

Data Association and Prediction for Tracking Multiple Objects

M. Chandrajit¹, R. Girisha², T. Vasudev¹ and M. Hemesh¹

¹Maharaja Research Foundation, Maharaja Institute of Technology, Mandya - 571438, Karnataka, India;
chandrajith.m@gmail.com, vasu@mitmysore.in, hemeshm18@gmail.com

²PET Research Foundation, PES College of Engineering, Mandya - 571401, Karnataka, India;
wrtite2girisha@gmail.com

Abstract

Objectives: Tracking moving objects is essential for high level computer vision analysis, like object behavior interpretation or gait recognition etc. In this paper, a new method to track multiple moving objects in the surveillance video sequence is proposed. **Methods/Statistical Analysis:** Object tracking is done by extracting color moments feature from the segmented foreground object and associating individual objects in the successive frames using nearest neighbor classifier and Chi-Square dissimilarity measure. Further, object tracking during occlusion has been addressed by predicting the missing state of the occluded object by applying Lagrange's polynomial extrapolation using the apriori knowledge. **Findings:** The proposed method is evaluated using several challenging sequences of the benchmark IEEE PETS, IEEE CHANGE DETECTION, and EPFL dataset. Further, comparative evaluation with contemporary methods has been carried out to corroborate the efficacy of the proposed method. **Application/Improvements:** This work focuses on devising a new method to track multiple moving objects using color moments and Lagrange's extrapolation framework which works for a complex environment and in the presence of occlusions. Further, the shadow elimination task after the motion segmentation could be considered as an improvement step in this work which will aid in better tracking results.

Keywords: Color Moments, Lagrange's Extrapolation, Multiple Objects Tracking, Object Tracking, Video Surveillance

1. Introduction

In recent years, visual surveillance has become one of the most active research areas in computer vision. An automatic visual surveillance system generally contains five stages, such as motion segmentation, object classification, tracking, recognition and high-level motion analysis^{1,2}. Motion segmentation stage detects and extracts moving objects from the video frames. The object classification stage classifies the detected objects into predefined classes such as human or vehicle or animal, etc. The object tracking stage estimates the objects position in the successive video frames. Further, tracking can be performed on a single object or multiple objects using intra-camera or inter-camera setup. The former approach uses one camera to track the objects within its Field of View (FOV) and latter approach uses multiple cameras to track objects in overlapping FOV³. Finally, the high-level

motion analysis stage interprets the behavioral activity of the object.

Object tracking has a wide variety of applications, like automatic video surveillance, traffic video monitoring, accident prediction and detection, motion-based recognition, human computer interaction, human behavior understanding, etc. Object tracking in general is a very challenging problem due to complex and varying object shapes and backgrounds, loss of information caused by representing the real-world 3D into a 2D scene, noise in images induced by image capturing device, illumination changes as a result of dynamic environment, occlusions, shadows, etc.¹⁻⁶. The usage of object tracking in a variety of applications and challenges involved motivates to develop a robust algorithm for effective object tracking.

In order to address the challenges this paper presents a new method which uses color moments and Lagrange's

* Author for correspondence

polynomial extrapolation for tracking multiple moving objects in the video sequence captured in a complex environment. The proposed method makes an attempt to overcome some of the above mentioned problems, such as partial occlusions, complex and varying object shapes and background. This paper is organized as follows: Section 2 reports the related works on object tracking, Section 3 presents the overview of proposed work, Section 4 presents the proposed object tracking algorithm followed by the experimental results and conclusions in Section 5 and 6 respectively.

2. Related Works

Multiple objects tracking involves matching objects in successive frames of video sequence using features such as centroid, motion field, stereo disparity, edges, texture, colors and gradients. Effective and efficient tracking of moving objects from unconstrained surveillance video is a challenging task. Different methods reported in the literature for object tracking can be categorized into region based⁷⁻¹⁰, feature based¹¹⁻¹⁵, model based¹⁶⁻²⁴ and hybrid²⁵⁻²⁸. The region based method track objects exploiting changes in region of moving objects overtime. The feature based method track objects by extracting features corresponding to moving objects and then matching the features between frames of a video. The model based method track objects using apriori model of the moving object. In the case of hybrid method, combination of the above mentioned methods are devised to track the objects in frames of a video.

Kalman filter or particle filter are the mathematical predictive tools used to estimate the object location in the case of occlusion^{1,5}. Kalman filter is employed when the data associated to the object is assumed to be Gaussian distributed. Kalman filter uses three steps namely initialization, prediction and update to predict the state of the object in Bayesian framework. The particle filter is a variant of Kalman filter which is used for non-Gaussian object state and usually employed for tracking multiple objects.

Region based tracking across three cameras using Kalman filter is proposed in⁷. Optical flow based method for object tracking using two-way ANOVA is proposed in^{8,9}. However, the algorithm does not maintain the identity of tracked objects. A feature based tracking method in which corner points of extracted vehicle are

used for tracking is proposed in¹⁰. A feature based head tracker that uses color information and particle filter is proposed in¹² for single object tracking. Another feature based method that uses edges of object is proposed in¹³. However, the method is designed only for tracking the face of a person. A scene adaptive feature based method to track multiple objects is proposed in¹⁴. Object tracking method that use color and texture features and support vector machines is proposed in¹⁵. However, the object that needs to be tracked has to be manually selected in the first frame by the user, which makes the method not suitable for real-time applications.

In¹⁶ a combination of adaptive particle filter, appearance and motion information of the object is used for tracking. An object tracking method which use the local motion of object and color based particle filter is proposed in¹⁷. However, the method cannot be used for all practical applications as the object has to be selected manually and also the method is computationally expensive. A learning based method that learns color, size and motion to track objects across cameras using Kalman filter is proposed in¹⁸. A well-known early work namely Hydra¹⁹ tracks head of human by template matching approach. A model based method that uses Principle axis of an object and Homography constraints to match and track the object across different views of camera in Kalman filter framework is proposed in²⁰. However, the method works in controlled environment and not applicable for practical applications. A multiple objects tracking algorithm in which object association in successive frames is based on integer programming flow model is proposed in²¹. However, the algorithm cannot be used for real-time scenario as the accuracy decreases in case of dynamic environment. An algorithm to track single moving object is proposed in²², wherein the object appearance and motion was modeled using Kalman filter. However, the user needs to provide the objects' position in the first frame for tracking. A grid based modified inverse observation model for people detection followed by a multi hypothesis model for tracking of motion objects is proposed in²³. Particle filter based tracking by detection model for multiple objects tracking is proposed in²⁴.

