International Journal of Control and Automation
Vol. 8, No. 9 (2015), pp. 303-312
http://dx.doi.org/10.14257/ijca.2015.8.9.29

A Survey of Video Object Tracking

Meng Li, Zemin Cai[1], Chuliang Wei and Ye Yuan

Department of Electronic Engineering, College of Engineering, Shantou University, China
Guangdong Provicial Key Laboratory of Digital Signal and Image Processing Techniques, China
[1]Corresponding author, zmcai@stu.edu.cn

Abstract

Due to the numerous important applications of video surveillance and monitoring, video object tracking has been an active research topic in the last decade. This paper makes a survey of approaches to high quality object tracking by looking at theoretical backgrounds and practical results, which are categorized into four groups. The principle, the evolution processes and the latest progresses of these approaches are identified to form a conclusion for future directions of object tracking algorithms.

Keywords*: video object tracking; feature matching; Bayesian estimation; classifier; real-time*

1. Introduction

Video object tracking [1] is an important task within the field of computer vision. As an interdisciplinary frontier technology, it combined with image processing, pattern recognition, artificial intelligence, automatic control and other areas of theory and knowledge. Video object tracking has broad application prospect in many fields [2-5]: video surveillance, human-computer interaction, intelligent traffic, robot vision navigation, precision guided weapons, *etc*. The research of tracking algorithms is of important theoretical value and practical significance.

Video object tracking refers to the detection, extraction, recognition and tracking of moving object in video image sequences, in order to obtain accurate motion information parameters (such as position, velocity, *etc*), and carries on the analysis to the corresponding processing, so we can further implement object behavior understanding. Video object tracking can be a very complicated task due to: complex object shapes, irregular movements, scene illumination changes, object occlusion and real-time requirements.

In recent years, with the rise of research on digital image technology and computer vision, video object tracking technology rapid development, many excellent algorithms have been proposed. This survey categorizes and describes the existing tracking algorithms and explains their strengths and weaknesses at the end of each category. Finally, some future directions of object tracking algorithms are also addressed shortly.

2. Video Object Tracking Algorithms

Since the 1980s, numerous algorithms for video object tracking have been proposed. According to the different methodologies for object tracking algorithms, we categorize the existing tracking methods into four forms, including Matching-based tracking, Filtering-based tracking, Class-based tracking and Fusion-based tracking. Then we will introduce each type of the tracking algorithms respectively.

ISSN: 2005-4297 IJCA

2.1. Matching-Based Tracking

Matching-based tracking algorithms are established for object model based on matching before tracking, depending on the matching relation to select the best matching point as the tracking result in the current video sequences. With the different descriptions of the object attributes, we can divide the Matching-based tracking algorithms into the following four basic methods:

2.1.1. Region-Based Tracking

The fundamental thought of region-based tracking algorithms [6] is: the initial object region of the image as the object template, matching the object template with all the possible location of the candidate images, the highest matching degree of the position is judged to be the best match point, and the region identified by the point is the object region. The most common of correlation matching criterion is Sum of Square Difference (SSD).

The object templates are fixed in the early region-based tracking algorithms. Lucas et al. [7] present a method by using spatial gradient of gray image to find the best matching region, which uses gray gradient values of each point in the object region to update candidate object region. This method can be applied to affine motion or non-rigid object tracking. It is feasible to use a fixed template in a short-period tracking, but it is malfunction in a long-period tracking due to the object appearance changes. If the object templates can change along with the object appearances, the reliability of the tracking algorithms would be improved. Jepson *et al.*, [8] propose an adaptive appearance model based on texture features, which provides robustness in occlusions. And an online EM-algorithm is used to update the object model during tracking. Recently, Zoidi *et al.*, [9] employ background subtraction to detect object， and update object model by similarity of color histogram. The experiments results show that the proposed tracking scheme is successful in tracking objects under scale and rotation variations and partial occlusion.

Mean Shift algorithm [10] is a typical representative of region-based tracking algorithms. The core idea is: building the object model by manual or target detection algorithm, establishing a cost function to describe the similarity between the object model and the object candidate model, using optimization methods to find the maximum cost function, and the maximum value corresponds to the location of the object in video sequences. It is a classical object tracking algorithm due to it is easy to implement and it has the advantage of good real-time performance. Since Mean Shift algorithm is applied by Comaniciu *et al.*, [11, 12] to the field of computer vision, a large number of improved algorithms [13, 14] based on Mean Shift are proposed.

