2010 IEEE International Conference on Robotics and Automation
Anchorage Convention District
May 3-8, 2010, Anchorage, Alaska, USA

Visual Tracking for Non-Rigid Objects using Rao-Blackwellized Particle Filter

Jungho Kim, Chaehoon Park and In-So Kweon

Abstract— **Particle filters have been used for visual tracking during long periods because they enable effective estimation for non-linear and non-Gaussian distributions. However, particle filter-based tracking approaches suffer from occlusion and deformation of the target objects, which result in the large difference between the current observations and the target model. Thus, we present a Rao-Blackwellized particle filter (RBPF)-based tracking algorithm that effectively estimates the joint distribution for the target state and the target model; in the proposed method, the target object is tracked by using the particle filter while the target model is simultaneously updated on the basis of the on-line approximation of a mixture of Gaussians. To ensure the robustness to occlusion, we represent the target model by 16 orientation histograms that are spatially divided, and individually update each histogram through a video sequence. We demonstrate the robustness of the proposed method under occlusion and deformation of the target objects.**

I. INTRODUCTION

Visual tracking is an important element in surveillance, guidance, or obstacle avoidance systems, whose role is to decide the position and the movement of the target objects. The problem of visual tracking is defined as sequentially estimating the state of a dynamic system using a sequence of noisy observations. For this task, there has been immense attention on particle filters [1][2][3] because, for any non-linear or non-Gaussian dynamic estimation problem, we can design an accurate, reliable, and fast recursive Bayesian filter by using the particle filter.

However, particle filter-based tracking algorithms have suffered from severe occlusion and large deformation of the target objects. Moreover, if the target objects move randomly and we cannot use prior information on smooth motion, this problem is more difficult. Therefore if the target model is fixed through a sequence of images, tracking is prone to failure because the appearances of the target objects can be easily changed by illumination conditions, pose changes or occlusion. Thus, in our approach, we estimate the joint distribution for the target state and the target model through the image sequence to cope with the large variations between the current observations and the appearance for the initial target model. However, estimation of the joint distribution for the target state and the target model is a high-dimensional problem, and it makes the particle filter infeasible especially when the target model is represented by high-dimensional visual information. Thus, we present an efficient tracking algorithm that adopts a Rao-Blackwellized particle filter [4], an effective method of reducing the dimensionality of a problem.

When one object is occluded by other moving objects, tracking results are unstable, and unreliable.Moreover, these unreliable results can affect the model update and tracking is prone to failure in the subsequent frames. To prevent it, we represent the target model by 16 orientation histograms [5] computed from 16 sub regions and individually update the distribution for each histogram by using the current observations.

This paper is organized as follows. Related works are presented in Section II. In Section III, the details of the Rao-Blackwellized particle filter are addressed. In Section IV, the tracking method is presented. In Section V, we present how to update the target model by using the on-line approximation of a mixture of Gaussians. In Section VI, the performance of the proposed method under occlusion and deformation of the target objects is demonstrated.

II. RELATED WORKS

Many approaches have been proposed to solve visual tracking problems. Particle filter-based visual tracking [1][6] has been extensively studied because it has shown to yield improvements in terms of performance over some conventional methods such as Kalman filter-based approaches. In [7][8], to solve the problem of tracking curves in a dense visual clutter, the authors proposed a condensation algorithm that uses factored sampling and learned dynamical models to propagate the random samples over time.

Comaniciu et al. [9] proposed a new method for the real-time tracking of non-rigid objects by using the mean shift algorithm. In [10], the author proposed an efficient technique for tracking 2D blobs through an image by using the mean shift algorithm.

Recently, the RBPF has been applied to tracking problems. For examples, in [11], the RBPF is used to integrate the subspace coefficient in an eigen tracking problem. The RBPF is also used to estimate the locations and identities of multiple objects [12]. Xu et al. [13] proposed an adaptive Rao-Blackwellized particle filter to improve the efficiency and accuracy of a regular particle filter. In [14], the authors proposed a tracking method to deal with face pose changes during tracking by using the RBPF and pose-dependent probabilistic PCA.

Jungho Kim and Chaehoon Park are with the Department of Electrical Engineering and Computer Science, Korea Advanced Institute of Science and Technology, Korea jhkim@rcv.kaist.ac.kr, chpark@rcv.kaist.ac.kr

In-So Kweon is a Professor of Electrical Engineering and Computer Science, Korea Advanced Institute of Science and Technology, Korea iskweon@kaist.ac.kr

Some previous works have been proposed to solve the tracking problems for the deformation and occlusion of the target objects. In [15], a tracking approach is addressed by representing the template object by multiple image patches to handle partial occlusion and pose changes. Wu et al. [16] proposed a dynamic Bayesian network that accommodates an extra hidden process for occlusion. Ross et al. [17] proposed a tracking method that incrementally learns the appearance of the target. Zhou et al. [18] proposed a tracking approach that incorporates adaptive appearance models in a particle-filter framework. Recently, Babenko et al. [19] addressed the problem of learning an adaptive appearance model by training a discriminative classifier in an online manner.

