ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 4, Issue 4, October 2014

Object Tracking under Heavy Occlusion based on Extended Kalman Filter, Particle Filter, and Color Matching

Youngsung Soh, Mudasar Qadir, Hadi Raja, Yongsuk Hae, Intaek Kim, Malik M. Khan, and Tayyab Wahab
soh@mju.ac.kr, mudasar.kalwar@gmail.com, nothan111@gmail.com, wise_sunys@nate.com, kit@mju.ac.kr, malikmk2003@yahoo.com, and tayyabwahab90@gmail.com
Myongji University, Yongin, South Korea

***Abstract*—Object tracking is useful in areas such as surveillance, access control, military applications, and health care. Objects may be human, vehicle, and other entities. A huge number of approaches have been proposed in this field. Different methods assume different environments and adopt various techniques ranging from deterministic to probabilistic approaches. Excellent surveys can be found in [1]. In this paper, we proposed a method for tracking objects by exploiting probabilistic power of track filters including extended Kalman filter and particle filter in the presence of heavy occlusion. We also use color matching to improve on tracking results. We applied the method to real world videos and obtained successful results.**

***Index Terms*—Object tracking, Extended Kalmanfiter, Particle filter, Color matching.**

I. INTRODUCTION

Object tracking plays an important role in many areas such as surveillance, traffic monitoring, health care, etc. To track objects, objects should be extracted first. There are many ways in segmenting objects from the image. They include, background differencing, frame differencing, optical flow based method, multi resolution method, etc. After objects are extracted, they should be tracked. Jalal et al. [1] broadly categorize tracking methods into top down and bottom up approaches. Top down approaches usually assume the presence of external input to initialize the tracking process. Mean shift method and its various variants are examples of top down approaches. In bottom up approaches, objects are extracted first by any segmentation algorithm and tracked by mapping objects in between consecutive frames. Most of these approaches suffer when difficulties such as noise, illumination change, and camera jittering occur. Jalal et al. [1] mentioned that wavelet is a promising tool to overcome these difficulties. There is yet another major difficulty that makes tracking hard. When there are occlusions, the characteristics of objects changes dynamically. Jalal et al. [1] introduced three types of occlusion. They are self-occlusion, inter-object occlusion, and object-background occlusion. In this paper, we propose a new method for tracking objects when there are heavy occlusions. We basically allow all three kinds of occlusions, but most of them are inter-object occlusions. When different objects are occluded and then separated, there occur merge and split of objects. Here the term "split" does not mean the split of a single object into several pieces. Rather it means that the group of objects breaks into smaller groups of objects or individual object. The proposed method utilizes the combination of extended Kalman filter (EKF), particle filter(PF), and color matching(CM) depending on merge and split scenarios. This paper is organized as follows. Chapter II reviews the related works for tracking under heavy occlusions. Chapter III describes the proposed method in detail. The experimental results are shown in Chapter IV. Finally Chapter V concludes the paper.

II. RELATED WORKS

Many methods have been proposed for multiple object tracking under heavy occlusion. Marceno et al. [2] proposed multiple object tracking method under heavy occlusion based on Kalman filter (KF) and shape matching. They argued if an object retains constant speed, KF alone works. If speed constraint is violated, shape matching should also be used. The shape they used is a binary sub image of individual object and match score is computed by simple correlation. Since their method is based on linear KF, it has an inherent limitation for nonlinear motion and the occlusion in their test data does not seem heavy. Liu et al. [3] used color appearance model instead of whole object appearance model to support occlusion cases. They combine R, G, and B color components with various linear coefficients to get feature set for target and non-target objects. To select reliable features, they apply online feature selection method based on Ada Boost proposed in [4]. After selecting reliable features, they extract regions having discriminative power measured by log likelihood of selected features. Finally they applied PF to track regions. They tested their method on station hallway sequences in PETS2006 database provided by University of Reading in U.K. They tested four different variations of the algorithm and showed that the one with feature selection capability and region based feature update and tracking performed best. Yang et al. [5] proposed a tracking method that can handle merges and splits of moving target. First, objects are segmented based on two level background maintenance model. Two levels are pixel level and frame level. Morphology is used to get final objects. Second, the presence of merges and splits are checked. If there is a merge, a group is created and is tracked as if it is a single target. If split is detected, feature correspondence is

