

Several algorithms were found in the literature to propose advances in object tracking. Target objects can be represented at fixed views by learning a subspace model offline [77] [78] [79]. Online Expectation Maximization (EM) with a Gaussian mixture model is proposed to handle target appearance variations during tracking [80]. Fragment-based appearance model [81] and template-based method [82] were proposed to overcome image noise, pose change and partial occlusion problems [81] [83]. From another point of view, discriminative methods handle object tracking as a binary classification problem by recognizing target regions as separated regions from the background. Support vector machine classifiers are enhanced with optical flow through Gaussian Pyramid to create a Support Vector Tracking (SVT) [84]. Color features [85] online boosting and multiple-instance learning (MIL) [86] enhance the tracking through its rich information. Furthermore, the underlying structure in unlabeled sampled data could be exploited to select positive and negative samples for update [87]. Moreover, several criterions, parameters, and features were found in use in the object tracking literature (e.g. success rate and center location error) for performance evaluation [30] [88] [89]. A major limitation in the reviewed literature is the use of limited sequences of frames [90], and therefore, there are not any clear distinctions between which algorithms are most suitable for certain applications. However, it depends mainly on understanding the behavior and restrictions of the target objects like humans or vehicle as well as the camera in terms of location and capabilities like cameras mounted on aerial vehicles [91].

In [92] the authors proposed most significant algorithms, which is related to the Object Tracking, and those algorithm allowing a robust and accurate detection of moving objects for a small-cost in memory expenditure and computational complexity, that allow the user to use normal cameras and computers instead of expensive and high-performance computers. Although the set faces are used to delays especially in vision process box and connections (LAN). The authors in this paper discussed the existing target tracking algorithms, which can be roughly divided into three categories [93]. Firstly, the Point tracking detects a target in consecutive frames using a point representation, the points are associated based on the earlier state of the target object, which can include both object motion and object position. A Kalman particle filter is a representative type of point tracking method, which is used to estimate the state of a linear or non-linear system and update it to achieve target tracking. As well as, kalman particle filter normally defined as a comparison between a target and its template [94] [95]. Secondly, the kernel-based tracking is achieved by computing the motion of a kernel across consecutive frames, which mean shift method is a typical kernel tracking method that maximizes the appearance similarity of a largest in successive frames iteratively by comparing the object features and features in a window around the next hypothesized object location, which can be considered as the kernel [96]. Finally, the silhouette based object tracker is used to find the object region in each frame by an object model produced from the earlier frames. The shape matching is an ordinary approach used in silhouette-based object tracking methods. Shape-matching is presented by computing the similarity of an object with a model generated from the hypothesized silhouette of an object based on previous frames [97].

In [68] the authors proposed a Deformable Multiple Kernel (DMK) tracking algorithms takes, which efficiently combine the Deformable Part Model (DPM) into the multiple kernel tracking. On the other hand, the Multiple Kernels search the for the local optimal based on color and deformable part model (DPM) information, and kernels are bound with each other, owing to the deformation costs. However, the advantage of the proposed work of not only low computation owing to the kernel-based tracking, but also robustness of the deformable part model (DPM) detector, as a result as to successfully track objects more accurately. In [98] the authors proposed an approach to perform the object tracking for a mobile robot traveling in crowded urban environments, building on the previously proposed Deep Tracking framework [75] [76]. However, unlike classical techniques, which employ a multi-stage pipeline, the method is learned end to end with limited architectural choices. Furthermore, through the employing a spatial transformer module, the model is able to exploit noisy estimates of visual ego-motion as a proxy for true vehicle motion. The results show that the method achieves favorably to deep tracking in terms of accurately predicting future states, and demonstrate that the model can capture the location and motion of cars, buses, cyclists, and pedestrians, even when incomplete occlusion. In [99] the authors proposed a deep learning based approach for robust outdoor vehicle tracking. The proposed approach firstly, a stacked denoising autoencoder is pre-trained to learn the feature representation way of images. After that, the k-sparse constraint is added to the stacked denoising auto-encoder and the encoder of k-sparse Stacked Denoising Auto Encoder (kSSDAE). The kSSDAE is attached with a classification layer to build a classification Neural Network (NN). Moreover, after fine-tuning, the classification NN is applied to online tracking under particle filter framework. The wide tracking experiments are conducted on a challenging single object online tracking evaluation platform benchmark to verify the effectiveness of our tracker. The results show that the tracker outperforms most state-of-the-art trackers.