III. Rao-Blackwellized Particle Filter for Tracking

Particle-filter-based tracking approaches estimate the distribution of the target state, x_t, with given observations up to time t, $Z_{1:t}$ and an initial target model, M, as shown in Eq. (1).

$$p(x_t|Z_{1:t},M) \tag{1}$$

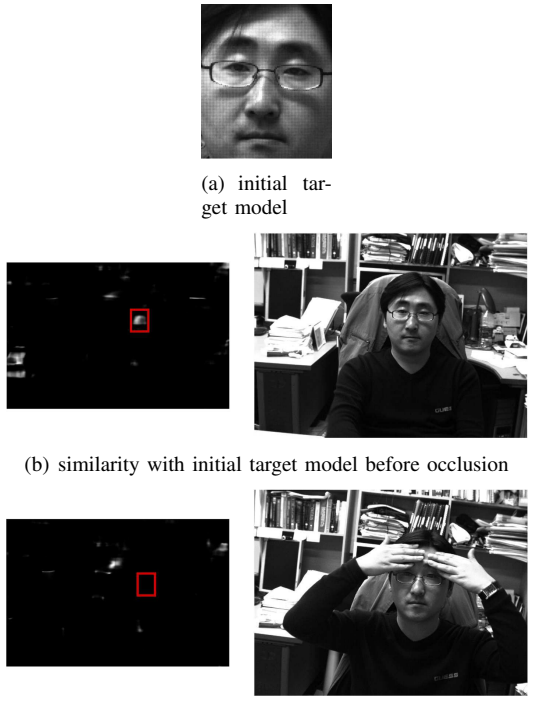
To verify the week point of the particle filter under occlusion and deformation, we compute the likelihood values for all possible locations in the input image. Each left image of Fig. 1 shows the likelihood values for each corresponding image at the right. The brighter pixels represent the higher likelihood values. As shown in Fig. 1(c), the locations that may correspond to the target object do not have the higher likelihood values due to occlusion of the target object and complex clutter in the background.

In order to track the non-rigid objects under occlusion, we estimate the joint distribution for the target state and the target model. In our algorithm, the target model is represented by the 128-dimensional orientation histogram [5], and that makes the particle filter infeasible.

To solve this problem, an effective method of reducing the dimensionality of a problem is to exploit the analytic structure, known as Rao-Blackwellization [20]. The resulting algorithm is referred as a Rao-Blackwellized particle filter. The system state vector consisting of the target state x_t and the target model $M_{1:t}$ is first repartitioned, as referred in Eq. (2). Thus, the problem of estimating the joint distribution is decomposed into an estimation problem for the target state and estimation problems for the target model that are conditioned on state estimates.

$$\begin{aligned} p(x_t,M_{1:t}|Z_{1:t},d_{1:t}) &= \frac{p(x_t,M_{1:t},Z_{1:t},d_{1:t})}{p(Z_{1:t},d_{1:t})} \\ &= p(x_t|Z_{1:t},d_{1:t})\,p(M_{1:t}|x_t,Z_{1:t},d_{1:t}) \end{aligned} \tag{2}$$

where $M_{1:t}$ is the target model that is updated up to time t, and $d_{1:t}$ is data association between the target state and the target model. In the Rao-Blackwellized particle filter, each particle x_t^i has its own target model $M_{1:t}^i$, as shown in Eq. (3). Thus, $d_{1:t}$ simply means the previous target model belonging to each particle and the weight for each particle



(a) initial target model

(b) similarity with initial target model before occlusion

(c) similarity with initial target model after occlusion and illumination changes

Fig. 1. The tracking results by using the conventional particle filter when occlusion occurs

is evaluated by comparing the current observations with the corresponding target model in the tracking stage. We update the target model at time t using the previous target model for each particle and the corresponding current observations in the model update stage.

$$\begin{aligned} &p(x_t,M_{1:t}|Z_{1:t},d_{1:t}) \\ &= \sum_{i=1}^{N} p\left(x_t^i|Z_{1:t},d_{1:t}\right) p\left(M_{1:t}^i|x_t^i,Z_{1:t},d_{1:t}\right) \delta\left(x_t - x_t^i\right) \end{aligned} \tag{3}$$

IV. Target State Estimation

In the tracking stage, we estimate the probability (weight) for each particle $p\left(x_t^i|Z_{1:t},d_{1:t}\right)$ involved in Eq. (3). This posterior distribution is decomposed by using Bayes theorem, that is the mechanism for updating the knowledge about the target state from new data, as shown in Eq. (4).

$$p\left(x_t^i|Z_{1:t},d_{1:t}\right) = \eta\, p\left(Z_t|x_t^i,d_{1:t},Z_{1:t-1}\right) p\left(x_t^i|d_{1:t-1},Z_{1:t-1}\right) \tag{4}$$

where η is a normalization term that makes the sum of the weights 1.