A hybrid approach is proposed in²⁵ where statistical method is combined with CRF (Conditional Random Field) framework and sliding window optimization algorithm for labeling objects. Another hybrid approach that combines SSD (Square Sum of Difference) and color

based mean-shift tracker in Kalman filter framework is proposed in²⁶. A combination of region based and appearance based methods to track moving object is proposed in²⁷. Adaboost classifier and particle filter based method to track multiple humans is proposed in²⁸. Face detection based on Haar-Like features (machine learning based clustering technique) and mean-shift tracking is proposed in²⁹.

The region based tracking methods cannot reliably handle occlusions as well as do not accurately track in case of a complex environment with multiple moving objects. The feature based methods overcome the occlusion limitation of region based methods. However, the efficiency of feature based method decreases in case of distortion during perspective projection. The model based methods have some shortcoming such as the necessity of constructing model and high computational cost. The hybrid based methods are suitable for tracking objects in case of a complex environment. However, a suitable combination of methods to make a hybrid model is a challenging task.

Although there are several methods proposed in the literature for object tracking. These methods are based on constraints and assumptions such as: requires a training step for tracking, works for constrained environment, designed to track only part of an object like face, hand etc. and assumes no occlusion scenario. These constraints and assumptions may not be suitable for real-time environment. An attempt is made in this work by devising a new method to track multiple moving objects using color moments and Lagrange's extrapolation framework which works for a complex environment and in the presence of occlusions.

3. Overview of the Proposed Method

The proposed multiple objects tracking block diagram is shown in Figure 1 as an iterative sequence of the following phases: motion segmentation phase and multiple objects tracking phase. The Motion segmentation is done on the inputted frames of a video sequence by adopting our earlier work³⁰. The method segments the motion pixels by finding dissimilarities in pixel values considering its neighborhood values in successive frames using Chi-Square statistics. A blob hole filling step is required after

the motion segmentation because the blobs extracted from motion segmentation contains holes due to temporal differencing. Following the blob hole filling phase, a morphological erosion operation is performed to eliminate any noise in the frame. Subsequently, the features namely color moments, area and centroid are extracted from the segmented motion object blob using connected component analysis. The color moments and centroid features are used for the data association of the objects in successive frames. Further, in the case of occlusion the centroid feature of the object blob is used for predicting the missing state. Finally, the object blob area feature is used to eliminate the noise which may be present in the motion segmented frame.

The tracking phase of the proposed work is treated as a data association problem. That is, multiple objects tracking is treated as the correspondence of objects in successive motion segmented frames. This is achieved by correlating the features of objects using Chi-Square dissimilarity measure with nearest neighbor classifier and thus identity of an individual object is established. In addition, each tracked object within the FOV of a static camera will be given an identity. Consequently, it may happen that the moving object which was tracked until the previous frame within the FOV of a camera may not be tracked in the current frame because of occlusion. An occlusion in this context is detected when an already tracked object in the previous frame is not detected in the current frame. In such a case, the missing state of motion object blob in the current frame is estimated using the Lagrange's extrapolation framework.

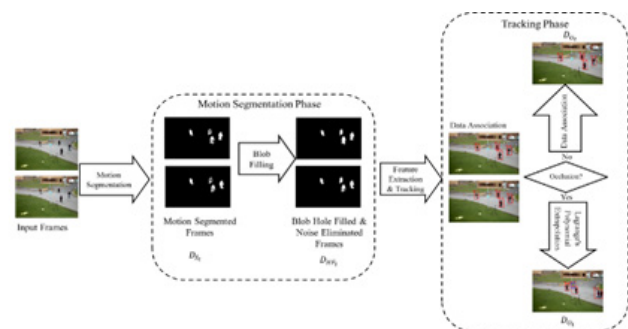


Figure 1. Block diagram of the proposed method.

The occlusion handling process consists of three steps namely initialization, prediction, and update. In the initialization step, the object centroid feature of three successive previous frames is used as aprior knowledge and

the prediction step uses this aprior knowledge to predict the missing state. Lastly, in the update step the predicted centroid value of the occluded object is recorded in the aprior knowledge base. Thereby, the proposed method tracks the multiple objects in video sequence based on data association and prediction.

4. The Proposed Method

4.1 Motion Segmentation using Chi-Square Statistics

The fundamental step of an automatic video surveillance is the segmentation of motion objects from video frames. Motion Segmentation is performed using Chi-Square test³⁰ for each pixel in temporal domain considering the 3×3 neighborhood and the frame is generated, where, t is the temporal frame. Subsequently a blob hole filling procedure³¹ is employed in order to fill any holes in the moving blobs. Further, morphological erosion operation is applied to eliminate any remaining noise and the hole filled noise suppressed frame is generated.

4.2 Feature Extraction

In general, the feature selection of an object is a challenging task. Moreover, the choice of features varies according to the tracking applications. In order to track an object which is very small, the centroid feature is more useful. On the other hand, to track an object which is big in size, a combination of various features could be used. In this work, object representation is done by using centroid and color moments features. Further, the area of the object blob is also extracted for the elimination of object which is less than the predefined size. A connected component analysis is applied on the object blob to extract color moments, centroid and area features as shown in Figure 2, Table 1 shows an example of extracted features from the segmented objects in the frame shown in Figure 2.

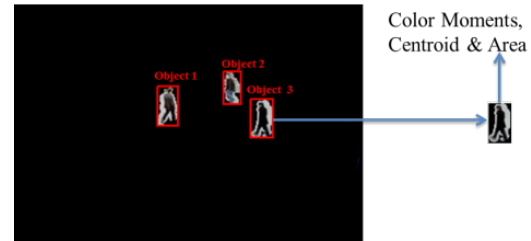


Figure 2. Feature extraction.

