دوفصلنامه محاسبات و سامانه‌های توزیع شده
www.jdcs.ir

تاریخچه مقاله:
تاریخ ارسال: ۹۸/۰۹/۱۰
تاریخ اصلاحات: ۹۸/۱۱/۱۲
تاریخ پذیرش: ۹۸/۱۲/۱۵
تاریخ انتشار: ۹۸/۱۲/۲۰

Combining Sequential and Parallel Tracking Strategies in Motion Mining: New Approach

Ghahari Bidgoli Milad[1*], Shakarami Ali[2]

[1]Sama Technical and Vocational Training College, Islamic Azad University, Islamshahr Branch, Islamshahr, Iran

[2]Islamic Azad University, Qom, Iran

Abstract

In recent years, research into motion mining and tracking of moving objects in real-time have attracted the attention of many researchers. Therefore, a new model for motion mining based on a combination of sequential and parallel tracking strategies has been presented in this paper in order to take advantage of them and reducing their shortcomings simultaneously. In fact, combining tracker-level model helps to choose the right motion mining algorithm based on input data features, and also reduces tracking error by synchronizing tracker activity with parallel and series strategies simultaneously. In comparison with other existing solutions, this model provides important advantages such as decreasing the response time, improving the speed and increasing accuracy for tracking moving objects in the higher layers.

Keywords:

Motion mining, objects tracking, parallel trackers, Sequential tracking

روش ارجاع به مقاله:
Combining Sequential and Parallel Tracking Strategies in Motion Mining: New Approach, Journal of Distributed Computing and Systems (JDCS) Vol.2, No.2, PP.178-186, 2020

Email: m.ghahari@siiau.ir

1. Introduction

Despite associated hardware for video data collection has grown very fast, requests of human have always a more rapid growth. To overcome this problem, we need a comprehensive computer system for analyzing and deciding on the image data set. Thus, it is vital to build a system to extract important data from the image data set and analyze it correctly. Motion tracking is a prevalent technique to record the movement of objects or people for immediate or delayed motion analysis and reuse. [1]

Knowledge Discovery in Databases (KDD) is a process that aims at finding valid, useful, novel and understandable patterns in data [2]. Data mining is a process that extracts understandable knowledge or patterns in data [3]. As a result, data mining is a part of KDD. In the praxis, KDD and data mining are used as synonym. Motion mining is also defined as a part of data mining [4].

There are many definitions for the motion mining in the literature. One of them defines motion mining as finding important patterns in image dataset which we could not achieve by searching and retrieving data set simply [5]. These patterns are used to improve decision making.

From other point of view, motion mining has been defined as knowledge discovery, patterns and the occurrences mining from image data set to explore the semantic structure of the images [6].

Motion mining identifies the important moving objects and tracks objects frame by frame as well as analysis objects to identify their behavior [7]. The existing architecture for motion mining includes three layers: Data preprocessing, Mining layer and Visualization layer [8].

In the first layer, data collect and clean for the next phases. In the Mining layer, some operations are performed to track and predict the future path of moving objects. In previous methods, due to lack of an efficient strategy and intelligent database, speed and accuracy of motion analysis is low. Sometimes there are plenty of similarities between target objects and non-target objects or background. Consequently, detecting moving objects is probably erroneous action. Because of this reason, it is necessity to isolate foreground data from background data in the mining process.

The remainder of this paper is structured as follows: In Section 2 challenges and related works in the field of motion mining are explored. The fusion strategies are presented in Section 3. In the next section, the architecture of a tracker system for motion mining is described. Section 5 suggests a new approach with combining fusion strategies and finally results have been proposed and the paper is concluded.

2. Related works and challenges

There are many applications which motivate us for motion mining field such as traffic analysis, MoveMine, moving animals and climate changes. Therefore, standard databases were established in the aforementioned fields to make a framework for adjusting parameters and visualization.