Suppose that the required distribution $p(x_{t-1}^j|d_{1:t-1},Z_{1:t-1})$ at time t is available. The prior distribution is formulated by Eq. (5).

$$\begin{aligned} p\left(x_t^i|d_{1:t-1},Z_{1:t-1}\right) &= \sum_j p\left(x_t^i,x_{t-1}^j|d_{1:t-1},Z_{1:t-1}\right) \\ &= \sum_j p\left(x_t^i|x_{t-1}^j\right) p\left(x_{t-1}^j|d_{1:t-1},Z_{1:t-1}\right) \end{aligned} \tag{5}$$

The prediction stage involves the motion model $p(x_t^i|x_{t-1}^j)$ which evolves the target candidates at time t, and when observations Z_t become available, the update stage is carried out.

A. Motion Model

To define a probabilistic model for the state evolution between time steps, we adopt a random walk that is based on a uniform density of the previous state and the uncertainties as follows:

$$p\left(x_t^i|x_{t-1}^j\right) \sim U\left[x_{t-1}^j - \mathrm{b}, x_{t-1}^j + \mathrm{b}\right] \tag{6}$$

where b represents our uncertainty about the incremental changes of the target object in the image. To reduce the computational cost in the computation of the orientation histograms for newly generated particles, we calculate the minimum range of the image that includes all predicted particles, as shown in Eq. (7). It is used when constructing the integral images, as introduced in Section 4.2.

$$\begin{aligned} &u_{max} = max\left[u_t^i + s_t^i\right], \quad u_{min} = min\left[u_t^i - s_t^i\right] \\ &v_{max} = max\left[v_t^i + s_r s_t^i\right], \quad v_{min} = min\left[v_t^i - s_r s_t^i\right] \\ &i = 1, 2, \cdots, N \end{aligned} \tag{7}$$

where u_t^i and v_t^i are the coordinate for a particle x_t^i, and s_t^i and s_r are the scale of the particle (a half of the width) and the fixed scale ratio (height / width) that is determined by the initial target region. N is the number of particles.

B. Observation Model

In the update stage, the observations Z_t are used to modify the prior distribution to obtain the required posterior distribution for the current target state. For this purpose we use the 128-dimensional orientation histogram z_t^i that assigns a consistent orientation based on local image properties within the region defined by each particle x_t^i. The orientation histogram is formed from the gradient orientations within a region around the target location and it has shown to give reliable feature matching over a wide range of scale and illumination changes. Recently, orientation histograms have also been used for human detection [21]. Moreover, we can reduce the computational cost for computing the orientation histograms by constructing the integral images. The integral images have been used for face detection and feature detection to reduce the computation time, as introduced in [22][23]. Even though we compute numerous histograms (3000 histograms when managing 3000 particles), the computation does not require much time because, in this algorithm, a majority of the computation time is devoted to computing integral images [1]. Moreover, when constructing the integral images, we use the minimum range defined in Eq. (7).

[1]In our implementation, the computational costs for computing 100 and 3000 histograms are 0.016 sec and 0.031 sec, respectively. (The image resolution is 640 × 480 pixels.)

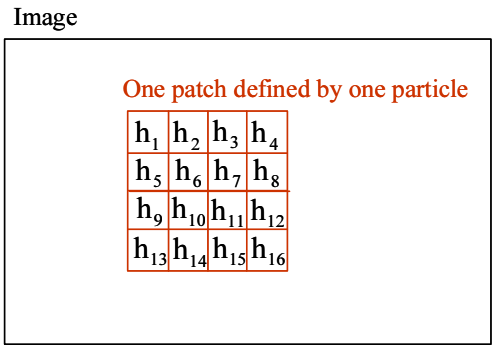


Fig. 2. Observation $z = [h_1 h_2 \cdots h_{16}]^T$, h_i is an 8-dimensional vector, and we individually normalize the orientation histogram for each sub region *i.e.* $\|h_i\| = 1/16$

Because our target model is represented by a mixture of Gaussians (MoG), as mentioned in Section V-A, the likelihood is determined by

$$p\left(Z_t|x_t^i, M_{1:t-1}\right) = \exp\left(-\frac{1}{2G}\sum_{j=1}^{B}\sum_{k=1}^{G}\left(z_t^{i,j} - \mu_{1:t-1}^{i,j,k}\right)^T\left(\Sigma_{1:t-1}^{i,j,k}\right)^{-1}\left(z_t^{i,j} - \mu_{1:t-1}^{i,j,k}\right)\right) \tag{8}$$

where G is the number of Gaussians for one MoG, and B is the number of sub regions. $z_t^{i,j}$ is one histogram corresponding to the j^{th} sub region for the i^{th} particle (*i.e.* h_j). $\mu_{1:t-1}^{i,j,k}$ and $\Sigma_{1:t-1}^{i,j,k}$ are the mean and the covariance corresponding to the k^{th} Gaussian for the j^{th} sub region, that are updated up to time $t-1$, respectively.