Region-based tracking algorithms use global informations of the objects, such as gray informations, texture features and so on. So they have high credibility, even if minor deformation of the objects does not affect the tracking performance. Unfortunately, they are time-consuming when the search regions are large. And the decline of the accuracy would cause the loss of the objects when the objects have large deformations or serious occlusions. In recent years, region-based tracking algorithms focus on how to deal with the variations of the templates. The variations are caused by various poses of the moving objects. If we can predict the various poses, steady tracking can be realized. The papers [15, 16] study this issue and achieve good tracking effects.

2.1.2. Feature-Based Tracking

The fundamental thought of feature-based tracking algorithms [17] is: Using the feature to identify the object, and match the object in the video sequences based on image feature. This algorithm usually has two main steps which include feature extraction and feature match. The first step is to extract salient features from the video sequences, such as corner, boundary, and centroid, *etc.* The second step is to find the most similar object in the next frame according to the feature matching criterion. The method can determine the position of the object in the video sequences with the two steps.

Early papers solve the problem of identifying the corresponding feature points between the adjacent frames. The common method is to exert some constraint conditions in the motion of the feature points. Sethi *et al.*, [18] assume that the motion of the feature points is smooth, and propose a concept of path coherence and the corresponding iterative algorithm. According to the similarity hypothesis, Rangarajan *et al.*, [19] present a non iterative polynomial time approximation algorithm based on optical flow to solve the correspondence problem. But these algorithms have high computational complexity and terrible error in the scene of occlusions. Tissainayagam *et al.*, [20] propose an object tracking algorithm based on point feature. The method searches the local maximum corner points, and defines them as the key points which are tracked by MHT approach. It has good performance for simple geometry objects, but has poor performance due to the difficulty of extracting stable corner points for complex objects. Recently, Li *et al.*, [21] explore a new feature subset evaluation method for feature selection in object tracking. The work gives a novel idea for how to select the useful features, and improves the real-time performance of the algorithm.

Feature-based tracking algorithms don’t use global features of the objects. Although the objects are partial occlusions, they can use other visible features to accomplish the tracking task. But the algorithms can’t deal with full occlusions and overlaps effectively. In recent years, numerous papers have sprung up in the field of feature-based tracking algorithms. See some influential papers on [22, 23].

2.1.3. Deformable Template-Based Tracking

The fundamental thought of deformable template-based tracking algorithms [24] is: Using the surface or curve which has good elasticity and deformation property of the contour or edge of the moving object as object bounding contour, and updating this contour to match the object.

Kass *et al.*, [25] propose Active Contour Models which are the most commonly used deformable templates in object tracking. Snakes are Active Contour Models which lock onto nearby edges, localizing them accurately. They define that a Snake is an energy-minimizing function guided by external constraint forces and influenced by image forces. Numerous improved algorithms based on Snake have been proposed. I-Snake [26], B-Snake [27], G-Snake [28], and AI-Snake [29] are the representative ones. Applications show that Snakes are suit for tracking rigid and non-rigid motions [30, 31], and only for single object tracking. Osher *et al.*, [32] propose an Active Contour Model based on level sets apply to multiple object tracking.

In the situations of noise interferences, occlusions, and edge blur, it is very difficult to get accurate contour by the above models. Importing the priori knowledge of color [33], texture [34], and shape [35] et al., and use them to constrain the contours to obtain exact objects edges. Yilmaz et al. [33] propose a contour tracking algorithm which uses kernel density estimation and Gabor wavelet model to guide the contour evolution based on the color feature and texture feature respectively. Ning *et al.*, [36] present a novel framework by joint registration and active contour segmentation (JRACS), which has good robustness for non rigid object tracking.

By contrast, deformable template-based tracking algorithms describe objects simply, reduce computational complexity and have high robustness in the situations of partial occlusions. However, the algorithms are difficult to start tracking automatically, because they are highly sensitive to the initialization of tracking. A further weakness is that the tracking precision is limited at the contour level. In recent years, the state of the art tracking algorithms based on deformable template are proposed by the papers [37, 38].