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 4, Issue 4, October 2014

conducted to find the match. They considered three classes of features which are motion, appearance, and color. Among them, color was chosen due to its immunity to scene dynamics. They applied the method to a few public databases and reported good results. Pan et al. [6] proposed the content based tracking scheme where they performed content-adaptive progressive occlusion analysis to combine spatio-temporal context, target, and motion constraints. To track occluded target, they used variant-mask template matching. To prevent template from deterioration, drift-inhibitive masked Kalman appearance filter was introduced. They adopted a local best match authentication algorithm to handle complete occlusion.

III. THE PROPOSED METHOD

Fig. 1 shows the block diagram of the proposed method. Given two consecutive frames Ft and Ft+1, objects in two frames are segmented by background subtraction. Next the number of blobs and the size of blobs in two frames are compared to detect the presence of merges and splits. Depending on the situations, one of three possible scenarios is selected and performed. Those scenarios consist of three basic components. They are extended Kalman filter(EKF), particle filter(PF), and color matching(CM). By appropriately combining basic components, merges and splits can be handled and objects are tracked successfully. We explain each component in turn in detail along with the background subtraction method we adopted.

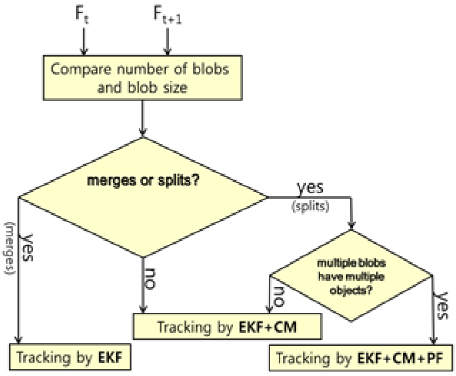


Fig 1.Block diagram of the proposed method

A. Background Subtraction

Background subtraction is used to extract objects. We used spatio-temporal Gaussian mixture model(STGMM) proposed by Soh et al. [7]. STGMM is based on Gaussian mixture model (GMM) [8]. STGMM considers both spatial and temporal variations of the image contents, whereas GMM takes into consideration only temporal variations. STGMM performed far better than GMM especially when there are background dynamics such as swaying tree branches, fluttering flags, and sea waves.

B. Extended Kalman Filter

Kalman filter (KF) is frequently used to model linear dynamics by linear evolution functions as in (1).

$$x_{k+1} = F_{k+1} x_k + w_k$$

$$z_k = H_k x_k + \qquad (1)$$

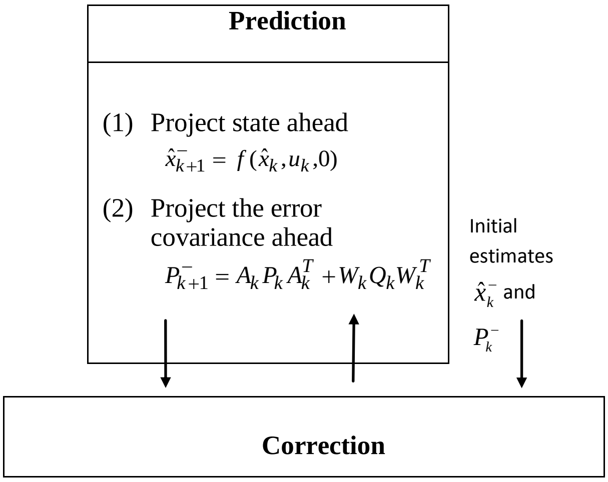
Here, :represents state vector, F_k a process linear function, and a process noise vector. Represents observation vector, a observation linear function, and a measurement noise vector. Human and vehicles are common objects considered in tracking. In real world situations, they do not necessarily follow linear motion. They change speed and direction either smoothly or abruptly, thus violating the linearity. To overcome this difficulty, EKF [9,10] was introduced as in (2).