V. Target Model Estimation

A. Target Model Representation

While the Gaussian distribution has some important analytical properties, it suffers from significant limitations when it comes to modeling the real dataset. Thus, we represent the target model by a mixture of Gaussians.

When the target objects are partially occluded or deformed, that causes large dissimilarity between the previous target model and the current observations. What is worse, if the occluding object contains highly textured regions, the occluded parts strongly affect the tracking results because we generally use the normalized histograms. In this case, occlusion of the small portion causes a large change in the orientation histograms. However, by maintaining the spatially divided histograms for each target model that are individually normalized and by individually updating each histogram, the proposed algorithm is not strongly influenced by the effect of occlusion.

For each particle, we individually update the distribution for its target model by using the corresponding current observations [2]. The target model $M_{1:t}$ is composed of 16 MoG that are equally weighted, *i.e.* $M_{1:t} = [m_{1:t}^1, m_{1:t}^2, \cdots, m_{1:t}^{16}]$, as shown in Eq. (9).

[2]For simplicity, we omit the index of the particle

$$p(M_{1:t}|x_t, Z_{1:t}, d_{1:t}) = \prod_{j=1}^{B} p\left(m^j_{1:t}|x_t, Z_{1:t}, d_{1:t}\right)$$
$$p\left(m^j_{1:t}|x_t, Z_{1:t}, d_{1:t}\right) = \frac{1}{G}\sum_{k=1}^{G} g\left(m^j_{1:t}, \mu^{j,k}_{1:t-1}, \Sigma^{j,k}_{1:t-1}\right)$$
$$g(m, \mu, \Sigma) = \frac{1}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}(m-\mu)^T\Sigma^{-1}(m-\mu)\right) \tag{9}$$

Also, the covariance matrix is assumed to be the form:

$$\Sigma^{j,k}_{1:t-1} = \left(\sigma^{j,k}_{1:t-1}\right)^2 \mathrm{I} \tag{10}$$

B. Target Model Update

We update the posterior distribution for the target model by using on-line approximation of a mixture of Gaussians introduced in [24]. When computing likelihood values for particles, as explained in Section IV-B, we compare the current observations for each sub region z^j_t with the G Gaussian distributions that are represented by $\mu^{j,1}_{1:t-1}, \Sigma^{j,1}_{1:t-1}, \cdots, \mu^{j,G}_{1:t-1}, \Sigma^{j,G}_{1:t-1}$. In this step, as shown in Fig. 3, we select the n^{th} and l^{th} Gaussian distributions corresponding to most and least probable distributions with the current observations by Eq. (11).

$$\begin{aligned} n &= argmax_c\left[g\left(z^j_t, \mu^{j,c}_{1:t-1}, \Sigma^{j,c}_{1:t-1}\right)\right] \\ l &= argmin_c\left[g\left(z^j_t, \mu^{j,c}_{1:t-1}, \Sigma^{j,c}_{1:t-1}\right)\right] \end{aligned} \tag{11}$$

If none of the G distributions match the current observations, which means $g(z^j_t, \mu^{j,l}_{1:t-1}, \Sigma^{j,l}_{1:t-1}) < T_g$, this least probable distribution is replaced with a distribution defined by the current histogram z^j_t as its mean and the pre-defined initial variance. T_g means a threshold value for determining whether the Gaussian distribution matches the current observations or not.

Otherwise, the μ and σ parameters of the n^{th} Gaussian distribution that matches the new observations are updated as follows:

$$\begin{aligned} \mu^{j,n}_{1:t} &= (1-\rho^{j,n})\,\mu^{j,n}_{1:t-1} + \rho^{j,n} z^j_t \\ (\sigma^{j,n}_{1:t})^2 &= (1-\rho^{j,n})(\sigma^{j,n}_{1:t-1})^2 + \rho^{j,n}(z^j_t - \mu^{j,n}_{1:t})^T(z^j_t - \mu^{j,n}_{1:t}) \end{aligned} \tag{12}$$

where $\rho^{j,n} = \beta g(z^j_t, \mu^{j,n}_{1:t-1}, \Sigma^{j,n}_{1:t-1})$ and β is a constant for the learning rate.

To prevent the drift problem caused by inaccurate tracking, the first Gaussian distribution which is computed from the initial model image remains same, *i.e.* we find two Gaussian distributions in Eq. (11) among the G distributions except the first Gaussian distribution.

The model update step is performed when generating the new particles at the time t because we do not use all the particles maintained at time $t-1$. We generate new particles x^i_t according to the weights of the previous particles, w^i_{t-1}. Which means that the previous particles that are evolved to the current state are determined by the re-sampling process [1] whose role is to eliminate samples with low importance weights and to multiply samples with high importance weights. Therefore, to reduce the computation time, the update for the target model is performed during the re-sampling process.