2.1.4. Model-Based Tracking

The fundamental thought of model-based tracking algorithms [39] is: establishing geometric model of object according to the priori knowledge, tracking object by matching the candidate region model and object model, then in the tracking process, so it can determine the scale parameters, the pose parameters, and the motion parameters by the image characteristics. The models can be divided into: hierarchical model [40], 2D models [41] and 3D models [42]. The 3D models are the most widely used tracking algorithms in practical applications. The 3D models: obtaining the 3D structure model and object model according to the prior knowledge, collecting the video sequences to determine the parameters of 3D model, in order to obtain the object motion parameters. The 3D model-based tracking algorithms can precisely analyze the 3D motion trajectories of objects. Even in the situations of different types of poses, they can reliably track the objects. However, a mass of parameters are required to build the 3D models, the models matching processes are quite complex, and the 3D tracking systems are time-consuming and have poor real-time performances.

Yang *et al.*, [43] present a novel 3-D model-based vehicle localization algorithm, which can efficiently and robustly determine the poses of vehicles in traffic scenes by calibrated cameras. The method uses edge points in the images as features, and measures the degree of matching between the edge points and the projected model by a pose evaluation function. Chen *et al.*, [44] propose a dynamic spatial bias appearance model (DSBAM) to track an object by partitioning the object into regions with different confidences. Yang et al. [45] introduce an online conditional random field (CRF) model, and transform the tracking problem into an energy minimization problem. The energy functions include a set of unary functions that are on account of motions and appearance models. The method is powerful to distinguish spatially close objects with similar appearances.

Model-based tracking algorithms are not easily affected by observation perspectives, so they are intrinsically robust to various motions. The algorithms have high precision by using the information of 3D. Ineluctably, model-based tracking algorithms have some insufficiencies such as the necessity of constructing the models, high computational cost, complex model updating mechanisms, and poor real-time performances. In the recent years, the papers [46, 47] are focus on how to improve the real-time performances.

2.2. Filtering-Based Tracking

Filtering-based tracking algorithms regard the tracking problems as state estimation problems [48]. The states of the objects can include all motions characteristics of our concern objects. The key of the objects tracking is how to infer the posterior probability density of the objects state and effectively represent them under the condition of giving the observation data. In order to deduce the posterior probability density functions of the states, two objects state models are introduced as follows:

State model: $$s_k = f_k(s_{k-1}, w_k) \quad (1)$$

Observation model: $$z_k = h(s_k, v_k) \quad (2)$$

hereinto, $\{s_k, k \in N\}$ and $\{z_k, k \in N\}$ represent object state and observation data, respectively. State model is used for the description of the evolution of the system, and observation model is used for the description of the relationship between observation and state.

Posterior probability density function includes all the statistical information about the object state which can be obtained in the tracking process, so it is a complete solution of estimation problem. Maximum value estimation of object state, minimum variance estimation, maximum a posteriori estimation, all of them can be obtained by the posterior probability density function. Kalman Filter and Particle Filter are two typical Bayesian filtering methods which would be discussed in the following sections.

2.2.1. Kalman Filter

Kalman Filter (KF) [49] is an effective method for estimating the states, which predicts the object states by state model, and estimates the posterior probability density function by observation model. When the dynamic characteristics of the object states (state model and observation model) are satisfy with both Gauss and linearity conditions, Kalman Filter can obtain the optimal solution with the significance of the minimum mean square error. The thought of Kalman Filter [50] is widely used since it is introduced in the field of video object tracking. In video object tracking, Kalman Filter usually use Gauss noise to represent the uncertainty of state model and observation model, and use the uncertainty to automatic balance the effect of observations and predictions to tracking results. If the video sequences interferences are weak, it is reasonable to approximate the uncertainty of the object states by Gauss model, Kalman Filter can obtain good tracking results. However, if the video sequences interferences are strong, Kalman Filter is invalid due to multiple peaks contained by the posterior probability density function. In order to solve the tracking problem of non-linear systems, Extended Kalman Filter (EKF) algorithms [51] are presented. EKF provides a linearized method for the non-linear systems, and then tracking by KF. EKF has the advantage of simple calculation, but when the systems have a high degree of non-linearity, EKF may be invalid. Unscented Kalman Filter (UKF) [52] is another approach for non-linear systems. It introduces Monte Carlo Filter which uses a set of discrete sample points to describe the mean and variance of the posterior probability density function of the state. It can acquire good filtering effects in the non-linear systems on account of avoiding the direct processing of the linearization of the non-linear systems.