$$x_{k+1} = F(k, x_k) + w_k$$

$$z_k = h(k, x_k) + \qquad (2)$$

Here $F(k,$ xand $h(k,$ xare nonlinear functions of state variables and evolutions. EKF is performed through two steps: prediction and correction. Fig. 2 shows the detail of these two steps. The measurement z_k is the external input and state estimate $\hat{x}_k$ is the final output. In order to derive the state estimate by prediction and correction, several parameters (A_k , H_k , Q , R) are required. A_k and H_k are $\frac{df}{dx}$ and $\frac{dh}{dx}$ respectively and they are extracted from the nonlinearity $\hat{x}_k = f(\hat{x}_{k-1})$ and $h(\hat{x}_k) = H_k \hat{x}_k$. Q is the covariance matrix of the process and R represents noise characteristics in the measurement.

As in KF, EKF finds the estimate $\hat{x}_k$ from the measurement z_k. The prediction step computes predicted value of estimate $\hat{x}_k$ and predicted error covariance. The correction step calculates Kalman gain and updates estimate and error covariance.



ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 4, Issue 4, October 2014

(1) Compute the Kalman Gain

$$K_k = P_k^- H_k^T (H_k P_k^- H_k^T + V_k R_k V_k^T)^{-1}$$

(2) Update estimate with measurement

$$\hat{x}_k = \hat{x}_k^- + K_k (z_k - h(\hat{x}_k^-, 0))$$

(3) Update the error covariance

$$P_k = (1 - K_k H_k) P_k^-$$

Fig 2. A complete operation of EKF with equations

C. Particle Filter

PF is a sequential Monte Carlo method for online learning based on a Bayesian framework. It has many other names such as bootstrap filter, condensation tracker, etc. It implements recursive Bayesian filter by Monte Carlo sampling where it represents the posterior density by a set of random particles with accompanying weights. Estimates are computed by generated samples and weights. There are many kinds of PFs in the literature. We choose to use sequential importance resampling (SIR) PF. To explain the algorithm, we use the notations used in Latecki [11] that are the extended version of Keith Copsey in Pattern and Information Processing Group, DERA Malvern. The mechanics of SIR PF is illustrated in Fig. 3.

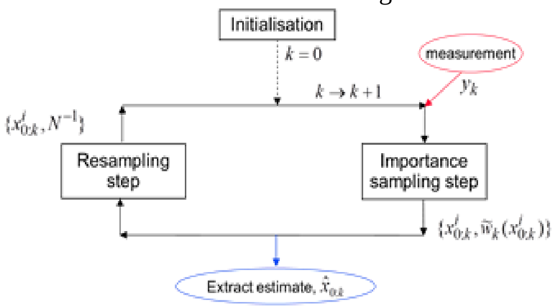


Fig. 3 SIR PF process

The algorithm consists of four steps. They are initialization, importance sampling, resampling, and iteration steps, and are described below [11].

1. Initialization
   - $k = 0$
   - For $i = 1, \ldots, N$ sample $x_0^i \sim p(x_0)$
   - and set $k = 1$
2. Importance sampling step
   - For $i = 1, \ldots, N$ sample $\tilde{x}_k^i \sim q(x_k | x_{k-1}^i)$

     and set $\tilde{x}_{0:k}^i = (x_{0:k-1}^i | x_k^i)$
   - For $i = 1, \ldots, N$ compute the importance weight w_k^i
   - Normalize the importance weight,

     $$\tilde{w}_k^i = w_k^i / \sum_{j=1}^{N} w_k^j$$
3. Resampling Step
   - Resample with replacement N particles: $(x_{0:k}^i; i = 1, \ldots, N)$
   - from the set: $(\tilde{x}_{0:k}^i; i = 1, \ldots, N)$
   - according to the normalized importance weights, $\tilde{w}_k^i$
4. Set $k \to k+1$
   - Proceed to the Importance Sampling step, as the next measurement arrives.