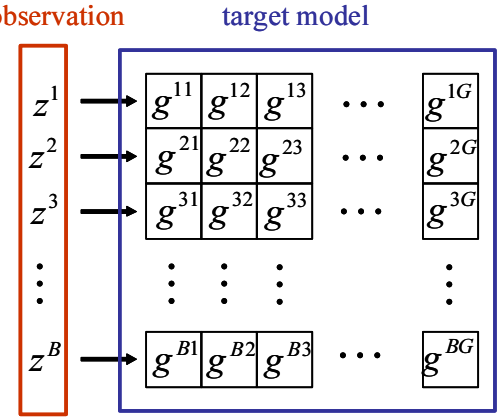


Fig. 3. Model update for each sub region by using the current observations, $z^1, z^2, \cdots, z^B$. where g^{jk} means one Gaussian whose mean and covariance are μ^{jk} and Σ^{jk}, respectively.

The proposed algorithm that simultaneously tracks the target object and updates the target model can be summarized as Algorithm 1.

Algorithm 1 Tracking and Model Update using a Rao-Blackwellized Particle Filter

Input : $\left\{x^i_{t-1}, w^i_{t-1}, M^i_{1:t-2}, z^i_{t-1}\right\}_{i=1,\cdots,N}$

- FOR i = 1 : N
  - Determine $i*$ for i by re-sampling process
  - Draw $x^i_t \sim p(x_t|x^{i*}_{t-1})$
  - If $M^{i*}_{1:t-1}$ is not updated by z^{i*}_{t-1}
    - Update $M^{i*}_{1:t-1}$ using $M^{i*}_{1:t-2}$ and z^{i*}_{t-1}
  - Copy $M^i_{1:t-1}$ from $M^{i*}_{1:t-1}$
- END FOR
- FOR i = 1 : N
  - Compute orientation histogram z^i_t for x^i_t
  - Calculate weight : $w^i_t \sim p(x^i_t|Z_{1:t}, M^i_{1:t-1})$
- END FOR
- FOR i = 1 : N
  - Normalize weight: $w^i_t = \frac{w^i_t}{\Sigma_j w^j_t}$
- END FOR

Output : $\left\{x^i_t, w^i_t, M^i_{1:t-1}, z^i_t\right\}_{i=1,\cdots,N}$

VI. Experimental Result

A. Tracking Result

We tested our tracking algorithm on several challenging sequences, and some of them are publicly available, using a 2.4 GHz CPU. For all sequences the number of Gaussians for one MoG was fixed ($G = 10$) and 500 particles were used for each target object. The proposed algorithm can process about 14 frames per second when it tracks 3 objects and a total of 1500 particles are used for 3 objects, as shown in Table I.