V. CONCLUSIONS AND DISCUSSIONS

To this end, this paper focuses on reviewing recent algorithms in the video analysis field Object Motion Detection, Classification and Tracking Algorithms Related to Video Analysis in Computer Vision [101]. Through the review, various algorithms and approaches have been discussed. Each of the techniques has its advantages and disadvantages with respect to the application objectives [85]. In optical flow motion detection techniques, several approaches, which utilized Horn-Schunck or Lucas-Kanade methods, were reviewed in order to justify the precedence of Horn-Schunck for vehicle detection in aerial videos. The advantages of Horn and Schunck method over Lucas and Kanade were depicted to be based on the scenarios having that Horn and Schunck method is more effective for scenarios with smooth flow over the entire frame (global constraints) i.e. motion of objects are not restricted to a certain neighborhood [32] [33], while Lucas and Kanade is more efficient for scenarios with local constraints [29] [30]. The classification of non-probabilistic methods was exemplified and literature for classifiers in each technique was reviewed in terms of multi-class handling, computational complexity and possible features to be included for classification were explicated [30]. Moreover, the significance of SVM classifiers and different applications and features were indicated in the literature based on choosing SVM as the most suitable classifier for the research at hand, because SVM employs the margin maximization theory to create a sparse model that clearly classifies object instances and having that basic limitation of SVM were found possible to be overcome for the target application. The literature also reviews related work in tracking moving objects from aerial surveillance using a variety of methods and thereby indicated that the tracking method could be designed based on the features used in the detection and classification methods like the centroid tracking which depends on detecting the object region in the frame.

REFERENCES

[1] S. C. Sen-Ching and C. Kamath, (2004),"Robust techniques for background subtraction in urban traffic video," in Proc. SPIE Visual Communications and Image Processing, vol. 5308, 2004, pp. 881–892.

[2] R. Feris, R. Bobbitt, S. Pankanti, and M.-T. Sun,(2015) "Efficient 24/7 object detection in surveillance videos," in IEEE AVSS, 2015, pp. 1–6.

[3] Kannappan Palaniappan,(2016)."Moving Object Detection for Vehicle Tracking in Wide Area Motion Imagery Using 4D Filtering". 2016 23rd International Conference on Pattern Recognition (ICPR) Cancún Center, Cancún, México, December 4-8, 2016.

[4] S. Sivaraman and M. M. Trivedi, (2014)., "Active learning for on-road vehicle detection: A comparative study," Machine Vision and Applications, vol. 25,