D. Color Matching

As Yang et al. [5] pointed out, color is an important feature that can discriminate different objects. In case of human, we extract color information of torso and leg parts as in Fig. 4. The numbers on both sides of black box represent the positional ratios when the box height is set to 1 measured from top to bottom. Thus the portion of the box from 0.25 to 0.375 belongs to torso and that from 0.625 to 0.75 belongs to leg. These positional ratios were obtained by manually dividing the body parts of accumulated object data and taking the average.

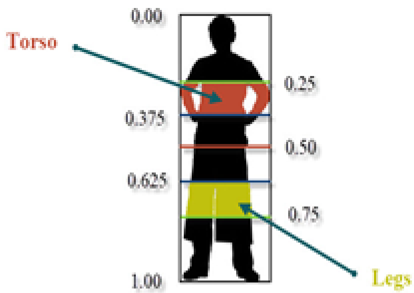


Fig. 4 Torso and leg parts of human body

To match the color similarity between two objects in consecutive frames, color histograms are built and compared. Here we use HSI representation of color since, unlike RGB, chromaticity and intensity components are well separated. We use only H and S to build histograms. Given histograms of objects under comparison, there are several ways to measure the distance between them. Equations (3) to (7) show some of the distance metrics using color histograms. Here and represent two color histograms under comparison.

Histogram intersection

$$H(H_Q, H_X) = \frac{\sum_i \min\left(H_Q(i), H_X(i)\right)}{\min\left(\sum_i H_Q(i), \sum_i H_X(i)\right)} \quad \textbf{(3)}$$

Euclidean distance

$$L_2(H_Q, H_X) = \sqrt{\sum_i \left(H_Q(i) - H_X(i)\right)^2} \quad \textbf{(4)}$$

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 4, Issue 4, October 2014

Bhattacharyya distance

$$\boxed{B(H_Q, H_X) = -\ln \sum_i \sqrt{H_Q(i) \times H_X(i)}} \quad (5)$$

Matusita distance

$$\boxed{M(H_Q, H_X) = \sqrt{\sum_i \left(\sqrt{H_Q(i)} - \sqrt{H_X(i)}\right)^2}} \quad (6)$$

Divergence

$$\boxed{D(H_Q, H_X) = \sum_i \left[\left(H_Q(i) - H_X(i)\right) \ln \frac{H_Q(i)}{H_X(i)}\right]} \quad (7)$$

We choose to use Bhattacharyya distance since it is known to perform best. Equations(8) to (10) show the Bhattacharyya distance we defined for our application.

$$B_H = \sum_{i=0}^{179} \sqrt{a_H[i] \times b_H[i]} \quad (8)$$

$$B_S = \sum_{i=0}^{255} \sqrt{a_S[i] \times b_S[i]} \quad (9)$$

$$B = \alpha B_H + (1 - \alpha) B_S \quad (10)$$

Here, and represent hue components of two histograms under comparison each having 180 bins, and and are saturation components each having 256 bins. is the weighting factor and is the final Bhattacharyya distance.

Since is the distance, similarity can be obtained by , where is a small constant to prevent dividing by zero.

IV. EXPERIMENTAL RESULTS

Test video for the experiment was captured in outdoor environment. Humans and vehicles are moving. Four human objects are experiencing heavy occlusion, in our case, many kinds of merges and splits. As depicted in Fig. 1, the proposed algorithmruns in three modes. They are merge, split, and none. Fig. 5 shows the tracking instance where merge has occurred. Among four individual humans, two having IDs 0 and 2 were merged into a single blob. Merged blob has label 2/0 to indicate that objects labeled 0 and 2 are in the same blob. Here we use the combination of EKF and CM. EKF was used for location prediction and CM was used to identify blobs having IDs 1 and 3. All the objects were tracked correctly.