2.2.2. Particle Filter

Particle Filter (PF) [53] is a sequential Monte Carlo Filter, which is used to solve the Bayesian estimation problem under the condition of non-linear and non-Gauss. The fundamental thought of PF is: evaluating the probability density function of the state by a set of weighted sampling particles, using Monte Carlo method to simulate the spread of the probability density, when the sampling quantity is big enough, the sampling can be a good approximation to probability density function, and the mean, covariance and other statistics can be easily calculated according to the sampling points. Because of the no constrains of the linear or Gauss, PF is widely used in the practical applications: radar target tracking, machine learning, computer vision, wireless communication, speech enhancement, robots, *etc*.

Isard and Blake are the first to introduce PF to be applied to visual tracking, and propose the famous Conditional Density Propagation algorithm [54] which obtains good tracking effects. Different kinds of improved PF algorithms are greatly developed. In order to get better Importance Density, numerous improved algorithms based on PF have been proposed, such as Auxiliary Particle Filter (APF), Extended Kalman Particle Filter (EKPF), and Unscented Particle Filter (UPF), *etc*. Markov Chain Monte Carlo (MCMC) algorithm [55] is introduced in PF to solve the problem of Sample Impoverishment. With the development of the theory of density estimation, Regularized Particle Filter (RPF) is also presented which uses continuous density to resample.

A large number of studies show that tracking in complex environment, Particle Filter has better performance than Kalman Filter. However, PF struggles with the problems of particle degeneracy phenomenon, tracking accuracy and occlusions. In recent years, Particle Filter tracking problems [56, 57] are mainly to discuss the three problems.

2.3. Class-Based Tracking

In recent years, some scholars [58-60] regard the tracking problem as a classification problem of foreground and background. They precisely locate the objects positions by constructing classifiers to classify the location areas.

Avidan [61] combines optical flow method and Support Vector Machine (SVM), and uses them for vehicle tracking. But it is difficult to form and train the positive and negative samples. In addition, he also proposes a boosting-based method [62] for object tracking, and the method produces satisfactory tracking results. Kalal *et al.*, [63] develop a novel learning method (P-N learning) to estimate the tracker's errors and updates it to avoid these errors in the next frame. The method shows a significant improvement over state-of–the-art approaches during in long-term tracking.

Class-based tracking algorithms have high tracking accuracy. Unfortunately, there are two drawbacks: on the one hand, constructing the classifiers needs vast positive and negative samples that become the difficulty of how to study and select the samples; on the other hand, the need to search for objects in a large range area. In recent years, the papers [64, 65] are concern about how to structure reasonable classifiers.

2.4. Fusion-Based Tracking

In practical applications, fusion-based tracking algorithms are proposed to achieve good tracking effects. The algorithms often combine variety of tracking algorithms or different sources of information to improve the accuracy of the tracking results. This method makes full use of complementary advantages of different methods to obtain high-quality tracking.

Fusion-based tracking algorithms can be divided into the following three basic parts:

The first method is based on multi-features fusion, which is the most common approach. It can not be stable tracking in a long time by using single feature information, due to the complex of the tracking scenarios or objects. Therefore, many scholars use multi-feature information fusion technology to improve the video object tracking performance. They give different weights by the ability of features to describe the moving objects. The method advances the tracking accuracy by the reliability of objects description. Zhou *et al.*, [66] present method which integrates spatial position, shape and color information to improve the performance of EKF. Brasnett *et al.*, [67] develop a consistent histogram-based framework for the analysis of color, edge and texture features. The method shows that tracking with multiple weighted features provides more reliable performance than single feature tracking.

The second method is based on multi-model fusion. This method integrates object models in different time of video sequences or combines object models with various angels of multiple cameras. It improves the robustness of tracking, on account of adapting the changes of the object and characterizing the effective features of the object. Khan *et al.*, [68] present a multiview approach to solve the tracking problems in crowed and cluttered scenes.