The color moments characterize both shape and color information of the object and used in image retrieval applications³². Moreover, the color moments are scaling and rotation invariant and hence, they can be used for varying illumination video sequences. Since the lower order color moments namely mean (Eq.1), standard deviation (Eq. 2) and skewness (Eq.3) are used as features in³², we are using the same features to represent the object blob in this work. The first three color moments are computed for the RGB color channels of each pixel by using the following equations.

$$\mu_c = \sum_{i=1}^N \frac{1}{N} P(w, h)_c \quad (1)$$

where, $C=\{R,G,B\}$, N =number of pixels in the object blob and (w, h) are the dimension of the blob representing width and height.

$$\sigma_c = \sqrt{\frac{1}{N} \sum_{i=1}^N (P(w, h)_c - \mu_c)^2} \quad (2)$$

$$S_c = \sqrt[3]{\frac{1}{N} \sum_{i=1}^N (P(w, h)_c - \mu_c)^3} \quad (3)$$

4.3 Multiple Objects Tracking

Tracking of multiple objects is the association of the segmented objects by using the features in the successive frames of a video. In this work, Chi-Square dissimilarity measure³³ is used to measure the association of objects features in successive frames. An object in frame $D_{(HF_t)}$ is

Table 1. Feature values

Object	Area	Centroid	Color Moments								
			μ_R	σ_R	S_R	μ_G	σ_G	S_G	μ_B	σ_B	S_B
1	592	(157.18,106.23)	65.79	72.98	0.69	64.12	73.24	0.73	63.17	72.46	0.75
2	433	(226.73,85.69)	57.71	62.40	0.62	58.26	63.70	0.60	58.58	64.02	0.57
3	434	(264.51,117.78)	67.81	78.60	0.53	68.46	79.68	0.54	68.62	79.68	0.53

assigned to an object in frame $D_{HF_{t+1}}$ which minimizes the Chi-Square distance. A nearest neighbor classifier with Chi-Squared distance is used to map the object association as shown in the following equation:

$$D(o_t, o_{t+1}) = \sum_{i=1}^F \frac{(o_{t_i} - o_{t+1_i})^2}{o_{t_i} + o_{t+1_i}} \quad (4)$$

where, $F = 10$ which represent the nine color moments and centroid features of an object and o_{t_i} and o_{t+1_i} are respectively the feature values of the object in the frames D_{HF_t} and $D_{HF_{t+1}}$ at the i^{th} feature.

The tracked objects in the FOV of a camera will be given the identity which is established by correlating the objects features in successive frames as mentioned above. In real-time scenario the objects entering or object exiting the FOV of camera are commonly encountered. Moreover, the object may enter or an existing object may exit the FOV of a camera from any direction. This situation is addressed by giving a new identity for the objects entry and eliminating the identity of the objects which exits the FOV respectively.

A partial and/or full occlusion of an object occurs in real-time situations when the scene is captured within the FOV of a single camera. In such cases, the prediction of object state in each input frame is very important for successful tracking. In the literature, most of the works on object tracking use either Kalman or particle filter for predicting the object state in a frame using the apriori knowledge. Further, Kalman filter works only when the data is assumed to be Gaussian while the latter is used for a non-Gaussian assumption. This paper focuses on exploring new method for object tracking which do not put any restriction on the assumption of data. The proposed work uses the Lagrange's polynomial extrapolation for object tracking. This tool is generally used to predict the values beyond the original observation range by using the apriori knowledge. In this work, we have used the centroid feature of the object blob and then predict the missing state (centroid) using Lagrange's polynomial extrapolation method^{34,35}. The general equation of Lagrange's n^{th} order polynomial extrapolation formula is given in the following equation.

$$f(x) = \sum_{i=1}^{n+1} f(x_i) \prod_{j=1, j \neq i}^{n+1} \frac{(x - x_j)}{(x_i - x_j)} \quad (5)$$

where, $f(x)$ is the prediction value that has to be

determined, x_i is the data point value at i^{th} position and x is data point to be predicted.

The occlusion handling stage in the proposed algorithm consists of three steps namely initialization, prediction and update. In the initialization stage, object centroid feature of three successive previous frames is recorded in the feature vector $C_{x,y}$ which will be the apriori knowledge of the individual object. This process of initialization in which the apriori knowledge is built, enables to predict the location of an object in the current frame accurately. In the prediction stage, the second order Lagrange's polynomial equation 6 and 7 are used to predict the centroid $(C_{x_{\text{Predicted}}}, C_{y_{\text{Predicted}}})$ of the occluded object as shown in Figure 3. Finally, in the update stage, the centroid values of objects are reinitialized in the knowledgebase. This iterative process is repeated for all the segmented objects.

The occlusion handling stage is employed only whenever a detected object in the previous frame is not detected in the current frame within the FOV of the camera. That is, if an existing object from previous frame is not assigned an identity in the current frame then we assume that the object got occluded by some other object and hence it is not detected. This situation is handled using the above mentioned process to predict the location of the occluded object within the FOV of a camera.

The first order Lagrange's polynomial ($n=1$) is a linear equation and prediction using a linear equation will be same as prediction using a straight line equation. On the other hand, if we use the third order Lagrange's polynomial ($n=3$) then the apriori knowledge has to be built by using the data from the four previous frames. Flexibility of building the apriori knowledge is questionable using the four previous frames in the case of low frame rate video sequences where object motion displacement will be high as shown in Figure 4. Hence, in this work we have used second order Lagrange's equation which uses three frames to construct the apriori knowledge.

$$C_{x_{\text{Predicted}}} = \sum_{i=1}^3 C_{x_i} \prod_{j=1, j \neq i}^3 \frac{(x - x_j)}{(x_i - x_j)} \quad (6)$$

$$C_{y_{\text{Predicted}}} = \sum_{i=1}^3 C_{y_i} \prod_{j=1, j \neq i}^3 \frac{(y - y_j)}{(y_i - y_j)} \quad (7)$$

where, x_i, y_i are the coordinate values of centroid at i^{th} frame.

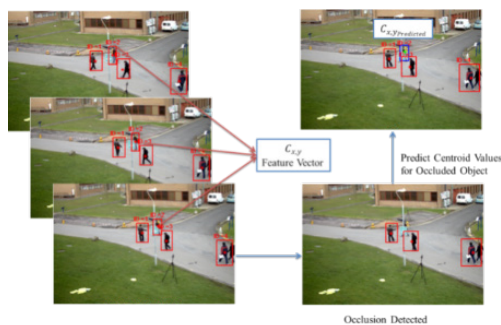


Figure 3. Occlusion handling.

