Occlusion Activity Detection Algorithm Using Kalman Filter for Detecting Occluded Multiple Objects

Heungkyu Lee* and Hanseok Ko**

* Dept. of Visual Information Processing

** Dept. of Electronics and Computer Engineering,

Korea University, Seoul, Korea

hklee@ispl.korea.ac.kr, hsko@korea.ac.kr

Abstract. This paper proposes the detection method of occluded moving objects using occlusion activity detection algorithm. When multiple objects are occluded between them, a simultaneous feature based tracking of multiple objects using tracking filters fails. To estimate feature vectors such as location, color, velocity, and acceleration of a target are critical factors that affect the tracking performance and reliability. To resolve this problem, the occlusion activity detection algorithm is addressed. Occlusion activity detection method provides the occlusion status of next state using the Kalman prediction equation. By using this predicted information, the occlusion status is verified once again in its current state. If the occlusion status is enabled, an object association technique using a partial probability model is applied. For an experimental evaluation, the image sequences for a scenario in which three rectangles are moving within the image frames are made and evaluated. Finally, the proposed algorithms are applied to real image sequences. Experimental results in a natural environment demonstrate the usefulness of the proposed method.

1 Introduction

The importance of detection, tracking, and recognition problems has received increased attention since visual tracking began to play an important role in surveillance systems, virtual reality interfaces and a variety of robotic tasks. But many key issues are not solved yet. The tracking of non-rigid objects and classifying their appearance model is a challenging problem in visual surveillance system. Especially, the monitoring and visual surveillance of human activity [1][2][3][4] requires complex tracking algorithms because of the unpredictable situations which occur whenever multiple peoples are moving, stopping, hiding behind obstacles and interacting with each other. Human actions within the field of view have no consistent rules concerning their movement. In addition, when multiple peoples are interacting with each other in a natural scene, a variety of events can occur such as occlusion, partial occlusion or short-time stopping. In such cases, the general tracking filter like a Kalman filter tends to fail because of sudden movements or sudden variation of speed.

Some tracking algorithms have a weakness according to the given specific situation. Feature-based tracking has a weakness in that it's difficult to estimate a centroid or the velocity of moving targets when the targets are occluded from each

other. A region-based approach has also same situation. The parametric method needs the calculation of an optimal fitting of the model to pixel the data in a region, and it should be updated continuously to the change of the appearance and the intensity variation while non-rigid objects are moving. However, when multiple peoples are occluded from each other, a parametric estimation of the individual person model is inaccurate. The view-based method to find the best match of a region in a search area with a reference template also has similar weaknesses with a region based approach in occlusion situations. These problems are due to the inaccurate estimate of state information such as centroid, color, velocity and acceleration of targets. Thus, in this paper, we propose the detection method for an accurate estimate of the state information using occlusion activity detection and an object association algorithm as shown in Figure 1. The occlusion activity detection algorithm provides the occlusion status information. By using this information, when the occlusion is activated, the proposed object association algorithm [5] can be applied to estimate the accurate state of information of the occluded multiple objects respectively, and then the general tracking algorithm is applied. On the contrary, if the occlusion is not activated, the general tracking algorithm is applied. This is due to the fact that the tracking algorithm is reliable when the occlusion between the objects has not occurred. Thus, the proposed algorithms can provide the reliable feature vectors for simultaneous multiple tracking algorithms even in the occlusion time.

This paper is organized as follows. In Section 2, we describe the proposed algorithms. It describes the detection method using the occlusion activity detection and object association. In Section 3, we evaluate our proposed algorithms having image sequences. Finally, the conclusive remarks are presented in Section 4.

Fig. 1. System block-diagram for detecting occluded multiple objects

2 Detection of Occluded Multiple Objects

2.1 Behavior of Moving Objects

The key issue of multiple targets tracking in the image sequences is the occlusion problem. During the occlusion time, unpredictable events can be activated. Thus, the tracking failure and missed tracking often happened. To cope with this problem, a behavior analysis of multiple targets is required the first time. Figure 2 describes all the possible action flows of targets within the field of view. A specific target enters into the specific field of view, and then it is moving, stopping, interacting with other targets. And finally, it leaves the field of view. We can classify the behaviors of multiple targets according to the action flow as in Figure 2.: (1) A specific target

enters into the scene. (2) Multiple targets enter into the scene. (3) A specific target is moving and forms a group with other targets, or just moves beside other targets or obstacles. (4) A specific target within the group leaves a group. (5) A specific target continues to move alone, or stops moving and then starts to move again. (6) Multiple targets in a group continue to move and interact between them, or stop interacting and then start to move again. (7) (8) A specific target or a group leaves a scene. The events of (1), (4), (5), and (7) can be tracked using general tracking algorithms. However, the events of (2), (3), (6) and (8) cannot be tracked reliably. Thus, to resolve this problem, we propose the occlusion reasoning method that detects occlusion activity status using Kalman Filter.