no. 3, pp. 599–611, 2014.
[5] F. Schubert and K. Mikolajczyk, (2014)., “Robust registration and filtering for moving object detection in aerial videos,” in IEEE ICPR, 2014, pp. 2808–2813.
[6] W. R. Thissell, R. Czajkowski, F. Schrenk, T. Selway, A. J. Ries, S. Patel, P. L. McDermott, R. Moten, R. Rudnicki, G. Seetharaman, I. Ersoy, and K. Palaniappan, (2015) “A scalable architecture for operational FMV exploitation (AVAA),” IEEE ICCV Video Summarization for Large-scale Analytics.
[7] R. Gargees, B. Morago, R. Pelapur, D. Chemodanov, P. Calyam, Z. Oraibi, Y. Duan, G. Seetharaman, and K. Palaniappan, (2016), “Incidentsupporting visual cloud computing utilizing software-defined networking,” IEEE Trans. on Circuits and Systems for Video Technology, 2016.
[8] P.M. Jodoin, M. Mignotte, C. Rosenberger (2007), Segmentation framework based on label field fusion, IEEE Trans. Image Process. 16 (10) 2535–2550.
[9] Wu-Chih Hua, Chao-Ho Chen b, Tsong-Yi Chen b, Deng-Yuan Huang c, Zong-Che Wud, (2015)., Moving object detection and tracking from video captured by moving camera, J. Vis. Commun. Image, Elsevier 164 180.
[10] Y. Wang, (2010). Joint random field model for all-weather moving vehicle detection, IEEE Trans. Image Process. 19 (9) 2491–2501.
[11] A. Ghosh, B.N. Subudhi, S. Ghosh, (2012)., Object detection from videos captured by moving camera by fuzzy edge incorporated Markov random field and local histogram matching, IEEE Trans. Circuits Syst. Video Technol. 22 (8) 1127–1135.
[12] W.-C. Hu, C.-H. Chen, C.-M. Chen, T.-Y. Chen,(2014)., Effective moving object detection from videos captured by a moving camera, in: Proceedings of the First Euro China Conference on Intelligent Data Analysis and Applications, vol. 1, 2014, pp. 343–353.
[13] J. Shi, C. Tomasi,(2016), Good features to track, in: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 1994, pp. 593–600.
[14] Y. Yuan, Z. Xiong, and Q. Wang, (2017)., “An incremental framework for video based traffic sign detection, tracking, and recognition,” IEEE Trans. Intell. Transp. Syst., vol. 18, no. 7, pp. 1918–1929, Jul. 2017.
[15] Peng Chen, Yuanjie Dang, Ronghua Liang., (2018), Real-Time Object Tracking on a Drone With Multi-Inertial Sensing Data, IEEE Transactions On Intelligent Transportation Systems, VOL. 19, NO. 1, JANUARY 2018.
[16] Ramakrishnan, V., Prabhavathy, A. K. and Devishree, J. (2012). A Survey on Vehicle Detection Techniques in Aerial Surveillance, International Journal of Computer Applications 55(18).
[17] Horn, B. and Schunck, B. (1981). Determining Optical Flow, Artificial intelligence 17(1): 185– 203.
[18] Lucas, B. and Kanade, T. (1981). An Iterative Image Registration Technique with an Application to Stereo Vision, Proceedings of the 7th international joint conference on Artificial intelligence.
[19] Denman, S., Fookes, C. and Sridharan, S. (2009). Improved Simultaneous Computation of Motion Detection and Optical Flow for Object Tracking,

Digital Image Computing: Techniques and Applications, DICTA 2009., IEEE, pp. 175–182

[20] Kim, J., Ye, G. and Kim, D. (2010). Moving Object Detection Under Free-Moving Camera, Image Processing (ICIP), 2010 17th IEEE International Conference on, IEEE, pp. 4669–4672.

[21] Shafie, A., Hafiz, F. and Ali, M. (2009). Motion Detection Techniques Using Optical Flow, World Academy of Science, Engineering and Technology 56.

[22] Yokoyama, M. and Poggio, T. (2005). A Contour-Based Moving Object Detection and Tracking, Visual Surveillance and Performance Evaluation of Tracking and Surveillance, 2005. 2nd Joint IEEE International Workshop on, IEEE, pp. 271–276.

[23] Zhang, P., Cao, T. and Zhu, T. (2010). A Novel Hybrid Motion Detection Algorithm Based on Dynamic Thresholding Segmentation, Communication Technology (ICCT), 2010 12th IEEE International Conference on, IEEE, pp. 853–856.

[24] Girisha, R. and Murali, S. (2011). Tracking Humans using Novel Optical Flow Algorithm for Surveillance Videos, Proceedings of the Fourth Annual ACM Bangalore Conference, ACM, pp: 7.

[25] Tzagkarakis, G., Charalampidis, P., Tsagkatakis, G., Starck, J.-L. and Tsakalides, P. (2012). Compressive Video Classification for Decision Systems with Limited Resources, Picture Coding Symposium (PCS), 2012, IEEE, pp. 353–356.

[26] Braillon, C., Pradalier, C., Crowley, J. L. and Laugier, C. (2006). Real-Time Moving Obstacle Detection using Optical Flow Models, Intelligent Vehicles Symposium, IEEE, pp. 466–471.