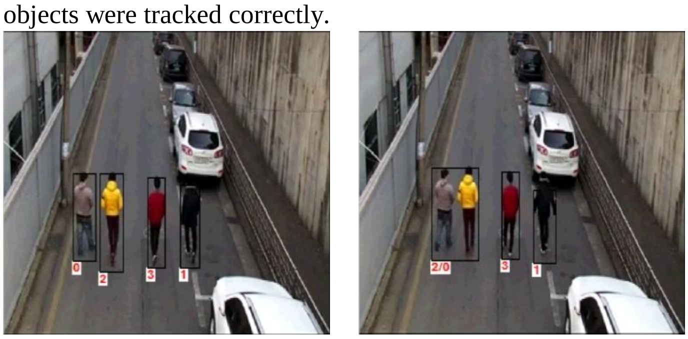


Fig. 5 Merge example

Fig. 6 shows the tracking instance where split has occurred. A single blob having four individual humans breaks into two blobs, one having three and the other having one object. Since there is only one blob having multiple objects, we use EKF+CM combination. Firstly, EKF predicts the location and then CM identifies the object having ID 1. Since there remains a big blob having multiple objects, all other IDs were allocated to that blob.

Fig. 6 Split example1

Fig. 7 depicts the tracking instance where another type of split has occurred. In this case, four in a group breaks into two groups of two objects. Since there is no blob having a single object, CM cannot be used. PF is applied to find which objects are in which blob. To do that, we exploit particles with associated likelihood values as weights that were computed using color histograms. When we compute likelihood value we use the equations (8) to (10) provided for CM. In the right image of Fig. 7, particles for four objects were displayed in different colors. All the objects were tracked correctly. By analyzing the particle locations in merged blobs, we even can find out the relative positioning of merged objects inside a blob.

Fig. 7 Split example2

Fig. 8 shows the tracking instance where none has occurred. In this case, only EKF is sufficient to track all the objects. All four objects were tracked successfully.

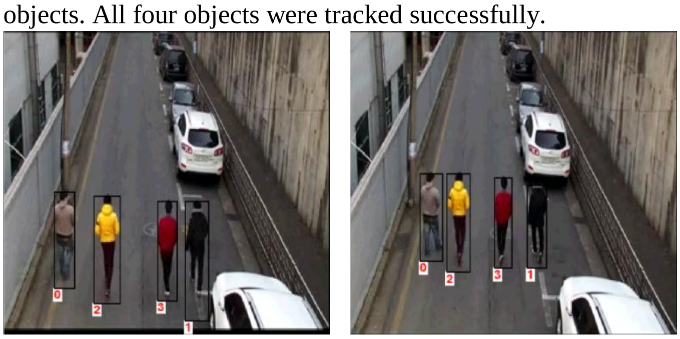


Fig. 8 No merge and split example

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 4, Issue 4, October 2014

V. CONCLUSIONS

Object tracking plays an important role in many areas such as surveillance, access control, health care, etc. Many approaches were proposed assuming various environments. One of the major difficulties for tracking is occlusion. In this paper, we proposed a new method for tracking under heavy occlusion. Depending on the occlusion scenario that can happen in real world, the proposed method used various combinations of EKF, PF, and CM. We tested the method with real world video and obtained successful results. Like many other tracking algorithms, the proposed method assumes that objects to be tracked have different colors. Difficulty is expected when objects with a similar color interact in nearby locations. This is intended for future research.

ACKNOWLEDGMENT

This work (Grant No.C0210325) was supported by Business for Cooperative R&D between Industry, Academy, and Research Institute funded by Korea Small and Medium Business Administration in 2014.

REFERENCES

[1] A. S. Jalal and V. Singh, "The state-of-the-art in visual object tracking,"Informatica, vol.36, pp.227-248, 2012.

[2] L. Marceno, M. Ferrari, L. Marchesotti, and C. S. Regazzoni,"Multiple object tracking under heavy occlusion by using Kalman filters based on shape matching,"IEEE International Conf. on Image Processing, pp.III-341-III-344, 2002.