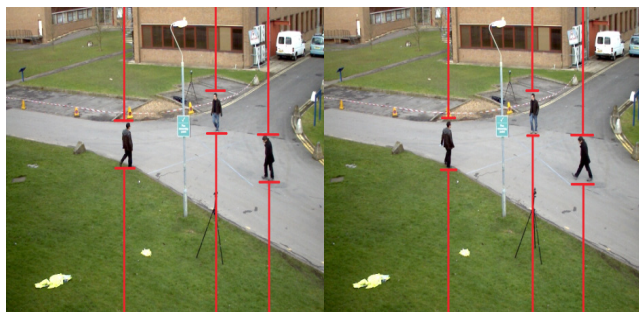


Figure 4. Example of low frame rate sequence IEEE PETS 2009 S2.L1 View_001 frames frame 00003-4.

5. Experiments

The algorithm has been tested on more than 70,000 frames of various sequences of IEEE PETS 2006, PETS 2009^{36,37}, IEEE CHANGE DETECTION (CD) 2014³⁸ and EPFL Multi-camera Pedestrian Videos^{39,40} dataset. Further, the dataset consists of a varied number of objects in both indoor and outdoor environments. The examples of results are shown in Figures 5–10 where, the row (a) is the output of motion segmentation stage, (b) is the result of blob hole filling processing stage and (c) is the output of object tracking stage. In the output frame, the tracked objects are numbered according to identity once they enter the FOV of a camera and shown using red asterisk inside the red bounding box. The predicted centroid for the object in the case of occlusion is shown in green asterisk inside the blue bounding box.

In the video frames shown in Figure 5(a), some motion object pixel region on the left top part of the frame (shown using red arrow) is not tracked as shown in Figure 5(c). This is because of the minimum area constraint considered for the object in the proposed method. However, even the objects with the smallest area can be tracked by varying the threshold.

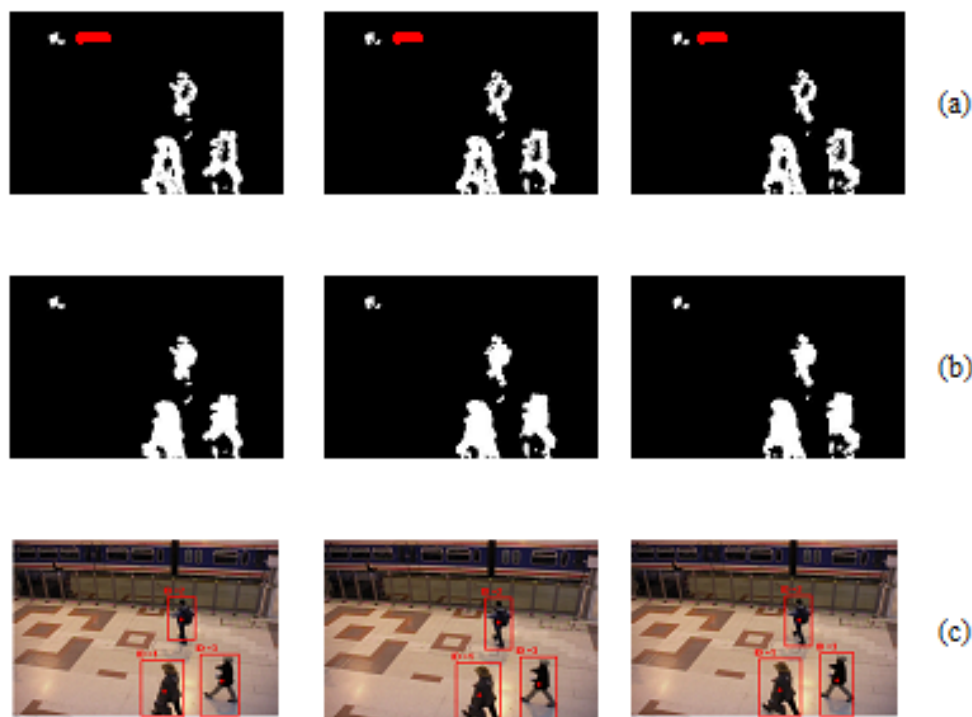


Figure 5. Tracking results of IEEE PETS 2006 dataset for frames in000076-78.

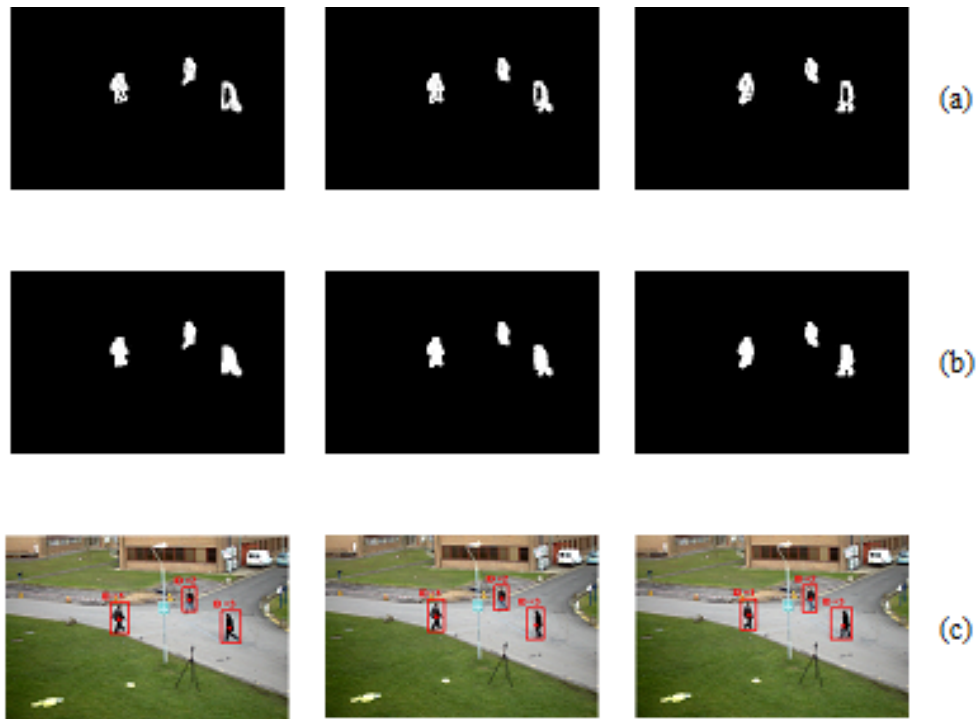


Figure 6. Tracking results of IEEE PETS 2009 S2.L1 View_001 dataset for frames frame00005-7.