[27] Hrabar, S. and Sukhatme, G. S. (2004). A Comparison of Two Camera Configurations for Optic-Flow Based Navigation of a UAV Through Urban Canyons, Intelligent Robots, and Systems, 2004. (IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on, Vol. 3, IEEE, pp. 2673–2680.

[28] Hrabar, S., Sukhatme, G. S., Corke, P., Usher, K. and Roberts, J. (2005). Combined Optic-Flow and Stereo-Based Navigation of Urban Canyons for a UAV, Intelligent Robots and Systems, 2005. (IROS 2005). 2005 IEEE/RSJ International Conference on, IEEE, pp. 3309–3316.

[29] Xiao, J., Cheng, H., Feng, H. and Yang, C. (2008). Object Tracking and Classification in Aerial Videos, Proceedings of SPIE, the International Society for Optical Engineering, Society of Photo-Optical Instrumentation Engineers, pp. 696711–1.

[30] Basavaiah, M. (2012).Development of Optical Flow Based Moving Object Detection and Tracking System on an Embedded DSP Processor, Journal of Advances in Computational Research: An International Journal 1(1-2).

[31] Hu, W., Tan, T., Wang, L. and Maybank, S. (2004). A Survey on Visual Surveillance of Object Motion and Behaviors, Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on 34(3): 334–352.

[32] Baker, S., Scharstein, D., Lewis, J., Roth, S., Black, M. and Szeliski, R. (2011). A Database and Evaluation Methodology for Optical Flow,

International Journal of Computer Vision 92(1): 1–31.
[33] Galvin, B., McCane, B., Novins, K., Mason, D. and Mills, S. (1998). Recovering Motion Fields: An Evaluation of Eight Optical Flow Algorithms, British Machine Vision Conference, Vol. 1, sn, pp. 195–204.
[34] Shuigen, W., Zhen, C. and Hua, D. (2009). Motion Detection Based on Temporal Difference Method and Optical Flow Field, Electronic Commerce and Security, 2009. ISECS'09. Second International Symposium on, Vol. 2, IEEE, pp. 85–88.
[35] JM McHugh, J Konrad, V Saligrama, P Jodoin (2009), Foreground-adaptive background subtraction. IEEE Signal. Process Lett. 16, 390–393 .
[36] CR Wren, A Azarbayejani, T Darrell, A Pentland, Pfinder: real-time tracking of the human body. IEEE Trans. Pattern Anal. Mach. Intell. 18(7), 780–785 (1997).
[37] A Elgammal, R Duraiswami, D Harwood, L Davis, Background and foreground modeling using nonparametric kernel density for visual surveillance. Proc. IEEE 90(7), 1151–1163 (2002).
[38] T Aach, A Kaup, (1995), Bayesian algorithms for adaptive change detection in image sequences using Markov random fields. Signal Process Image Comm. Vol. 7, No. (2), 147– 160.
[39] T Aach, A Kaup, R Mester, (1993)., Statistical model-based change detection in moving video. Signal Process. 31, 165– 180.
[40] Elham Kermani and Davud Asemani. (2014)., A robust adaptive algorithm of moving object detection for video surveillance, EURASIP Journal on Image and Video Processing 2014, 2014:27, pp:2-9.
[41] Roy, A.; Shinde,S.; Kang, K.-D. (2012); An Approach for Efficient Real Time Moving Object Detection, International Journal of Signal Processing, Image Processing and Pattern Recognition, 5(3), 2012.
[42] Philip, F.M.; Mukesh R.(2016); 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, 9(5): 888-899, 2016.
[43] H. Bay, T. Tuytelaars, L. Van Gool, "SURF: Speeded up robust features," in Proc. Eur. Conf. Comput. Vis. (ECCV), 2006, pp. 404–417.
[44] X. Yang and K.-T. Cheng, (2014).,"Local difference binary for ultrafast and distinctive feature description," IEEE Trans. Pattern Anal. Mach. Intell., vol. 36, no. 1, pp. 188–194, Jan. 2014
[45] M. Calonder, V. Lepetit, C. Strecha, and P. Fua, "BRIEF: Binary robust independent elementary features," in Proc. Eur. Conf. Comput. Vis. (ECCV), 2010, pp. 778–792.
[46] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," in Proc. Int. Conf. Comput. Vis. (ICCV), Barcelona, Spain, Nov. 2011, pp. 2564–2571.
[47] Zakaria Moutakki, Imad Mohamed Ouloul, Karim Afde, Abdellah Amghar, (2018)., "Real-Time System Based On Feature Extraction For Vehicle Detection And Classification", Transport and Telecommunication, 2018,