The third method is based on multi-algorithm fusion. Different object tracking algorithms have their own advantages for certain scenes. The method makes full use of the strengths of various algorithms by integrating the appropriate algorithms, so it can overcome the weaknesses of individual algorithm. Shan *et al.*, [69] proposes a Mean Shift embedded Particle Filter (MSEPF) algorithm which improves the sampling efficiency considerably by incorporating the Mean Shift into PF to move particles to local peaks in the likelihood.

The association of more objects cues or methods, the final fusion algorithms may have more accuracy results. But regretfully, the objects also have more constraints, the scenes also become more special, the computing time also is longer. Therefore, the fusion algorithms are often use eclectic principle to deal with the tracking problems. Recently, the papers [70, 71] focus on how to improve the tracking accuracy and satisfy the efficiency and applicability at the same time.

3. Conclusion

Nowadays, numerous video object tracking algorithms are proposed, but most of them are only suit for specific objects or fixed scenes owing to the application constrains. Due to the diversity of tracking objects, the complexity of scenarios and the variety of applications, there are enormous challenges for the current video object tracking algorithms. To implement a robust, accurate and real-time visual tracking system is the direction to the current tracking algorithms. It is very difficult to embrace every challenge of the tracking technologies, in that the development of the technologies contact with the research of people's perception. We all realize that there is not a mainstream theory for the people's perception, and its mathematical model is difficult to establish. However, video object tracking technologies have a wide application prospects. With the development of mathematical methods and computer technologies, it is increasingly convenience to implement whether in theories or in applications. Video object tracking technologies will have a profound impact on people's future lives.

Acknowledgements

This research was supported by the National Science Foundation of China (under Grant 61203249), Foundation for Distinguished Young Talents in Higher Education of Guangdong, China(Grant No. LYM11066, LYM11067) and Engineering and Technology Research Center of Guangdong Higher Education Institutes(GCEX-A1306).

References

[1] A. Yilmaz, O. Javed and M. Shah, "Object tracking: A survey", ACM Computing Surveys, vol.38, **(2006)**.

[2] J. J. Deng, X. H. Xie and B. Daubney, "A bag of words approach to subject specific 3D human pose interaction classification with random decision forests", Graphical Models, vol.76, **(2014)**, pp.162-171.

[3] L. Calderoni, D. Maio and S. Rovis, "Deploying a network of smart cameras for traffic monitoring on a 'city kernel", Expert Systems with Applications, vol.41, **(2014)**, pp.502-507.

[4] D. Lockery, J. F. Peters and S. Ramanna, "Store-and-feedforward adaptive gaming system for hand-finger motion tracking in telerehabilitation", Information Technology in Biomedicine, IEEE Transactions, vol.15, **(2011)**, pp.467-473.

[5] C. W. Tang, "Spatiotemporal visual considerations for video coding", Multimedia, IEEE Transactions, vol.9, **(2007)**, pp.231-238.

[6] G. D. Hager and P. N. Belhumeur, "Efficient region tracking with parametric models of geometry and illumination", Pattern Analysis and Machine Intelligence, IEEE Transactions, vol.20, **(1998)**, pp.1025-1039.

[7] B. D. Lucas and T. Kanade, ed., "An iterative image registration technique with an application to stereo vision", IJCAI, vol.81, **(1981)**, pp.674-679.

[8] A. D. Jepson, D. J. Fleet and T. F. El-Maraghi, "Robust online appearance models for visual tracking", Pattern Analysis and Machine Intelligence, IEEE Transactions, vol.25, **(2003)**, pp.1296-1311.

[9] O. Zoidi, A. Tefas and I. Pitas, "Visual object tracking based on local steering kernels and color histograms", Circuits and Systems for Video Technology, IEEE Transactions, vol.23, **(2013)**, pp.870-882.

[10] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis", Pattern Analysis and Machine Intelligence, IEEE Transactions, vol.24, **(2002)**, pp.603-619.

[11] D. Comaniciu, V. Ramesh and P. Meer, eds., "Real-time tracking of non-rigid objects using mean shift. Computer Vision and Pattern Recognition", Proceedings. IEEE Conference, **(2000)**.