In [9], the paper addresses the two major challenges: motion representation and action recognition. Thus a novel motion descriptor based on optical flow has been proposed to describe the activity of the persons though they are in various views. The new motion descriptor has been framed by fusing shape features and motion features together to increase the performance of the system.

Figure 1 shows main logical components of a video tracker and main logical components of a tracking algorithm. As it is depicted, components of a tracking algorithm consist of five logical components [10]:

a) The definition of a technique to discover relevant information from an image area occupied by a target object. This technique can be based on object classification algorithms, detecting changes, or extracting couple of features such as color, gradient and edges.

b) Specifying a representation to encode the appearance and the target shape. This representation defines the characteristics of the target object that will be used by the tracker. Generally, the representation is a trade-off between accuracy of the description and invariance: it should be descriptive enough to cope with clutter and to discriminate false targets, while allowing a certain degree of flexibility to cope with changes of target scale, pose, illumination and partial occlusions.

c) Third component is a method that propagates state of the target. This method should use information from the feature extraction method. By this way, different instances of the same object over the time are linked.

d) The next logical component has to define a strategy to manage targets appearing and disappearing from the imaged scene. This step, referred to a track management, initializes the track for a new object of interest and terminates the trajectory associated with a disappeared target. When a new target shows itself in the scene, the tracker must initialize a new trajectory.

A new target usually happens:

- At the edge of the field of view of the camera,
- At specific entry areas (e.g. doors),
- In the far-field of the camera (when the size of the projection onto the image plane increases and the target
- becomes visible), or
- In another object (when a target separate from another target e.g. a driver get off a car).

Similarly, a trajectory must be stopped when the target:

- leaves from view of the camera, or
- Joins or disappears at another object (e.g. a building).

In addition, it is desirable to terminate a trajectory when the tracking performance is expected to degrade under a predefined level, thus generating a track loss condition.

e) The discovery of meta-data from the state in a compact and unambiguous form.

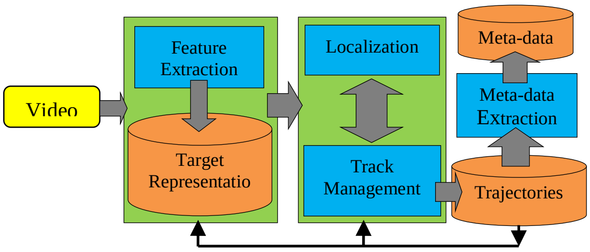

Fig.1: The video-tracking pipeline.

Figure 2 shows a summary of the main challenges in video tracking.

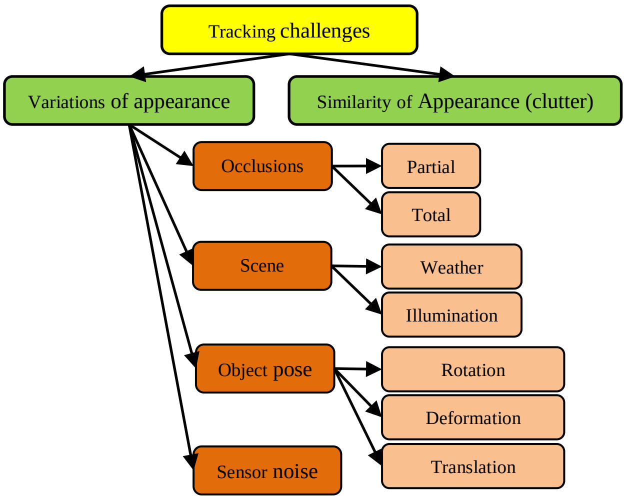

Fig.2: The main challenges in video tracking are due to temporal variations of the target appearance and to appearance similarity with other objects in the scene.

3. Fusion Strategies

Multi-feature fusion in video tracking can be performed both at the tracker level and at the measurement level. While tracker-level fusion enables using of a different range of tracker, fusion at the measurement-level avoids running multiple single-feature trackers, thus for reducing the problem should merge possible inconsistent or redundant tracking hypotheses [11].