Figure 7. Tracking results of IEEE PETS 2009 S2.L1 View_007 dataset for frames frame00013-15.

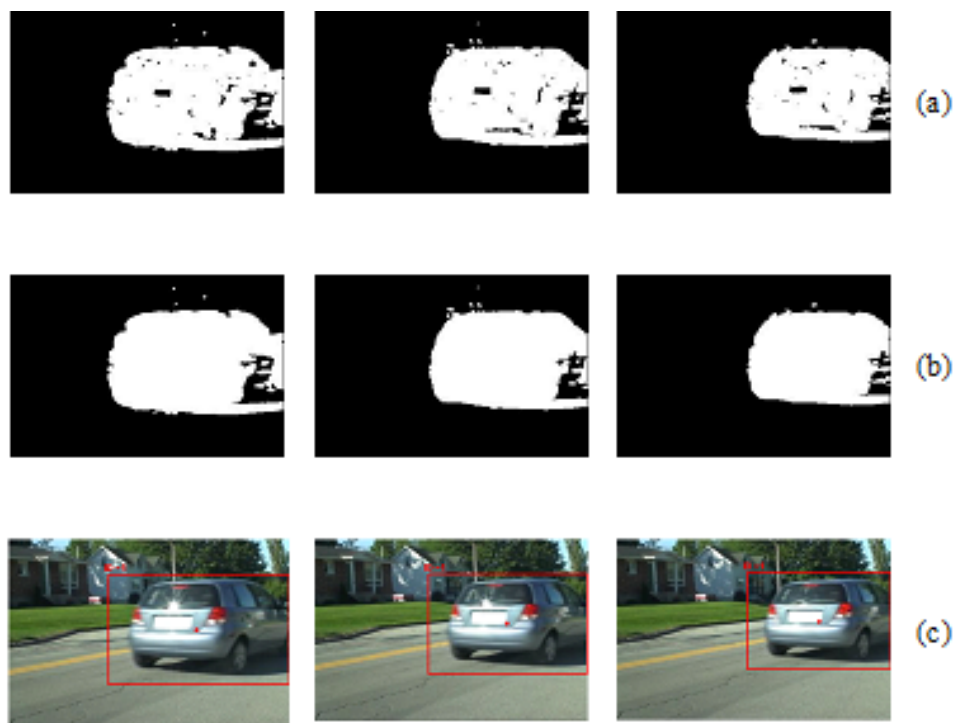


Figure 8. Tracking results of IEEE CD Bungalows dataset for frames in000033-35.

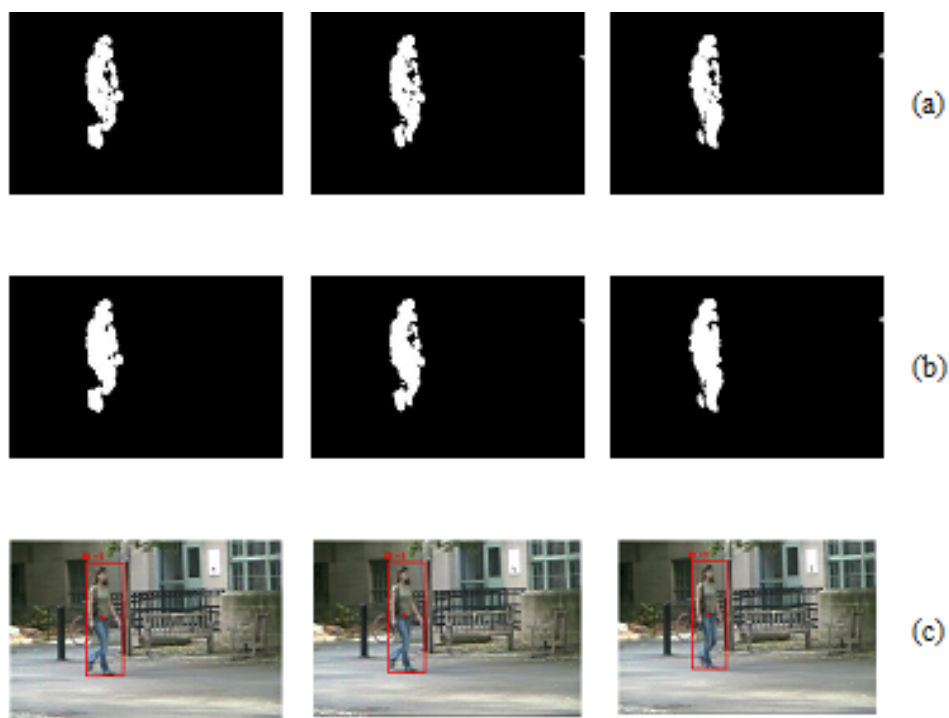


Figure 9. Tracking results of IEEE CD Backdoor dataset for frames in000297-299.