volume 19, no. 2, 93–102 .

[48] Lai, J., Huang, S. and Tseng, C. (2010)., Image-Based Vehicle Tracking and Classification on the Highway, Green Circuits and Systems (ICGCS), 2010 International Conference on, IEEE, pp. 666–670.

[49] Liu, X., Dai, B. and He, H. (2011). Real-Time On-Road Vehicle Detection Combining Specific Shadow Segmentation and SVM Classification, Digital Manufacturing and Automation (ICDMA), 2011 Second International Conference on, IEEE, pp. 885–888.

[50] Sonka, M., Hlavac, V. and Boyle, R. (1999). Image Processing, Analysis, and Machine Vision, PWS Pub.

[51] Tso, B. and Mather, P. (2009). Classification Methods for Remotely Sensed Data, CRC.

[52] Niu, F. and Abdel-Mottaleb, M. (2004). View-invariant Human Activity Recognition Based on Shape and Motion Features, Multimedia Software Engineering, 2004. Proceedings. IEEE Sixth International Symposium on, IEEE, pp. 546–556.

[53] Gallo, I. and Nodari, A. (2011). Learning Object Detection Using Multiple Neural Netwoks, Proceedings of International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications. INSTICC Press.

[54] Witten, D. M. and Tibshirani, R. (2011). Penalized Classification using Fisher's Linear Dis- criminant, Journal of the Royal Statistical Society: Series B (Statistical Methodology) 73(5): 753–772.

[55] Park, J., Choi, H. and Oh, S. (2010). Real-Time Vehicle Detection in Urban Traffic Using AdaBoost, Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on, IEEE, pp. 3598–3603.

[56] Han, F., Shan, Y., Cekander, R., Sawhney, H. and Kumar, R. (2006). A Two-Stage Approach to People and Vehicle Detection with Hog-Based SVM, Performance Metrics for Intelligent Systems Workshop in conjunction with the IEEE Safety, Security, and Rescue Robotics Conference, pp. 133–140.

[57] Chen, H., Tino, P. and Yao, X. (2009). Probabilistic Classification Vector Machines, Neural Networks, IEEE Transactions on 20(6): 901–914.

[58] Chen, Z., Pears, N., Freeman, M. and Austin, J. (2009). Road vehicle classification using support vector machines, Intelligent Computing and Intelligent Systems, 2009. ICIS 2009. IEEE International Conference on, Vol. 4, IEEE, pp. 214–218.

[59] Asha, G., Kumar, K. A. and Kumar, D. D. N. P. (2012). A Real Time Video Object Tracking Using SVM, International Journal of Engineering Science and Innovative Technology (IJESIT).

[60] Cao, X., Wu, C., Yan, P. and Li, X. (2011). Linear SVM Classification using Boosting HOG Features for Vehicle Detection in Low-Altitude Airborne Videos, Image Processing (ICIP), 2011 18th IEEE International Conference on, IEEE, pp. 2421–2424.

[61] Song, M. and Civco, D. (2004). Road extraction using svm and image segmentation, Photogrammetric engineering and remote sensing 70(12):

1365–1371.

[62] Palaniappan, K., Bunyak, F., Kumar, P., Ersoy, I., Jaeger, S., Ganguli, K., Haridas, A., Fraser, J., Rao, R. M. and Seetharaman, G. (2010)., Efficient Feature Extraction and Likelihood Fusion for Vehicle Tracking in Low Frame Rate Airborne Video, Information Fusion (FUSION), 2010 13th Conference on, IEEE, pp. 1–8.