[12] D. Comaniciu, V. Ramesh and P. Meer, "Kernel-based object tracking", Pattern Analysis and Machine Intelligence, IEEE Transactions, vol.25, **(2003)**, pp.564-577.

[13] O. D. Nouar, G. Ali and C. Raphael, eds., "Improved object tracking with CamShift algorithm", Acoustics, Speech and Signal Processing, ICASSP Proceedings, IEEE International Conference, **(2006)**.

[14] I. Leichter, "Mean shift trackers with cross-bin metrics", Pattern Analysis and Machine Intelligence, IEEE Transactions, vol.34, **(2012)**, pp.695-706.

[15] K. J. Cannons and R. P. Wildes, "The applicability of spatiotemporal oriented energy features to region tracking", Pattern Analysis and Machine Intelligence, IEEE Transactions, vol.36, **(2014)**, pp.784-796.

[16] R. He, B. Yang and N. Sang, “Integral region-based covariance tracking with occlusion detection”, Multimedia Tools and Applications, **(2014)**, pp.1-22.
[17] S. N. Sinha, J. M. Frahm, M. Pollefeys and Y. Genc, “Feature tracking and matching in video using programmable graphics hardware”, Machine Vision and Applications, vol.22, **(2011)**, pp.207-217.
[18] I. K. Sethi and R. Jain, “Finding trajectories of feature points in a monocular image sequence”, Pattern Analysis and Machine Intelligence, IEEE Transactions, vol.1, **(1987)**, pp.56-73.
[19] K. Rangarajan and M. Shah, “Establishing motion correspondence”, CVGIP: image understanding, vol.54, **(1991)**, pp.56-73.
[20] P. Tissainayagam and D. Suter, “Object tracking in image sequences using point features”, Pattern Recognition, vol.38, **(2005)**, pp.105-113.
[21] G. R. Li, Q. M. Huang, J. B. Pang, S. Q. Jiang and L. Qin, “Online selection of the best< i> k</i>-feature subset for object tracking”, Journal of Visual Communication and Image Representation, vol.23, **(2012)**, pp.254-263.
[22] E. J. Ong and R. Bowden, “Robust facial feature tracking using shape-constrained multiresolution-selected linear predictors”, Pattern Analysis and Machine Intelligence, IEEE Transactions, vol.33, **(2011)**, pp.1844-1859.
[23] X. Yang and K. T. Cheng, “Local Difference Binary for Ultrafast and Distinctive Feature Description”, Pattern Analysis and Machine Intelligence, IEEE Transactions, vol.36, **(2014)**, pp.188-194.
[24] Y. Zhong, A. K. Jain and M. P. Dubuisson-Jolly, “Object tracking using deformable templates”, Pattern Analysis and Machine Intelligence, IEEE Transactions, vol.22, **(2000)**, pp.544-549.
[25] M. Kass, A. Witkin and D. Terzopoulos, “Snakes: Active contour models”, International Journal of Computer Vision, vol.1, **(1988)**, pp.321-331.
[26] R. Ronfard, “Region-based strategies for active contour models”, International Journal of Computer Vision, vol.13, **(1994)**, pp.229-251.
[27] P. Brigger, J. Hoeg and M. Unser, “B-spline snakes: a flexible tool for parametric contour detection”, Image Processing, IEEE Transactions, vol.9, **(2000)**, pp.1484-1496.
[28] N. Paragios and R. Deriche, “Geodesic active regions and level set methods for supervised texture segmentation”, International Journal of Computer Vision, vol.46, **(2002)**, pp.223-247.