[3] T. Liu and P. Xu,"Reliable multiple object tracking under heavy occlusion,"International Symposium on Intelligence Information Processing and Trusted Computing, pp.88-92, 2010.

[4] Y. J. Yeh and C. T. Hsu, "Online Selection of Tracking Features Using AdaBoost, IEEE Trans. Circuits and Systems for Video Technology," vol. 19, no. 3, pp. 442-446, April 2009.

[5] T. Yang, Q. Pan, J. Li, and S. Z. Li," Real-time Multiple Objects Tracking with Occlusion Handling in Dynamic Scenes,"IEEE International Conf. Computer Vision and Pattern Recognition, vol. 1, pp.970-975, 2005.

[6] J. Pan, B. Hu, and J. Q. Zhang," Robust and Accurate Object Tracking under Various Types of Occlusions,"IEEE trans. On Circuits and Systems for Video Technology, vol. 18, issue 2, pp.223-236, 2007.

[7] Y. Soh, Y. Hae and I. Kim, "Spatio-temporal Gaussian Mixture Model for Background Modeling," IEEE International Symposium on Multimedia (ISM), pp. 360-363, Dec. 2012.

[8] C. Stauffer and W. Grimson, "Adaptive background mixture models for real-time tracking," IEEE International Conference on Computer Vision and Pattern Recognition, pp.246-252, 1999.

[9] Julier, S.J.; Uhlmann, J.K., "Unscented filtering and nonlinear estimation," Proceedings of the IEEE, pp. 401–422, 2004.

[10] Branko Ristic, Sanjeev Arulampalam, Neil Gordon, Beyond the Kalman Filter, Artech House, 2004.

[11] https://www.google.co.kr/?gfe_rd=cr&ei=qRfHUa2OMW6mgWm5IHQBA&gws_rd=ssl#newwindow=1&q=particle+filter+powerpoint.

ABOUT AUTHORS

Youngsung Sohgot BS in electrical engineering in 1978 from Seoul National University in Seoul, Korea. He obtained MS and PhD in computer science from the University of South Carolina in Columbia, South Carolina, USA in 1986 and 1989, respectively.
His current interest of research includes object tracking, stereo vision, and parallel algorithms for image processing.

Mudasar Qadir got BS in computer science in 2011 from FAST-NU University in Peshawar, Pakistan. He entered a Master course in information and communication engineering in Myongji Universityin 2013.
His current interest of research includes object tracking, stereo vision and parallel algorithms for image processing.

HadiRaja got BS in electrical engineering in 2010 from Govt. University in Lahore, Pakistan. He entered a Master course in information and communication engineering in Myongji Universityin 2012..
His current interest of research includes object tracking, stereo vision and parallel algorithms for image processing.

YongsukHae got BS and MS in information and communication engineering from Myongji University in Yongin, Koreain 2009 and 2012, respectively. He entered a PhD course in information and communication engineering in Myongji University in Yongin, Koreain 2012.
His current interest of research includes object tracking, stereo vision, and parallel algorithms for image processing.

ISSN: 2277-3754
ISO 9001:2008 Certified
International Journal of Engineering and Innovative Technology (IJEIT)
Volume 4, Issue 4, October 2014

Intaek Kim received BS and MS in electronics engineering from Seoul National University in Seoul, Korea in 1980 and 1984 respectively. He obtained PhD in electrical engineering from Georgia Institute of Technology in Atlanta, Georgia, USA in 1992.
His research interest includes pattern recognition, image processing and smart grid area.

Malik M. Khan got BS in electrical engineering in 2011 from Govt. College University in Lahore, Pakistan. He entered a Master course in information and communication engineering in Myongji University in 2012.
His current interest of research includes vehicle to grid, Kalman filtering, background modeling and multi-viewpoint tracking.

TyyabWahab got BS in telecommunication engineering in 2012 from National University of Computer and Emerging Sciences in Peshawar, Pakistan. He entered a Master course in information and communication engineering in Myongji University in 2013.
His current interest of research includes vehicle to grid, Kalman filtering, particle filtering and multi-viewpoint tracking.