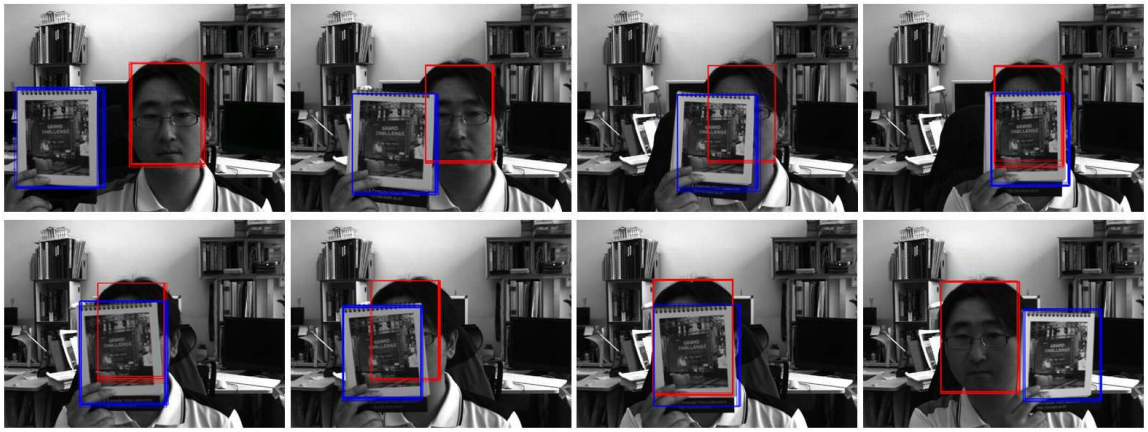


Fig. 4. The tracking results under occlusion

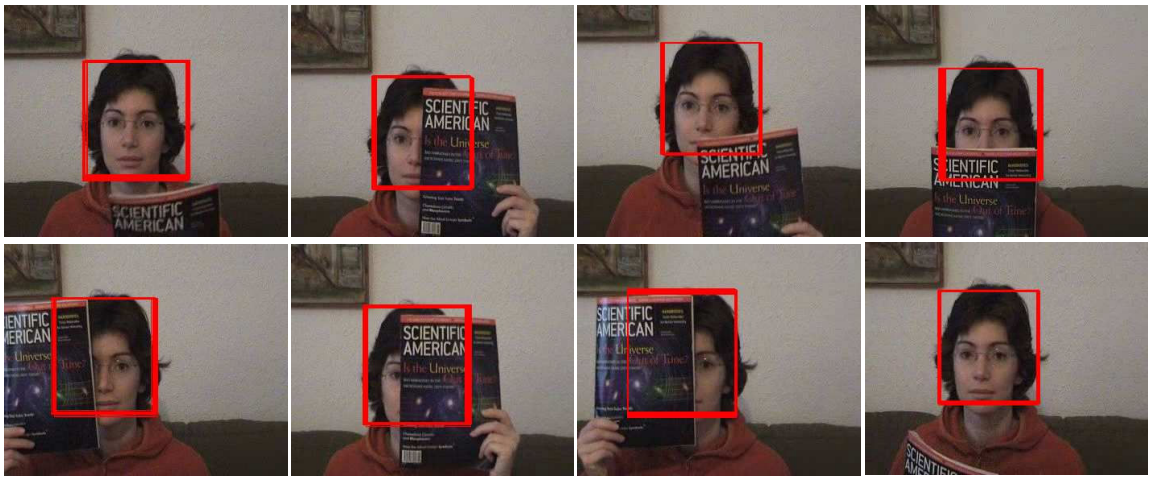


Fig. 5. The tracking results under occlusion using the face sequence used in [15](Frame 167, 216, 235, 305, 369, 568, 745, 791)

TABLE I
COMPUTATIONAL COST WHEN USING 2.4GHZ CPU AND 640 × 480 IMAGES (MS)

the number of frames	total	mean	max	min
793	56164	70.825	94	31

The first two sequences, as shown in Fig. 4 and 5, show the robustness to occlusion. The image sequence in Fig. 5 was obtained at the website of [15]. In both cases, our tracker was not strongly influenced by severe occlusion during about 300 and 800 frames in that the first result, as shown in the top-left figure is almost consistent with the last result in the bottom-right figure. In Fig. 6, we present more experimental results from two more sequences that are from the CAVIAR database [25]. The first is a sequence that includes two occluded persons by another person, and the second shows results for scale changes of the target object with partial occlusion.

B. Performance Evaluation

To analyze the performance in terms of the robustness to occlusion, deformation and random motion, we used a video sequence that has three moving people, as shown in Fig. 8. All of the people wandered randomly, and each person was occluded by the others or trees. For this sequence we manually marked the ground truth (every third frame). In Fig. 9, we plotted the position errors of our tracker and of other methods[3] which are the mean shift-based tracker [26], two methods introduced in [18], [19], and the RBPF that did not divide the region into 16 sub regions.

For human 1 who is inside the red box in Fig. 8, all other methods failed to track after occlusion with human 2 who is inside in the blue box. For human 2, three methods had the similar performance until 300 frames because he was not occluded. However, the proposed method shows the better performance in the last frame, as shown in Fig. 7 and 8. In the case of human 3 who is inside the green box in Fig. 8, some methods tracked a part of a tree instead of human 3 because the initial region partially contained the tree, thereby producing large errors for other methods, as shown in Fig. 9(c). Finally, the proposed method could successfully track

[3]For all particle filter-based approaches, we selected one particle corresponding to the maximum weight and compared it with the ground truth

(a) Tracking results of the CAVIAR database at http://homepages.inf.ed.ac.uk/rbf/CAVIARDATA1/OneLeaveShopReenter2cor.tar.gz

(b) Tracking results of the CAVIAR database at http://homepages.inf.ed.ac.uk/rbf/CAVIARDATA1/WalkByShop1cor.tar.gz

Fig. 6. The tracking results under occlusion using the CAVIAR database

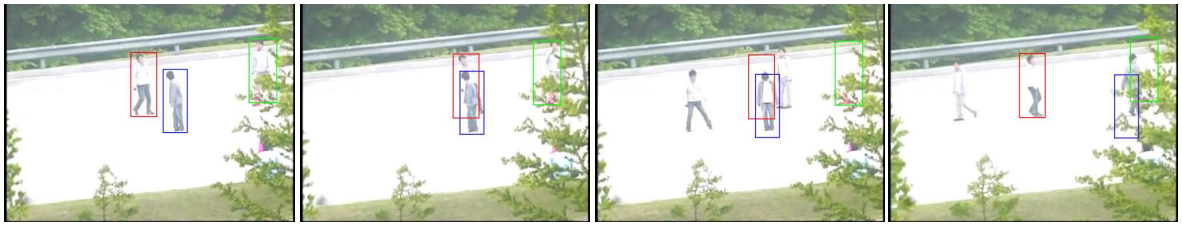


Fig. 7. The tracking results for a video sequence that has three people by using [19] (Frame 2, 40, 167, 793(last))

three people under these conditions.