Fusion at tracker level models single-feature tracking algorithms as black boxes. The video-tracking problem is redefined by modeling the interaction between outputs of black boxes, which can run in parallel (Figure 3) or in cascade (sequentially) [11-13].

An alternative is to perform the fusion sequentially, considering the features as if they were available at subsequent time instants (Figure 4).

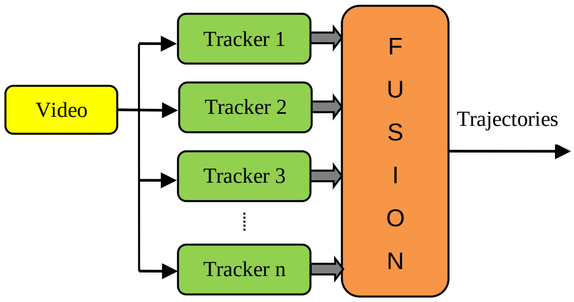

Fig.3: Fusion of independent estimates from parallel trackers.

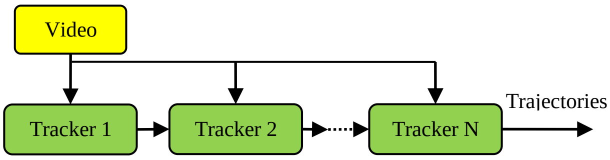

Fig.4: Sequential integration of tracking estimates in a cascade architecture.

The frame-by-frame measurement noise is used by the filter for each feature acts as a feature reliability estimator. The measurement noise can also be estimated in a training phase [14], and then it avoids adapting the feature contribution over the time, but reduces the flexibility of the tracker under changing scene conditions.

A summary of multi-feature tracking algorithms with fusion at the tracker level is given in Table 1. The table highlights the features and the fusion mechanism used by the trackers.

Table1: Comparison of feature types and fusion strategies used in video tracking algorithms that combine multiple features at the tracker level

Algorithm	Feature s	Fusion
Kanade–Lucas, particle filter [8]	Templat	Bayesian

	e	network
Condensation, Kalman filter [10]	Template, Blob, Color	Non-adaptive product
Particle filter [11]	Color, Contour	Product of pdfs
Condensation [12]	Template, Color	Covariance estimation
Extended Kalman filter [13]	Blob, Color, Geometry	Sequential integration

4. Architecture of Motion Mining

Motion mining tracks objects to extract useful patterns. The process analysis unknown important moving objects from the image data set using rule-based approach to detect or predict the semantic structure of a path without collision. The rich mining algorithms provide a new perspective in the analysis of image data set. This architecture has various components that work together to achieve high data purity.

At first, by using the methods of removal and separation of the background, important moving objects are separated. After removing irrelevant objects, method tracks target with helping estimation methods. Next, the results of the previous steps are used as the input data for motion mining.

Architecture of motion mining includes three layers as shown in Figure 5: collection and cleaning, mining and visualization layer.

The first layer collects the data of moving objects. In data mining, this layer is called data preparation step. Due to limitation in data collection techniques, there is the possibility of contradiction and confusion in the data set. So, data purity is required after data collection step.

In the second layer, mining layer, motion mining methods are performed on the data collection of moving objects. Among the existing algorithms in this field, periodic pattern, Swarm Pattern, Trajectory Clustering, Trajectory Classification and classification are the most important [8].

In the third layer, the results that are explored in the mining layer, are checked and then these results will be ready to use applications.