[63] Weston, J., Mukherjee, S., Chapelle, O., Pontil, M., Poggio, T. and Vapnik, V. (2001). Feature Selection for SVMs, Advances in neural information processing systems pp. 668–674.

[64] Gonsales, R., Woods, R. and Eddins, S. (2004). Digital Image Processing Using MATLAB, Pearson Education Inc., publishing as Prentice Hall.

[65] Goncalves, H., Corte-Real, L. and Goncalves, J. (2011). Automatic Image Registration Through Image Segmentation and SIFT, Geoscience and Remote Sensing, IEEE Transactionson 49(7): 2589–2600.

[66] Dalal, N. and Triggs, B. (2005). Histograms of Oriented Gradients for Human Detection, Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, Vol. 1, IEEE, pp. 886–893.

[67] Zohrevand, A.; Ahmadyfard, A.; Pouyan, A.; Imani, Z. (2014); A SIFT based object recognition using contextual information, Iranian Conference on Intelligent Systems (ICIS), 1-4, 2014.

[68] Li, Y.; Su G. (2015)., Simplified histograms of oriented gradient features extraction algorithm for the hardware implementation, International Conference on Computers, Communications and Systems (ICCCS), 192 -195, 2015.

[69] G. Jemilda, S. Baulkani, (2018)., Moving Object Detection and Tracking using Genetic Algorithm Enabled Extreme Learning Machine, International Journal Of Computers Communications & Control Issn 1841-9836, 13(2), 162-174, April 2018.

[70] Shingade, A.; Ghotkar A.(2014); Survey of Object Tracking and Feature Extraction Using Genetic Algorithm, International Journal of Computer Science and Technology, 5(1), 2014.

[71] Gu, Y., Li, P. and Han, B. (2009). Embedding Ensemble Tracking in a Stochastic Framework for Robust Object Tracking, Journal of Zhejiang University-Science A 10(10): 1476–1482.

[72] Yilmaz, A., Javed, O. and Shah, M. (2006). Object Tracking: A Survey, ACM Computing Surveys (CSUR), Vol 38, No. 4. pp: 13.

[73] Dedeoglu, Y. (2004). Moving Object Detection, Tracking and Classification for Smart Video Surveillance, Master’s thesis, Bilkent University.

[74] Zhao, Y., Gong, H., Jia, Y. and Zhu, S. (2012). Background Modeling by Subspace Learning on Spatio-Temporal Patches, Pattern Recognition Letters 33(9): 1134–1147.

[75] P. Ondr´uˇska and I. Posner, (2016)., “Deep tracking: Seeing beyond seeing using recurrent neural networks,” in The Thirtieth AAAI Conference on Artificial Intelligence (AAAI), Phoenix, Arizona USA, February 2016.

[76] P. Ondr´uˇska, J. Dequaire, D. Z. Wang, and I. Posner, (2016)., “End-to-end

tracking and semantic segmentation using recurrent neural networks," arXiv preprint arXiv:1604.05091, 2016.

[77] YZ Hsu, HH Nagel, G Refers, (2018)., "New likelihood test method for change detection in image sequences". Compute Vision Graph Image Process. 26, 73–106 .

[78] Ross, D., Lim, J., Lin, R. and Yang, M. (2008). Incremental Learning for Robust Visual Tracking, International Journal of Computer Vision 77(1): 125–141.

[79] Black, M. and Jepson, A. (1996). Eigentracking: Robust Matching and Tracking of Articulated Objects Using a View-Based Representation, Computer Vision ECCV'96 pp. 329–342.

[80] Jepson, A., Fleet, D. and El-Maraghi, T. (2003). Robust Online Appearance Models for Visual Tracking, Pattern Analysis and Machine Intelligence, IEEE Transactions on 25(10): 1296–1311.

[81] Adam, A., Rivlin, E. and Shimshoni, I. (2006). Robust Fragments-Based Tracking Using the Integral Histogram, Computer Vision and Pattern Recognition, IEEE Computer Society Conference on, Vol. 1, IEEE, pp. 798–805.