VII. Conclusion

We have presented an efficient tracking algorithm that simultaneously estimates the movement of the target object and the distribution for the target model by adopting a Rao-Blackwellized particle filter. In the proposed method, the particle filter-based tracking method and the model update based on the on-line approximation of a mixture of Gaussians are combined. To ensure the robustness to occlusion, we represent the target model by 16 histograms computed from 16 sub regions and individually update the posterior distribution for each histogram through the incoming observations. We have demonstrated the robustness of the proposed method through various experiments.

VIII. Acknowledgments

This work was supported by the project for development of the surveillance system for static and moving robots funded by Samsung Techwin and partially supported by the National Strategic R&D Program for Industrial Technology, Korea.

References

[1] B. Ristic, and S. Arulampalam, and N. Gordon, Beyond the Kalman Filter – Particle Filters for Tracking Applications, Artech House Publishers, 2004

[2] A. Doucet, and N. Freitas, and N. Gordon, Sequential Monte Carlo Methods in Practice, Springer, 2001

[3] A. Doucet, and A. M. Johansen, A Tutorial on Particle Filtering and Smoothing : Fifteen years later, Technical report, Department of Statistics, University of British Columbia, 2008

[4] A. Doucet, and N. Freitas, and K. Murphy and S. Russell, Rao-blackwellized particle filtering for dynamic bayesian networks, Proceedings of the 16th Conference on Uncertainty in Artificial Intelligence, 2000

[5] D. Lowe, Distinctive image features from scale invariant Keypoints, IJCV, vol. 60, pp.91–110, 2004.

[6] S. Arulampalam, and S. Maskell, and N. Gordon and T. Clapp, A tutorial on particle filters for on-line non-linear/non-Gaussian Bayesian tracking, IEEE Trans. Signal Processing, vol. 50, pp. 174–188, 2002

[7] M. Isard and A. Blake, Contour tracking by stochastic propagation of conditional density, ECCV, 1996

[8] M. Isard and A. Blake, CONDENSATION – conditional density propagation for visual tracking, IJCV, vol. 29, pp.5–28, 1998

[9] D. Comaniciu, V. Ramesh and P. Meer, Real-Time Tracking of Non-Rigid Objects using Mean Shift, CVPR, pp.1–8, 2000.

[10] R. T. Collins, Mean Shift blob tracking through scale space, CVPR, pp.234–240, 2003.

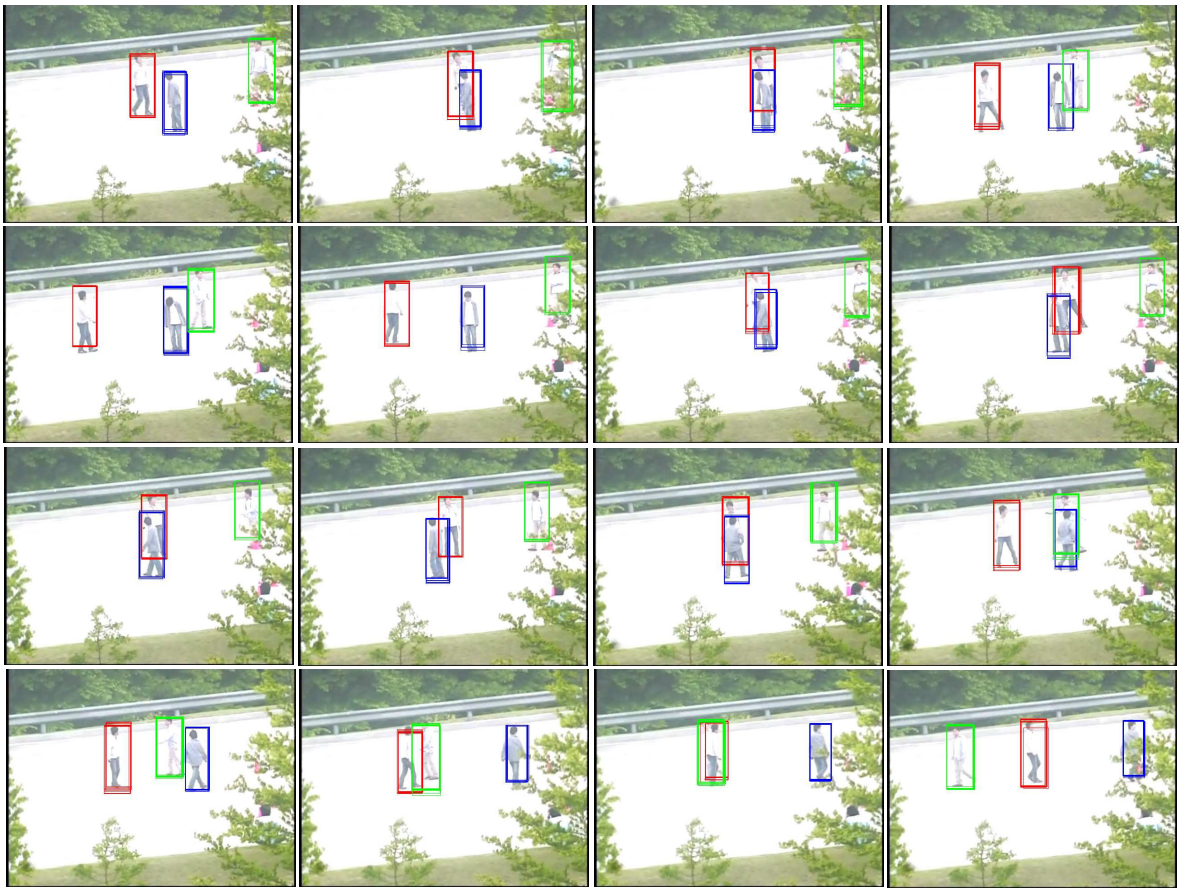
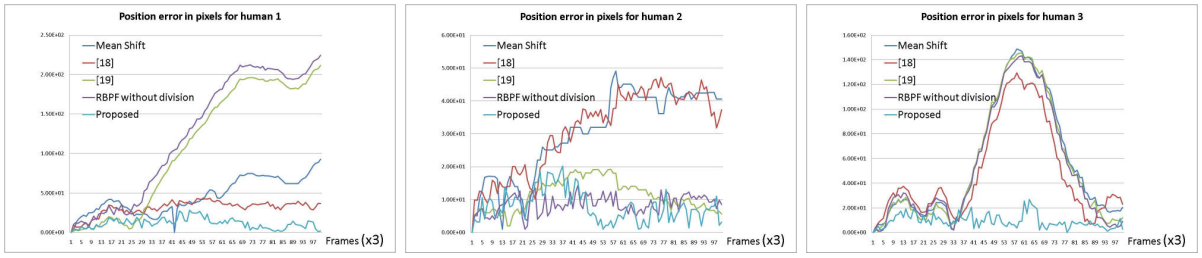