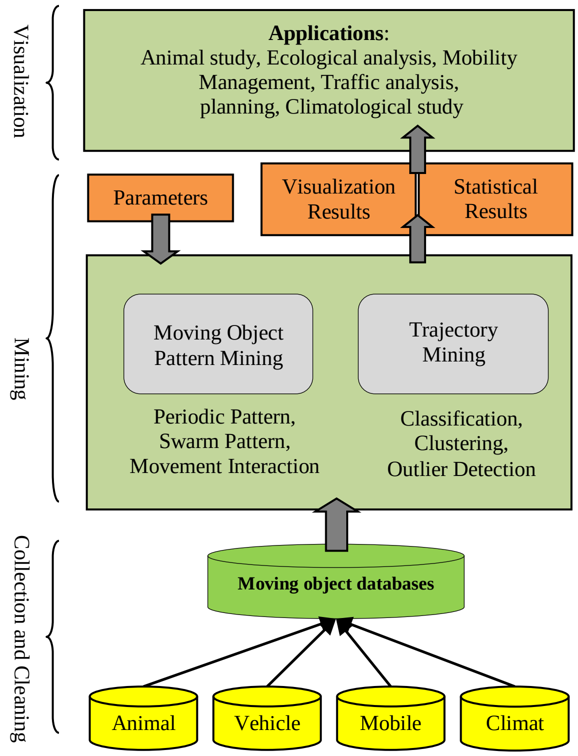

Fig.5: Architecture of a tracker system for motion mining.

5. Proposed Model

In the various models proposed for the architecture of the motion mining, results of each layer are delivered to upper layer. Errors over the bottom layers are published for the above layers. Ultimately speed and accuracy of mining are reduced. An additional layer has been used in the proposed model, called tracker-level fusion. This layer increases the accuracy of tracking results as shown in Figure 6. In the tracker-level fusion layer, the combination of series and parallel data fusion strategies are implemented simultaneously.

In the series data fusion strategy, combination of tracking is implicitly done by applying the characteristics of each tracker and results are sent to the next tracker. In the parallel data fusion strategy, all of trackers do their features on the initial data simultaneously and the final result will be achieved.

In the tracker-level fusion layer in which added for combination of results of these two strategies, the results of these two strategies are used in the tracker-level fusion layer and then the results of these two strategies are compared. Finally, the best result is selected between trajectories with low standard deviation and they are sent for using in the mining layer.

This results in sending precise results to the higher layers and thus helps in preventing the spread of errors in the higher layers. Using data mining algorithms in the mining layer, the output trajectories from layer composition has been investigated and these results are used in applications. By overlapping and simultaneous run time of strategies, improvement of tracking quality and prevention of wasting time in the mining layer is practicable. Consequently, it increases the tracking speed.

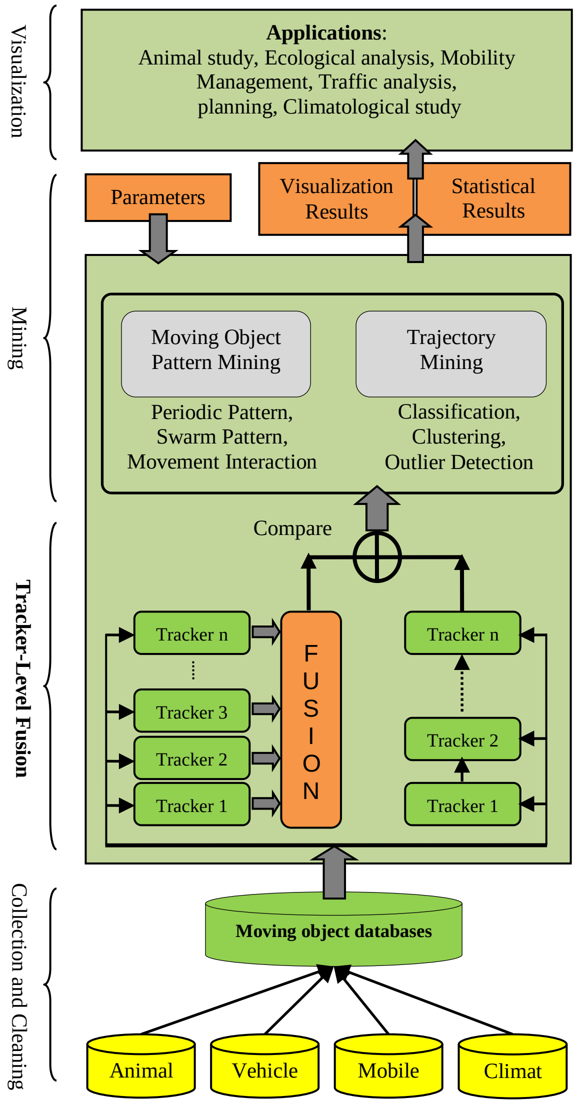

Fig.6: The proposed model of tracker system for motion mining.