[29] Z. Xue, S. Z. Li and E. K. Teoh, “AI-EigenSnake: an affine-invariant deformable contour model for object matching”, Image and Vision Computing, vol.20, **(2002)**, pp.77-84.
[30] S. Derrode, M. A. Charmi and F. Ghorbel, eds., “Fourier-based invariant shape prior for snakes”, Acoustics, Speech and Signal Processing, ICASSP Proceedings. IEEE International Conference, **(2006)**.
[31] H. Fang, S. H. Kim and J. W. Jang, eds., “A snake algorithm for automatically tracking multiple objects. Image Processing (ICIP), 18th IEEE International Conference, **(2011)**.
[32] S. Osher and J. A. Sethian, “Fronts propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations”, Journal of Computational Physics, vol.79, **(1988)**, pp.12-49.
[33] A. Yilmaz, X. Li and M. Shah, eds., “Object contour tracking using level sets”, Asian Conference on Computer Vision, **(2004)**.
[34] N. Houhou, J. Thiran and X. Bresson, eds., “Fast texture segmentation model based on the shape operator and active contour”, Computer Vision and Pattern Recognition, CVPR, IEEE Conference, **(2008)**.
[35] G. Charpiat, O. Faugeras and R. Keriven, eds., “Shape statistics for image segmentation with prior”, Computer Vision and Pattern Recognition, CVPR, IEEE Conference, **(2007)**.
[36] J. F. Ning, L. Zhang, D. Zhang and W. Yu, “Joint Registration and Active Contour Segmentation for Object Tracking”, Circuits and Systems for Video Technology, IEEE Transactions, vol.23, **(2013)**, pp.1589-1597.
[37] A. Nakhmani and A. Tannenbaum, “Self-Crossing Detection and Location for Parametric Active Contours”, Image Processing, IEEE Transactions, vol.21, **(2012)**, pp.3150-3156.
[38] W. M. Hu, X. Zhou, W. Li, W. H. Luo, X. Q. Zhang and S. Maybank, “Active Contour-Based Visual Tracking by Integrating Colors, Shapes, and Motions”, Image Processing, IEEE Transactions, vol.22, **(2013)**, pp.1778-1792.
[39] H. Dahlkamp, H. H. Nagel, A. Ottlik and P. Reuter, “A framework for model-based tracking experiments in image sequences”, International Journal of Computer Vision, vol.73, **(2007)**, pp.139-157.
[40] I. A. Karaulova, P. M. Hall and A. D. Marshall, “Tracking people in three dimensions using a hierarchical model of dynamics”, Image and Vision Computing, vol.20, **(2002)**, pp.691-700.
[41] S. Wu, L. Hong and J. R. Layne, “2D rigid-body target modelling for tracking and identification with GMTI/HRR measurements”, Control Theory and Applications, IEE Proceedings, vol.151, **(2004)**, pp.429-438.
[42] D. M. Gavrila and L. S. Davis, eds., “3-D model-based tracking of humans in action: A multi-view approach”, Computer Vision and Pattern Recognition, Proceedings CVPR, IEEE Computer Society Conference, **(1996)**.
[43] H. Yang, J. Lou, H. Z. Sun, W. M. Hu and T. N. Tan, eds., “Efficient and robust vehicle localization”, Image Processing, Proceedings International Conference, **(2001)**.
[44] D. T. Chen and J. Yang, “Robust Object Tracking via Online Dynamic Spatial Bias Appearance Models”, Pattern Analysis and Machine Intelligence, IEEE Transactions, vol.29, **(2007)**, pp.2157-2169.