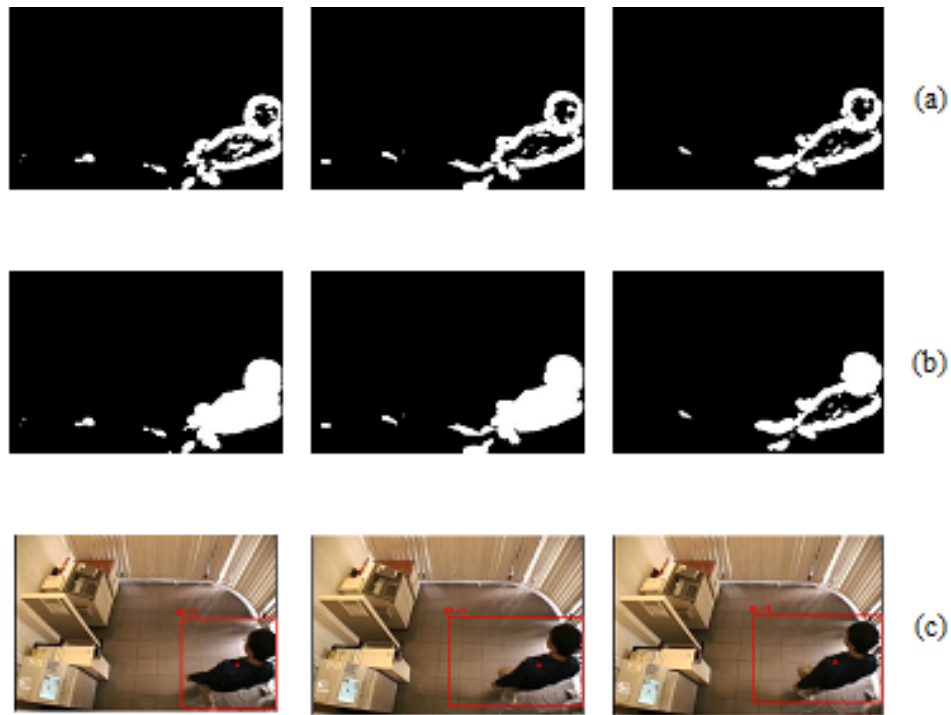


Figure 10. Tracking results of IEEE CD Copy Machine dataset for frames in000085-87.

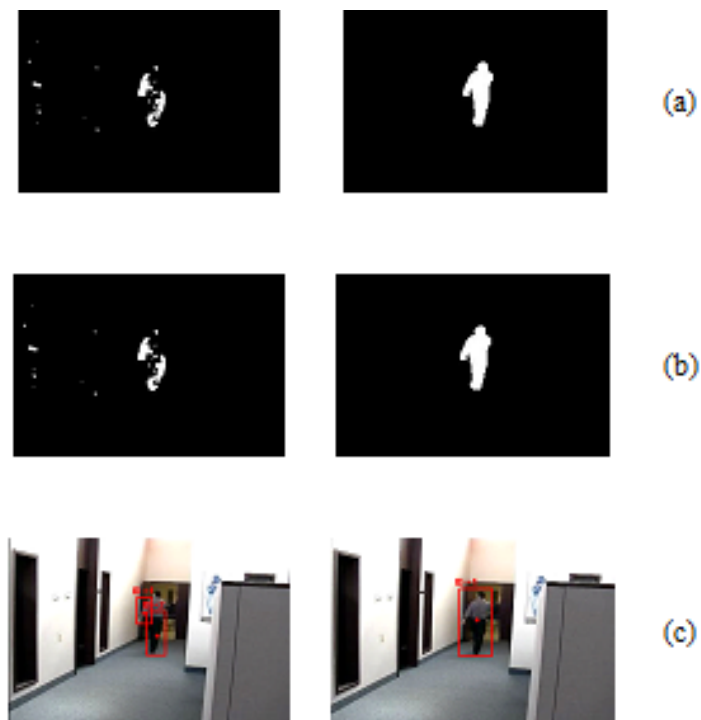


Figure 11. Tracking results of IEEE CD Cubicle dataset for frame in001268 (left) using motion segmentation method³⁰ (right) for manually segmented groundtruth frame.

The object tracking results are entirely dependent on the output of motion segmentation stage. Hence, the more accurate motion object blobs extracted the better tracking results by the proposed method as shown in Figure 11.

Figures 12–15 shows the occlusion results. The results validate the robustness of the algorithm when an occlusion occurs. Figure 15 shows an example in which the moving rack is given an identity and treated as a separate object even though it is moving along with the person. This situation happened because of the threshold value in the segmentation of the motion segmentation pixels which induces holes in motion objects.



Figure 12. Tracking under partial occlusion for frames in001873-74 of CD Backdoor dataset.



Figure 13. Tracking under partial occlusion for frames frame00018-19 of PETS 2009 S2.L1. View_001 dataset.

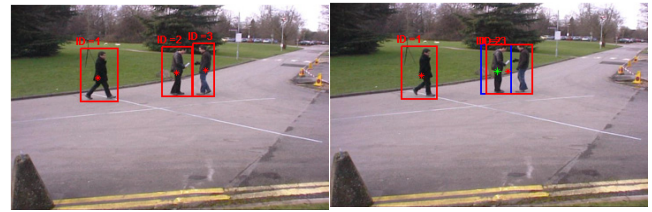


Figure 14. Tracking under partial occlusion for frames frame00027-28 of PETS 2009 S2.L1. View_007 dataset.



Figure 15. Tracking under partial occlusion for frames in000148-49 of PETS 2006 dataset.

Building a reliable model for prediction is a challenging task. In reality, the predicted value may be slightly deviated from the actual value. To show the efficacy of the proposed method, a graph for the predicted and the actual centroid values of the object for the IEEE CD Backdoor dataset is shown in Figure 16. The graph indicates that the algorithm is quite promising in predicting the objects missing state.

To corroborate the efficacy of the proposed system, performance has been ascertained by precision and recall accuracy metrics¹. The results obtained are shown in

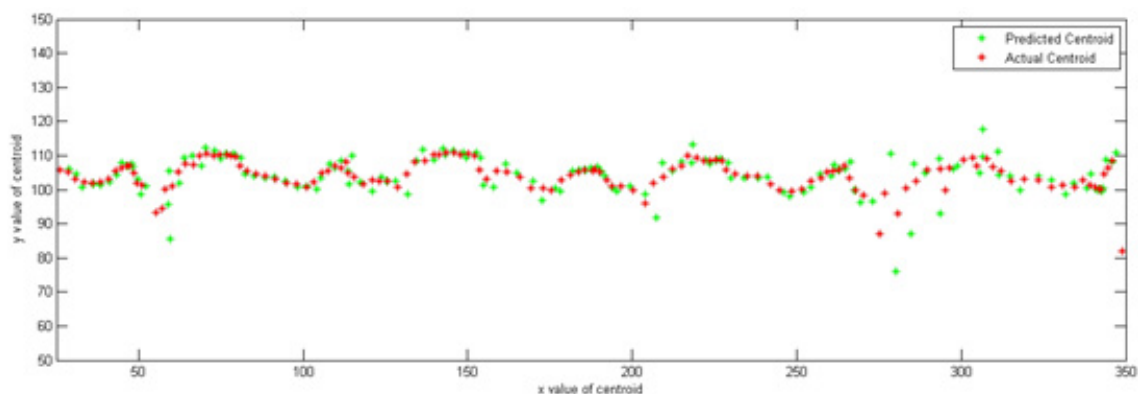


Figure 16. Graph of actual and predicted centroid for IEEE CD Backdoor dataset.