Fig. 8. The tracking results for a video sequence that has 3 persons by using the proposed method (Frame 2, 40, 55, 167, 198, 400, 409, 491, 535, 555, 645, 697, 710, 744, 756, 793(last))



(a) Tracking errors for human 1 (The person inside the red box in Figure 8)

(b) Tracking errors for human 2 (The person inside the blue box in Figure 8)

(c) Tracking errors for human 3 (The person inside the green box in Figure 8)

Fig. 9. Performance evaluation with other methods

[11] Z. Khan, T. Balch and F. Dellaert, A Rao-Blackwellized particle filter for Eigen tracking, CVPR, pp.980–986, 2004.

[12] D. Schulz, D. Fox and J. Hightower, People tracking with anonymous and id-sensors using Rao-Blackwellized particle filters, Internaitonal Joint Conference on Artificial Intelligence, 2003.

[13] X. Xu and B. Li, Adaptive Rao-Blackwellized Particle Filter and Its Evaluation for Tracking in Surveillance, IEEE Trans. Image Processing, vol. 16, pp.838–849, 2007.

[14] T. Wang, I. Y. H. Gu, A. Backhouse and P. Shi, Face tracking using Rao-Blackwellized particle filter and pose-depdent probabilistic PCA, IEEE International Conference on Image Processing, pp.853–856, 2008.

[15] A. Adam, and E. Rivlin and I. Shimshoni, Robust Fragments-based Tracking using the Integral Histogram, CVPR, 2006

[16] Y. Wu, and T. Yu, and G. Hua, Tracking Appearances with Occlusions, CVPR, 2003

[17] D. Ross, and J. Lim and R. Lin, Incremental Learning for Robust Visual Tracking, IJCV, vol. 77, pp. 125–141, 2008

[18] S. Zhou, and R. Chellappa and B. Moghaddam, Visual tracking and recognition using appearance-adaptive models in particle filters, IEEE Trans. Image Processing, vol. 13, 2004

[19] B. Babenko, and M. Yang and S. Belongie, Visual Tracking with Online Multiple Instance Learning, CVPR, 2009

[20] F. Gustafsson, F. Gunnarsson, N. Bergman, U. Forssel, J. Jansson, R. Karlsson, and P. J. Nordlund, Particle filters for positioning, navigation and tracking, IEEE Transactions on Signal Processing, vol. 50, pp. 425–437, 2002.

[21] N. Dalal and B. Triggs, Histograms of Oriented Gradients for Human Detection, CVPR, 2005

[22] P. Viola and M. Jones, Rapid Object Detection using a Boosted

Cascade of Sample Features, CVPR, pp.511–518, 2001.
[23] H. Bay, T. Tuytelaars and L. V. Gool, SURF:Speeded Up Robust Features, ECCV, pp.404–417, 2006.
[24] C. Stauffer and W.E.L. Grimson, Adaptive background mixture models for real-time tracking, CVPR, 1999
[25] Caviar datasets available at http://homepages.inf.ed.ac.uk/rbf/CAVIARDATA1/
[26] The mean-shift-based tracker code available at http://www.cs.bilkent.edu.tr/ismaila/MUSCLE/MSTracker.htm