[82] Mei, X. and Ling, H. (2009). Robust Visual Tracking Using L1 Minimization, Computer Vision, 2009 IEEE 12th International Conference on, IEEE, pp. 1436–1443.

[83] Kwon, J. and Lee, K. (2010). Visual Tracking Decomposition, Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on, IEEE, pp. 1269–1276.

[84] Avidan, S. (2004). Support Vector Tracking, Pattern Analysis and Machine Intelligence, IEEE Transactions on 26(8): 1064–1072.

[85] Teuliere, C., Eck, L. and Marchand, E. (2011). Chasing a Moving Target from a Flying UAV, Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on, IEEE, pp. 4929–4934.

[86] Babenko, B., Yang, M. and Belongie, S. (2011). Robust Object Tracking with Online Multiple Instance Learning, Pattern Analysis and Machine Intelligence, IEEE Transactions on 33(8): 1619–1632.

[87] Kalal, Z., Matas, J. and Mikolajczyk, K. (2010). P-N Learning: Bootstrapping Binary Classifiers by Structural Constraints, Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on, IEEE, pp.49–56.

[88] Xiao, J., Cheng, H., Sawhney, H. and Han, F. (2010). Vehicle Detection and Tracking in Wide Field-of-View Aerial Video, Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on, IEEE, pp. 679–684.

[89] Meng, L. and Kerekes, J. (2012). Object Tracking Using High Resolution Satellite Imagery, Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of 5(1): 146–152.

[90] Woo, H., Jung, Y., Kim, J. and Seo, J. (2010). Environmentally Robust Motion Detection for Video Surveillance, Image Processing, IEEE Transactions on 19(11): 2838–2848.

[91] Vishwakarma, S. and Agrawal, A. (2012). A Survey on Activity Recognition

and Behavior Understanding in Video Surveillance, The Visual Computer pp. 1–27.
[92] Moteaal Asadi Shirzi and M.R. Hairi-Yazdi, (2011)., Active Tracking Using Intelligent Fuzzy Controller And Kernel-Based Algorithm, 2011 IEEE International Conference on Fuzzy Systems June 27-30, 2011, Taipei, Taiwan.
[93] Yu Hang, Chen Derong, Gong Jiulu, (2018)., "Object tracking using both a kernel and a non-parametric active contour model, Neurocomputing 295 (2018) 108–117.
[94] A. Yilmaz , O. Javed , M. Shah , (2006)., Object tracking: a survey, ACM Comput. Surv. (CSUR) 38 (4) 13 .
[95] M.S. Arulampalam , S. Maskell , N. Gordon , T. Clapp , A tutorial on particle filters for online nonlinear/non Gaussian Bayesian tracking, IEEE Trans. Signal Pro- cess. 50 (2) (2002) 174–188 .
[96] D. Comaniciu , V. Ramesh , P. Meer , Kernel-based object tracking, IEEE Trans. Pattern Anal. Mach. Intell. 25 (5) (2003) 564–577 .
[97] D.P. Huttenlocher , J.J. Noh , W.J. Rucklidge , Tracking non-rigid objects in com- plex scenes, in: Proceedings of the Fourth International Conference on Com- puter Vision, 1993, pp. 93–101 .
[98] Julie Dequaire, Dushyant Rao, Peter Ondr´uˇska, Dominic Zeng Wang and Ingmar Posner, (2017)., " Deep Tracking on the Move: Learning to Track the World from a Moving Vehicle using Recurrent Neural Networks" 2017
[99] Jing Xin, Xing Du, Jian Zhang (2017), Deep Learning For Robust Outdoor Vehicle Visual Tracking, Proceedings of the IEEE International Conference on Multimedia and Expo (ICME) 2017.
[100] Cheng, H.-Y.; Weng, C.-C.; Chen Y.-Y.(2012); Vehicle Detection in Aerial Surveillance Using Dynamic Bayesian Networks, IEEE Transactions on Image Processing, 21(4): 2152- 2159, 2012.
[101] Szeliski, R. (2010). Computer Vision: Algorithms and Applications, Springer.