[45] B. Yang and R. Nevatia, "Multi-Target Tracking by Online Learning a CRF Model of Appearance and Motion Patterns", International Journal of Computer Vision, vol.107, **(2014)**, pp.203-217.
[46] X. Li, A. Dick, C. H. Shen, A. van den Hengel and H. Z. Wang, "Incremental Learning of 3D-DCT Compact Representations for Robust Visual Tracking", Pattern Analysis and Machine Intelligence, IEEE Transactions, vol.35, **(2013)**, pp.863-881.
[47] C. Wojek, S. Walk, S. Roth, K. Schindler and B. Schiele, "Monocular Visual Scene Understanding: Understanding Multi-Object Traffic Scenes", Pattern Analysis and Machine Intelligence, IEEE Transactions, vol.35, **(2013)**, pp.882-897.
[48] O. Williams, A. Blake and R. Cipolla, "Sparse bayesian learning for efficient visual tracking", Pattern Analysis and Machine Intelligence, IEEE Transactions, vol.27, **(2005)**, pp.1292-1304.
[49] R. E. Kalman, "A new approach to linear filtering and prediction problems", Journal of Basic Engineering, vol.82, **(1960)**, pp.35-45.
[50] R. J. Meinhold and N. D. Singpurwalla, "Understanding the Kalman filter", The American Statistician, vol.37, **(1983)**, pp.123-127.
[51] A. E.Nordsjo, "A constrained extended Kalman filter for target tracking", Radar Conference, Proceedings of the IEEE, **(2004)**.
[52] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation", Proceedings of the IEEE, **(2004)**.
[53] N. J. Gordon, D. J. Salmond and A. F. M. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation", IEE Proceedings F (Radar and Signal Processing), IET Digital Library, **(1993)**.
[54] M. Isard and A. Blake, "Condensation—conditional density propagation for visual tracking", International Journal of Computer Vision, vol.29, **(1998)**, pp.5-28.
[55] K. Choo and D. J. Fleet, "People tracking using hybrid Monte Carlo filtering", Computer Vision, ICCV, Proceedings Eighth IEEE International Conference, **(2001)**.
[56] W. Choi, C. Pantofaru and S. Savarese, "A general framework for tracking multiple people from a moving camera", Pattern Analysis and Machine Intelligence, IEEE Transactions, vol.35, **(2013)**, pp.1577-1591.
[57] P. Kumar and A. Dick, "Adaptive earth movers distance-based Bayesian multi-target tracking", Computer Vision, IET, vol.7, **(2013)**, pp.246-257.
[58] M. Andriluka, S. Roth and B. Schiele, "People-tracking-by-detection and people-detection-by-tracking", Proc. of IEEE Conference on Computer Vision and Pattern Recognition, **(2008)**.
[59] D. A. Ross, J. Lim, R. S. Lin and M. H. Yang, "Incremental learning for robust visual tracking", International Journal of Computer Vision, vol.77, **(2008)**, pp.125-141.
[60] B. Babenko, M. H. Yang and S. Belongie, "Robust object tracking with online multiple instance learning", Pattern Analysis and Machine Intelligence, IEEE Transactions, vol.33, **(2011)**, pp.1619-1632.
[61] S. Avidan, "Support vector tracking", Pattern Analysis and Machine Intelligence, IEEE Transactions, vol.26, **(2004)**, pp.1064-1072.
[62] S. Avidan, "Ensemble tracking", Pattern Analysis and Machine Intelligence, IEEE Transactions, vol.29, **(2007)**, pp.261-271.
[63] Z. Kalal, K. Mikolajczyk and J. Matas, "Tracking-Learning-Detection", Pattern Analysis and Machine Intelligence, IEEE Transactions, vol.34, **(2012)**, pp.1409-1422.
[64] G. Shu, A. Dehghan, O. Oreifej and E. Hand, "Part-based multiple-person tracking with partial occlusion handling", Computer Vision and Pattern Recognition , IEEE Conference, **(2012)**.
[65] L. Zhang and L. van der Maaten, "Preserving Structure in Model-Free Tracking", Pattern Analysis and Machine Intelligence, IEEE Transactions, vol.36, **(2014)**, pp.756-769.
[66] Q. M. Zhou and J. K. Aggarwal, "Object tracking in an outdoor environment using fusion of features and cameras", Image and Vision Computing, vol.24, **(2006)**, pp.1244-1255.
[67] P. Brasnett, L. Mihaylova, D. Bull and N. Canagarajah, "Sequential Monte Carlo tracking by fusing multiple cues in video sequences", Image and Vision Computing, vol.25, **(2007)**, pp.1217-1227.
[68] S. M. Khan and M. Shah, "Tracking Multiple Occluding People by Localizing on Multiple Scene Planes", Pattern Analysis and Machine Intelligence, IEEE Transactions, vol.31, **(2009)**, pp.505-519.
[69] C. F. Shan, T. N. Tan and Y. C. Wei, "Real-time hand tracking using a mean shift embedded particle filter", Pattern Recognition, vol.40, **(2007)**, pp.1958-1970.
[70] J. Cao, M. Xing, J. H. Wang and N. N. Zhu, "Information Fusion Based on the Improved DS Theory in the Speaker Tracking System", J. of Computational Information Systems, vol.9, **(2013)**, pp.3571-3578.
[71] M. Heber, M. Godec, M. Rüther, P. M. Roth and H. Bischof, "Segmentation-based tracking by support fusion", Computer Vision and Image Understanding, vol.117, **(2013)**, pp.573-586.