Fig. 2. State Transition Diagram; general action flow of objects within the FOV

2.2 Occlusion Activity Detection

We assume that we found the moving blobs, and then it starts to detect occlusion activity. Occlusion activity detection is an algorithm to provide the current status of occlusion between objects, which are just labeled blobs of a blob detection level. According to the occlusion status, a countermeasure to reliably track can be applied. We assumed that occluded objects from the first time have not appeared, and the objects are non-rigid. The procedure of occlusion activity detection is as follows.

- STEP 1: Occlusion Prediction

As in (a) of Figure 3, this step predicts the next positions (centroids) of blobs employing the usual Kalman prediction model used in JPDAF:

$$\hat{X}(k+1/k) = F(k)\hat{X}(k/k) + u(k) \quad (1)$$

where $X(k+1/k)$ is the state vector at time $k+1$ given cumulative measurements to time k , $F(k)$ is a transition matrix, and $u(k)$ is a sequence of zero-mean, white Gaussian process noise. Using the predicted position computed at equation (1), we can determine the redundancy of objects within the field of view using the intersection measure. The decision of the occlusion is computed by comparing if or not there is an overlapping region between the rectangular blocks MV_i in the predicted center points as follows.

$$F_{oc} = \begin{cases} 1 & \text{if } (MV_i \cap MV_j) \neq \emptyset \\ 0 & \text{otherwise} \end{cases}, \text{ where } i, j = 1, \dots, m \quad (2)$$

where F_{oc} is an occlusion alarm flag, the subscript i and j are the index of the detected target at the current frame, the block MBR_i that represents the validation region has the fixed range computed at the current frame and m is a number of a target. If a redundant region has occurred at the predicted position, the probability of occlusion occurrence in the next step increases. Therefore, the occlusion alarm flag is set to 1, and current status is maintained.

- STEP 2: Occlusion Status Update

In the current frame, the occlusion status is updated to decide the occlusion occurrence. The first time, the size of the labeled blobs is verified whether they are contained within the validation region or not. If the size of labeled blobs is contained within the validation region, the occlusion status flag is disabled. Then, the Kalman gain is then computed and the measurement equation is updated. If number of the predicted center point is greater than two, then the occlusion alarm flag is set to 1, we can conclude that the occlusion has occurred at the region. Thus, the occlusion status is enabled. At this time, the object association technique is applied to estimate the accurate center points of the respective blobs. From the predicted center points of the previous step including near that region, it is searched. In addition, the process transition mode is changed as in Figure 2.

Fig. 3. Occlusion prediction of next frame

2.3 Object Association

For the identity of the occluded blobs, the object association technique can be applied to associate a measured object with a real target when the occlusion status is enabled. That is due to the fact that the labeling procedure measures some targets as one target. In addition, the occlusion status may be maintained during some periods. This case causes a tracking failure due to miss-association if there is no association technique. To resolve this problem, an object association can be applied for not only the position decision in the occlusion state, but also for the decision of the identity of a target between frames. It can be a means for an attribute tracking method, which can be described as the process of combining a position and color information incorporating the data from a prior target model, a target dynamic model, and a feature measurement model through a buffering technique.

To do this, we applied the SEA to verify the target identification for an object association between a priori target model and a feature measurement model. The object pixel data buffered for the prior target model is used. This calculates the matching relationship between the buffered data in a queue and a candidate block. If we assume that the size of a blob is $N \times N$ pixels, the search window is of size $(2N+1) \times (2N+1)$ pixels in a basis of the predicted position. The mean absolute difference (MAD) is used to measure the match between two blocks. The match is performed on the current frame t using a previously stored blob model, a prior target model. The SEA algorithm as in [7] can be computed as

$$R - M(x, y) \leq MAD(x, y) \quad (3)$$

However, matched result of a hidden object behind a specific object may result in a false acceptance. Thus, we divide the reference block into N sub-blocks, then calculate a partial probability of candidate blocks as in Figure 4. It is an alternative evidential reasoning based approach for identity reasoning under the partial probability models. The concept of a typical sequence is defined in terms of a i, j -element partition, P_i , given the true target type i .

$$P_i = \{a_{i1}, \dots, a_{ij} / \text{target type } i\} \quad (4)$$

We consider it as a target if the sum of probability values of a sub-window is greater than and equal to a given threshold value as follows.

$$p(P_i) = \frac{1}{NM} \sum_{i=1}^N \sum_{j=1}^M \vartheta(a_{ij}) \geq Th \quad (5)$$

The matching probability of an occluded object is computed using an equation (6) after dividing into $i \times j$ partition window as in (b) of Figure 3.

$$\vartheta(i, j) = \begin{cases} 1 & \text{if } MAD(x, y) \geq R - M(x, y) \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

Thus, we can estimate the occupancy region of the occluded objects. By using this information, the center point of an individual object is calculated again.