6. Results

In this section, based on series and parallel detecting strategy, the result of executing single property and multi property detecting algorithms will be discussed. In executing single property algorithms, combination of two properties, say edge and fitting, are used that each of which has already been tested in single property detecting algorithms.

It has been used LabVIEW[1] to execute tests. The result of executing edge-property-based detecting algorithm by this software with the aim of detecting head of human in a sequence of images related to the data set of SPEVI[2] with a scrambled background is shown. This software is a powerful and flexible software to analyze measurement systems. Existence of a visual and straight connection between user of this software and the software itself results in writing and using existing programs in this software package and making this more exciting. Consequently, for using in industry, education, laboratory researches, it becomes a standard model to gather and process data and as a way to control and simulate virtual tools. In table 6-1 the important detail results of testing edge-property-based detecting algorithm have been summarized.

Table6-1: Summarized results of testing edge-property-based detecting algorithm

	Sample No	Time (ms)
Detecting the first edge	25	464
Detecting the last edge	4464	42930
Sum of samples	4464	
Sum of total detected edges	616	
Total time (ms)	42930	

6-1 Executing fitting-property-based detecting algorithm

In this section, by using LabVIEW, the result of executing fitting-property-based detecting algorithm is depicted. For recognizing the subject

[1] Laboratory Virtual Instrument Engineering Workbench
[2] Surveillance Performance Evaluation Initiative Dataset

by using the property of fitting, it has been used four samples as it is shown in figure 6-1.

Fig.6-1: Different samples of subject for applying fitting in the test using fitting-property-based detecting algorithm

In table 6-2 the important detail results of testing fitting-property-based detecting algorithm have been summarized.

Table6-2: Summarized results of testing fitting-property-based detecting algorithm

	Sample No	Time (ms)
Detecting the first successful fitting	0	332
Detecting the last successful fitting	447	47850
Sum of samples	448	
Sum of total successful fittings	446	
Total time (ms)	47850	

6-2 Multi-property combinational algorithms based on series-surface detecting

6-2-1 executing the detection test by combination of properties of edge and fitting

In this section, the results of multi-property combinational algorithms based on edge and fitting with series-surface detection in LabVIEW are shown. In this test, detecting head of human in a sequence of images related to the data set of SPEVI with a scrambled background is depicted. In series strategy, first detector recognizes the edge of incoming samples. Next, it fits the captured sample according to the subject on the recognized edges resulted from the last step.

In table 6-3 the important detail results according to the series strategy have been summarized.

Table6-3: Important detail results of moving object according to the series strategy with combining edge and fitting properties

The first stage of algorithm			The second stage of algorithm		
	Sample No	Time (ms)		Sample No	Time (ms)
Detecting the first edge	25	464	Detecting the first successful fitting	1	43106
Detecting the last edge	4464	42930	Detecting the last successful fitting	614	68482
Sum of samples	4480		Sum of incoming edge	616	
Sum of total detected edges	616		Sum of total successful fittings	418	
Total time (ms)	43024		Total time (ms)	68557	

In figure 6-2, the output of detecting algorithm with series strategy and based on combining the properties of edge and fitting over multi frame has been shown.