Table 2. The precision and recall can be defined as follows:

$$\text{Precision} = \frac{\text{No. correct correspondences}}{\text{No. established correspondences}} \quad (8)$$

$$\text{Recall} = \frac{\text{No. correct correspondences}}{\text{No. actual correspondences}} \quad (9)$$

where, the actual correspondence denotes the correspondence as per the ground truth data.

The sequence of IEEE PETS 2006 and IEEE CD Copy Machine contains the ghosts which are resulted because of non-stationary objects within the FOV of a camera. These ghost objects are considered as background in the motion segmentation process and therefore not segmented. Hence, recall is comparatively less for such datasets. On the other hand, the IEEE CD Bungalows dataset contains

moving cars captured from relatively small distance which has resulted higher precision compared to results of other sequences.

Comparative assessment with the state-of-the-art methods has been done and shown in Table 3. Further, comparison of the results with Kalman filter based method is presented in Figure 17 and Figure 18. The results validate that the proposed method performs better compared to Kalman filter based method.

For each stage of the proposed method the average computational time for a single frame having a pixel resolution of 240 x 360 pixels is shown in Table 4. The execution of the method was carried out using MATLAB R2011a on Intel Dual Core 1.8 GHz machine with Windows 7 Operating System.

Table 2. Performance accuracy of the proposed method

Dataset	Scene	No. of objects	Precision	Recall
IEEE PETS 2006	Indoor	15	0.80	0.62
IEEE PETS 2009	Outdoor	37	0.81	0.71
IEEE CD Bungalows	Outdoor	09	1.00	0.88
IEEE CD Backdoor	Outdoor	05	0.77	0.91
IEEE CD Copy Machine	Indoor	05	0.75	0.43
IEEE CD Cubicle	Indoor	02	0.88	0.78
Average			0.84	0.72

Table 3. Empirical comparison of tracking methods

Author	Approach	Entry	Exit	No. of Objects	Occlusion	Identity
Girisha and Murali ⁹	Region Based	No	No	Multiple	No	No
Nummiaro et al. ¹¹	Feature Based	-	-	Single	Yes	No
Martin and Martinez ¹³	Model Based	-	-	Multiple	Yes	Yes
Heili et al. ¹⁸	Hybrid	Yes	Yes	Multiple	No	Yes
Proposed	Hybrid	Yes	Yes	Multiple	Yes	Yes



Figure 17. Comparative tracking results for frames of IEEE PETS 2009 S2L1 dataset (top row) Kalman filter result (bottom row) proposed method result.

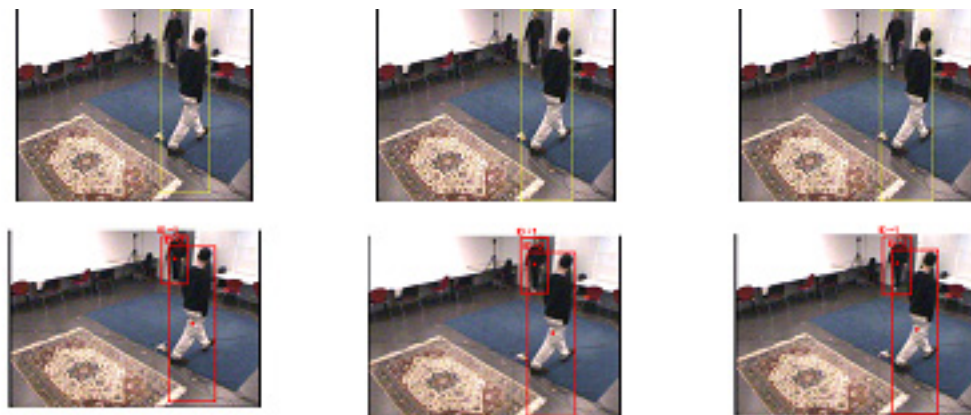


Figure 18. Comparative tracking results for frames of EPFL Laboratory dataset (top row) Kalman filter result (bottom row) proposed method result.

Table 4. Computational time

Stage	Time in secs
Motion Segmentation	0.09
Blob hole filling and Morphological processing	0.09
Tracking	0.24
Total	0.42

6. Conclusion

In this paper, a new algorithm to track multiple moving objects is proposed which use the color moments and Lagrange's polynomial extrapolation. The proposed method solves the object identity problem in successive frames by using the color moments feature extracted by the segmented foreground object blob and associating the object features using nearest neighbor classifier with Chi-Square dissimilarity measure. Further, the occluded object state is predicted using Lagrange's polynomial extrapolation which consists of three step process for establishing the missing state of object.

The method had been tested using the challenging IEEE PETS, IEEE CHANGE DETECTION and EPFL dataset which consists of both indoor and outdoor sequences with varied number of objects to prove its efficacy. Tracking accuracy using precision and recall accuracy measure is reported and the results obtained are encouraging. Further, comparative evaluation with the related methods has also been carried out to showcase the stability of the new method. The proposed method is able to perform significantly better compared to Kalman filter based method as per the experimental result. Incorporation of shadow segmentation model to improve tracking results will be considered as future directions.

7. References

1. Yilmaz A, Javed O, Shah M. Object tracking: A survey. *ACM Computer Surveys*. 2006; 38(4):1–13.
2. Hu W, Tan T, Wang L, Maybank S. A survey of visual surveillance of object motion and behaviours. *IEEE Transactions on Systems Man and Cybernetics, Part C: Applications and Reviews*. 2004; 34(3):334–52.
3. Wang X. Intelligent multi-camera video surveillance: A review. *Pattern Recognition Letters*. 2012; 34(1):3–19.
4. Soto ADM, Regazzoni CS. An overview on bayesian tracking for video analytics. *IEEE, Computer Vision and Image Understanding*. 2010; 104(1):90–126.
5. Yang H, Shao L, Zheng F, Wang L, Song Z. Recent advances and trends in visual tracking: A review. *Nerocomputing*. 2011; 74(18):3823–31.
6. Li L, Huang W, Gu IYH, Tian Q. Statistical modeling of complex backgrounds for foreground object detection. *IEEE Trans Image Process*. 2004; 13(1):1459–72.
7. Liem M, Gavril DM. Multi-person tracking with overlapping cameras in complex, dynamic environments. *Proceedings of the British Machine Vision Conference (BMVC)*; 2009. p. 1–1.
8. Girisha R, Murali S. Some new methodologies to track humans in a single environment using single and multiple cameras, *Doctoral Thesis, University of Mysore*; 2010. p. 1–3.
9. Girisha R, Murali S. Tracking humans using novel optical flow algorithm for surveillance videos. *Proceedings of the Fourth Annual ACM Bangalore Conference, Bangalore*; 2011. p. 1–7.
10. Dallalazadeh E, Guru DS. Feature-based tracking approach for detection of moving vehicle in traffic videos. *Proceedings of the First International Conference on Intelligent Interactive Technologies and Multimedia, U S A*; 2010. p. 254–60.
11. Naraghi MG, Koohi M, Sarbazy H. Object tracking based on estimation mesh method. *Indian Journal of Science and Technology*. 2012; 5(1):1–4.