3 Experimental Results

The proposed scheme was tested on real image sequences to assess its capabilities for tracking multiple moving targets (two people) in complex road scenes. Two different road scenes with an increasing complexity were considered. Acquired images were sampled at video rate: example 1 (total 180 frames, 15 frames per seconds, and its size is 240×320) and example 2 (total 70 frames, 15 frames per seconds, and its size is 240×320) which is a gray level image. These include the occlusion scenario. In addition, to show the robustness of the proposed methods, the image sequences including the occlusion scenario are depicted.

For the moving blob detection, an adaptive change detection algorithm is performed, and then a binarized algorithm is applied as in Figure 4. If the moving

blobs are detected, a labeling process is performed, then the center points of the respective blobs are calculated as feature vectors for tracking.

Fig. 4. Blobs detection experiments

To show the robustness of the proposed algorithm, the test image sequences having three rectangles moving are made as in (a) of Figure 5. The three rectangles are moving randomly with partial occlusion. This scenario file is simply binarized for the detection of moving blobs, and then labeled. Finally, the occlusion activity detection algorithm is applied, and if the occlusion status is enabled, the object association algorithm is applied. The (b) of Figure 5 shows the occlusion status of each rectangle in the image sequences. We can know that the occlusion activity detection can provide the precious information to process the occlusion problem.

Fig. 5. Test image sequences and its trajectories

Next, according to the occlusion status, the SEA using a partial probability for the object association is applied to the test image sequences. Table 1 describes the performance of occlusion activity detection rate and the root mean square (RMS) error of an object association algorithm.

The proposed algorithm is applied on the real image sequences having an occlusion scenario. The occlusion activity detection and objection association algorithm is applied according to the occlusion status. When the occlusion status is enabled, the object association algorithm using SEA gave the estimated center points. Using computed center points, the JPDA tracking filter [6] is performed. In the initial value of the JPDA algorithm to track multi-targets, the process noise variance = 10 and the

measurement noise variance = 25 are used. The initial position of two people are set to $A(17, 60)$, $B(254,147)$ and $A(16,115)$, $B(108,215)$ in Cartesian coordinates. In example 1, object A moved from left to right and object B moves from top to bottom. Object A is moved from the left-bottom to the right-top, and object B is move from the right-center to the left-center in example 2. An occlusion state is maintained for frames 34 and 24. Figure 6 depicts the tracking results and its trajectories. We know that the targets can be tracked reliably even while the occlusion status is enabled.

Table 1. Simulation results of test image sequences

Analysis of proposed methods	Results
Occlusion time	9
Total occlusion frame number	76
Accuracy of occlusion status	91.566
RMS error in estimated position x,y; object association	1.2

Fig. 6. Test image sequences and its trajectories

Fig. 7. RMS errors of test image sequences

We can evaluate the robustness of the object association algorithm using the RMS error of the computed position values as in Figure 7. In example 1, we know that the RMS error is high. This is due to the fact that the overlapping region is large. Meanwhile, (b) of Figure 7 shows that the RMS error is similar with that of the non-occlusion frames. This is due to that fact that this has a small overlapping region between targets.

4 Conclusions

In the proposed method, the occlusion activity detection gives the rules to cope with decision-making. In addition, the object association algorithm provides the estimated feature vectors even when the occlusion status is enabled. This is due to the fact that we know the target information such as velocity and acceleration. The proposed method reduced the computation time because the extra computation time for occlusion reasoning and association is spent when the occlusion status is enabled.

However, in the case of high variation of illumination change and shadow effects, some missing approximation is computed. Thus, the missing estimation of geometric information and an appearance model can give us a track failure. So, it needs a more accurate approximation and estimation method under a more complex situation and higher illumination variation environments, respectively.

Acknowledgements

This work was supported by grant No. 2003-218 from the Korea Institute of Industrial Technology Evaluation & Planning Foundation.

References

1. Haritaoglu, I, Harwood, D, Davis, L.S, "Hydra: multiple people detection and tracking using silhouettes" Visual Surveillance, Second IEEE Workshop on, pp 6 -13, June 1999.
2. S. J. McKenna, S. jabri and Z. Duric, A. Rosenfeld, H. Wechsler "Tracking Groups of people", Computer Vision and Image Understanding, pp42-56, 2000.
3. Romer Rosales and Stan Sclaroff, "3D Trajectory Recovery for Tracking Multiple Objects and Trajectory Guided Recognition of Actions", In Proc of IEEE on CVPR, June 1999.
4. Huwer, S.and Niemann, H., "Adaptive change detection for real-time surveillance applications", Visual Surveillance, IEEE International Workshop on, pp 37 -46, July 2000.
5. M. J. Swain and D. H. Ballard, "Colour indexing", International journal of Computer Vision, 7(1):11-32, 1991.
6. Y. Bar-Shalom and X. R. Li, Multitarget-multisensor tracking: principles and techniques, YBS Press, 1995.
7. W. Li and E. Salari, "Successive elimination algorithm for motion estimation, "IEEE Trans. Image processing. Vol 4, pp. 105-107, Jan. 1995.