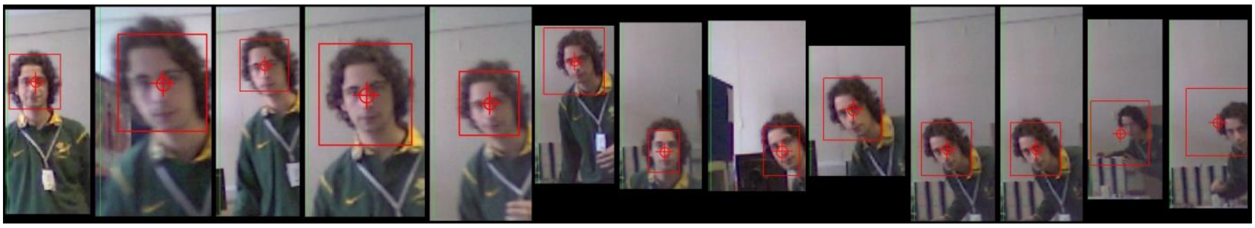

Fig.6-2: Detecting SPEVI multi frame data set according to detecting multi-properties detecting with properties of edge and fitting based on series-detector-surface strategy

6-3 Multi-properties combinational algorithms based on detector-surface-parallel strategy

6-3-1 executing the test of detecting with combining edge and fitting properties

In this section, results of executing multi-properties combinational algorithms based on edge and fitting properties with detector-surface-parallel strategy in LabVIEW software are discussed. In this test, detecting head of human in a sequence of images related to the data set of SPEVI with a scrambled background is aimed.

In series strategy, detector recognizes the edge. As soon as detection, according to the recognized edge, captured sample is fitted. The difference between series and parallel strategy is that in the former, all of edges are discovered. Then over this edge set, the fitting is done. In the latter, the fitting is done for any discovered edge.

In table 6-4, the important detail results have been summarized according to the parallel strategy.

Table6-4: Detail results of moving object according to the parallel strategy with combining edge and fitting properties

The first stage of algorithm for each sample			The second stage of algorithm for each discovered edge		
	Sample No	Time (ms)		Sample No	Time (ms)
Detecting the first edge	25	429	Detecting the first successful fitting	25	427
Detecting the last edge	4428	57399	Detecting the last successful fitting	4428	57419
Sum of samples	4480		Sum of incoming edge	547	
Sum of total detected edges	547		Sum of total successful fittings	405	
Total time (ms)	57680		Total time (ms)	57680	

In figure 6-3, the output of detecting algorithm with series strategy and based on combining the properties of edge and fitting over multi frame has been shown.

Fig. 6-3: Detecting SPEVI multi frame data set according to detecting multi-properties detecting with properties of edge and fitting based on parallel-detector-surface strategy

7. Conclusion

Motion mining tracks objects to extract useful patterns. The process analyze unknown important moving objects from the image data set using rule-based approach to detect or predict the semantic structure of a path without collision.

Architecture of motion mining includes three layers: collection and cleaning, mining and visualization layer. In the first layer, data are collected and cleaned for the next phases. In the mining layer, some operations are performed to track and predict the future path of moving objects. In the visualization layer, the results that are explored in mining layer are checked and then these results will be ready to use by applications.

In this paper, two strategies are combined in series and parallel data fusion strategies simultaneously. In addition, this way results in saving time and choosing the best possible result of data fusion strategies. Finding the best results increase the speed and quality in an accurate feature for the mining layer. It also increases the accuracy of the output of mining layer.

We propose starting the work in the mining model. According to the features of the input frame, a decision tree is used to select tracking algorithms in each tracker. This decision tree is built based on previous experiences and with helping an expert. The results of mining layer's operations can be used to update the decision tree.

We predict that using decision tree increases accuracy in tracking, and ultimately increases quality and speed of finding the future trajectories.