12. Nummiaro K, Koller-Meier E, Svoboda T, Roth D, Gool LV. Color-based object tracking in multi-camera environments. DAGM-LNCS. 2003; 2781:591–9.
13. Zhu G, Zeng Q, Wang C. Efficient edge-based object tracking. Pattern Recognition. 2006; 39(1):2223–26.
14. Wang C, Liu H, Gao Y. Scene-adaptive hierarchical data association for multiple objects tracking. IEEE Signal Processing Letters. 2014; 21(6):697–701.
15. Altaf A, Raeis A. Presenting an effective algorithm for tracking of moving object based on support vector machine. Indian Journal of Science and Technology. 2015 Aug; 8(17):1–7. DOI: 10.17485/ijst/2015/v8i17/70326
16. Martin A, Martinez A. On collaborative people detection and tracking in complex scenarios. Image and Vision Computing. 2012; 30(2–5):345–54.
17. Kristan M, Pers J, Kovacic S, Leonardis A. A local-motion-based probabilistic model for visual tracking. Pattern Recognition. 2009; 40(6):2160–8.
18. Gilbert A, Bowden B. Incremental, scalable tracking of objects inter camera. Computer Vision and Image Understanding. 2008; 111:43–58.
19. Haritaoglu I, Harwood D, Davis LS. Hydra: multiple people detection and tracking using silhouettes. International Conference on Image Analysis and Processing, Venice; 1999. p. 280–5.
20. Seo DW, Chae HU, Kim BW, Choi WH, Jo KH. Human tracking based on multiple view homography. Journal of Universal Computer Science. 2009; 15(13):2463–84.
21. Xi Z, Xu D, Song W, Zheng Y. Algorithm with dynamic weights for multiple object tracking in video sequence. Optik - International Journal for Light and Electron Optics. 2015; 126(20):2500–7.
22. Jahandide H, Mohamedpour K, Moghaddam KA. A hybrid motion and appearance prediction model for robust visual object tracking. Pattern Recognition Letters. 2012; 33(16):2192–7.
23. Chang FM, Lian FL, Chou CC. Integration of modified inverse observation model and multiple hypothesis tracking for detecting and tracking humans. IEEE Transactions on Automation Science and Engineering. 2016; 13(1):160–70.
24. Aleksandrovich KP, Vyacheslavovich SP, Olegovich SSS. Fast multi-object tracking-by-detection using tracker affinity matrix. Indian Journal of Science and Technology. 2016 Jul; 9(27):1–13. DOI: 10.17485/ijst/2016/v9i27/97696
25. Heili A, Chen C, Odobez J-M. Detection-based multi-human tracking using a CRF model. Proceedings of IEEE ICCV; 2011.
26. Babu RV, Perez P, Boutheimy P. Robust tracking with motion estimation and local Kernel-based color modeling. Image and Vision Computing. 2007; 25:1205–16.
27. Thome N, Merad D, Mignet S. Learning articulated appearance models for tracking humans: A spectral graph matching approach. Image Communication. 2008; 23(10):769–87.
28. Ali I, Dailey MN. Multiple human tracking in high-density crowds. Image and Vision Computing. 2012; 30(12):966–77.
29. Yamashita A, Ito Y, Kaneko T, Asama H. Human tracking with multiple cameras based on face detection and mean shift. Proceeding of IEEE International Conference on Robotics and Biometrics; 2011. p. 1664–71.
30. Chandrajit M, Girisha R, Vasudev T. Motion segmentation from surveillance video sequence using chi-square statistics. Proceedings of International Conference on Emerging Research in Computing, Information, Communication and Applications. (ERCICA-14). 2016; 8(7).
31. Soille P. Morphological image analysis: Principles and applications. New York, Inc. Secaucus, NJ: USA; 1999. p. 173–4.
32. Afifi AJ, Ashour WM. Image retrieval based on content using color feature. ISRN Computer Graphics. 2012; 2012:11.
33. Pele O, Werman M. The quadratic-chi histogram distance family. Daniilidis K, Maragos P, Paragios N, editors. Computer Vision-ECCV'10, Springer-Verlag: Berlin Heidelberg; 2010. p. 749–62.
34. Lagrange Polynomial [Internet]. [Cited 2016 Jul 25]. Available from: http://en.wikipedia.org/wiki/Lagrange_polynomial.
35. Brezinski C, Zalia MR. Extrapolation methods: Theory and practice, North Holland; 1991.
36. PETS: Performance Evaluation of Tracking and Surveillance [Internet]. [Cited 2006 Jun 18]. Available from: <http://www.cvg.reading.ac.uk/PETS2006/data.html>.
37. PETS: Performance Evaluation of Tracking and Surveillance [Internet]. [Cited 2006 Jun 18]. Available from: <http://www.cvg.reading.ac.uk/PETS2009/a.html>.
38. Goyette N, Jodoin PM, Porikli F, Konrad J, Ishwar P. Changedetection.Net: A new change detection benchmark dataset. Proceedings IEEE Workshop on Change Detection (CDW-2012) at CVPR-2012, Providence: RI; 2012 Jun. p. 1–8.
39. Berclaz J, Fleuret F, Türetken E, Fua P. Multiple object tracking using K-shortest paths optimization. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2011; 33(9):1806–19.
40. Fleuret F, Berclaz J, Lengagne R, Fua P. Multi-camera people tracking with a probabilistic occupancy map. IEEE Transactions on Pattern Analysis and Machine Intelligence.