References

1. *Yinfu Feng, Mingming Ji, Jun Xiao, Xiaosong Yang, Jian J. Zhang, Yueting Zhuang, and Xuelong Li, "Mining Spatial-Temporal Patterns and Structural Sparsity for Human Motion Data Denoising", Zhejiang University, China, Bournemouth University, U.K., Chinese Academy of Sciences, china, IEEE, 2014.*
2. *U. Fayyad, G. Piatetsky-Shapiro, P. Smyth, R. Uthurusamy and editors, "Advances in Knowledge Discovery and Data Mining," AAAI/MIT Press,1996.*
3. *R. S.Gyorodi, "A Comparative Study of Iterative Algorithm in Association Rules Mining," Studies in Information and control, Vol.12, No.3.*
4. *H. Sundaram et al. (Eds.): CIVR 2006, LNCS 4071, pp. 360–369, 2006.Springer-Verlag Berlin Heidelberg 2006.*
5. *Ch. Bhatt and M. Kankanhalli, "Probabilistic Temporal Multimedia Data Mining," National University of Singapore ACM Transactions on Intelligent Systems and Technology, Vol. 2, No. 2, Article 17, 2011.*
6. *Y. Ding and G. Fan, "Multi-channel Segmental Hidden Markov Models for Sports Video Mining," MM'08,Vancouver, British Columbia, Canada, Copyright 2008 ACM 978-1-60558-303-7/08/10, 2008.*
7. *A. Yilmaz, O. Javed, and M. Shah, "Object tracking: A survey," ACM Comput. Surv.38, 4, Article 13. 2006.*
8. *Z. Ming, J. Lee, L. Tang, Y. Jiawei, H. Roland Kays, "MoveMine: Mining Moving Object Databases," SIGMOD'10, Indianapolis, Indiana, ACM 978-1-4503-0032-2/10/06, 2010.*

9. *M. Sivarathinabala, S. Abirami and R. Baskaran, "View Invariant Human Action Recognition Using Improved Motion Descriptor", Anna University, India, springer, 2015.*
10. *G. Hua and Y. Wu. "Measurement integration under inconsistency for robust tracking," In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, New York, 650–657, 2006.*
11. *A. Cavallaro, E. Maggio, "Video tracking: theory and practice," John Wiley & Sons, Ltd. 2011.*
12. *I. Leichter, M. Lindenbaum and E. Rivlin, "A probabilistic framework for combining tracking algorithms," In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Vol. 2, Washington, DC, 445–451, 2004.*
13. *F. Moreno-Noguer, A. Sanfeliu and D. Samaras, "Integration of conditionally dependent object features for robust figure/background segmentation," In Proceedings of the International Conference on Computer Vision, Washington, DC, 1713–1720,2005.*
14. *J. Sherrah and S. Gong, "Fusion of perceptual cues using covariance estimation," In Proceedings of the British Machine Vision Conference, Nottingham, UK, 564–573, 1999.*
15. *H. Veeraraghavan, P. Schrater and N. Papanikolopoulos. "Robust target detection and tracking through integration of motion, color, and geometry," Computer Vision and Image Understanding, 103(2):121–138, 2006.*

Milad Ghahari Bidgoli received the B.Sc degree in Computer Software Engineering from IA (Islamic Aazd) University of Kashan, Isfahan, Iran, in 2008. He also received two M.Sc. degrees in Computer Software Engineering and Chief Architect of Technology (Designing of Services in Large Scale System) from the IAU South Tehran Branch and DPI University, Tehran, Iran, in 2011 and 2016, respectively, and he currently is a Ph.D candidate in Computer Software Engineering in IAU Pardisan Branch. He has been with the Islamshahr Branch of IA University as a member of Research Faculty and worked on various aspects of Mosion Mining, Edge Computing and IOT. His current research interests include Learning and Edge Computing for IoT scenarios, including task offloading, resource allocation and auto-Scaling.

Ali Shakarami received the B.Sc and M.Sc. degrees in computer software engineering and computer networks' engineering from Payam-e-Noor university of Tehran (PNU) and Iran University of science and Technology (IUST), Tehran, Iran, in 2002 and 2013, respectively; He is currently a Ph.D candidate in computer software engineering in Islamic Aazd University, Qom branch. His research interest is distributed computing paradigms.