A Robust Multiple Cues Fusion based Bayesian Tracker

Xiaoqin Zhang , Zhiyong Liu, Hong Qiao*

Institute of Automation,
Chinese Academy of Sciences,
Beijing 100080, China
Email: zhangxiaoqinnan@163.com
{zhiyong.liu, hong.qiao}@mail.ia.ac.cn

***Abstract*— This paper presents an efficient and robust tracking algorithm based on multiple cues fusion in the Bayesian framework. This method characterizes the object to be tracked using a MOG (mixture of Gaussians) based appearance model and a chamfer-matching based shape model. A selective updating technique for the models is employed to accommodate for appearance and illumination changes. Meantime, the mean shift algorithm is embedded as the prior information into the Bayesian framework to give a heuristic prediction in the hypotheses generation process, which also alleviates the great computational load suffered by the conventional Bayesian tracker. Experimental results demonstrate that, compared with some existing works, the proposed algorithm has a better adaptability to changes of the object as well as the environments.**

***Index Terms*— appearance model, chamfer distance, Bayesian tracker, template update**

I. INTRODUCTION

Object tracking [1], [2], [3], [4], as a basic task in the mobile robot system, is to locate the specified region in the video sequences. It received significant attention due to its crucial values in robotics applications.

In literature, there exist a variety of tracking algorithms from different perspectives, such as the snakes model [1], Condensation [2], mean shift [3], appearance model [4] and so on, and these algorithms have achieved great successes in this field. However, it is still a challenge to build a mobile robot tracking system that is robust to a wide variety of conditions. In the early applications, most approaches employed in tracking algorithm are based on a single cue [2], [3], [4], which is fragile to large changes of environment. For example, as applied on the head tracking, the color based approach works only on the front face or profile, but fails as the person turns around [3]. On the other hand, the edge or shape based trackers work well on the head tracking, but they are quite sensitive to the background clutter [1]. Thus, the combination of multiple cues based tracking approaches appear to achieve more reliable results [5], [6], [7], [8], [9]. Birchfield [5] uses a color histogram and intensity gradient of the target for robust head tracking. The primary limitation of such an integration strategy is that each cue has associated with the same fixed level of confidence, meaning that each cue is assumed to possess the same reliability in each frame of video. Isard [6] employed a two-level tracking in a stochastic framework, which consists of a fixed color distribution and a contour model. However, it is often invalid to assume a fixed color distribution in the dynamic environments. In [7], Wu et al. proposed a novel learning method as an approximation to a factorial graphical model where different shape and color distributions are interacted on-line in a co-inference way. However there lacks a proper scheme for decreasing the number of particles, because the particle filters suffer the problem of curse of dimensionality in high-dimension space.

In this paper, we propose an efficient and robust visual tracking algorithm in the Bayesian framework by integrating both the appearance and shape information of the target. While maintaining a low computational complexity, the proposed algorithm performs quite robustly in dynamically changing environments. The main features of our tracking approach are summarized as follows:

1) The appearance of the target is modeled by a mixture of Gaussians, and the parameters are calculated with an on-line EM algorithm, which is similar to [4], [10]. A selective adaptation scheme for updating the appearance model is adopted to accommodate for appearance and illumination changes.

2) The shape information is modeled by a chamfer transform, and we defines a similarity measure in the same metric as appearance model. It gets a significant improvement over the match measure used in Birchfield [5], and greatly improves the efficiency, as compared with [2], [7].

3) The mean shift algorithm is embedded into the particle filter framework to give a heuristic prediction to the hypotheses generation process. This strategy avoids the large number of particles when multiple cues are integrated in this framework.

The arrangement of this paper is as follows. A brief review of Bayesian based tracking algorithms is given in Section II. The detail of our algorithm is given in Section III. Experimental results are presented in Section VI, and Section VII is devoted to conclusion and the future works.

*Corresponding author. Tel.:+86-10-6257-3736.
E-mail address: hong.qiao@ia.ac.cn.

II. REVIEW OF BAYESIAN BASED TRACKER

Among those tracking algorithm mentioned above, one popular way is to take the tracking as a on-line Bayesian inference process for estimating the unknown state s_t from a sequential observations $o_{1:t}$ contaminated by noise. A dynamic state-space form is often employed in Bayesian inference framework, which contains two important components: system model(state transition model) and observation model [11], as follows,

$$system\ model: s_t = f_t(s_{t-1}, \epsilon_t) \leftrightarrow p(s_t|s_{t-1}) \quad (1)$$

$$observation\ model: o_t = h_t(s_t, \nu_t) \leftrightarrow p(o_t|s_t) \quad (2)$$

where s_t, o_t represent system state and real observation, ϵ_t, ν_t is the system noise and observation noise, $f_t(.,.)$ characterizes the kinematics of object, and $h_t(.,.)$ models the observer. Bayesian inference process evolves the sequence of probability distributions by extracting underlying information from sequence of noisy observations. When (1)(2) reduce to linear Gaussian case, the analytic filtering solution is given by the celebrated Kalman filter, in which the sufficient statistics of mean and state-error correlation matrix are calculated and propagated. Due to the non-linear and non-Gaussian essence in the real world, the sequential Monte Carlo approach which combines the powerful Monte Carlo sampling method with Bayesian inference is used in state estimation, that is usually called particle filter [12]. The key idea of particle filtering is to approximate the posterior probability distribution by a weighted sample set $\{(s^{(n)}, \pi^{(n)})|n = 1 \cdots N\}$. Each sample consists of an element $s^{(n)}$ which represents the hypothetical state of an object and a corresponding discrete sampling probability $\pi^{(n)}$, where $\sum_{n=1}^{N} \pi^{(n)} = 1$. First, the sample set is resampled to avoid the degeneracy problem, and the new sample is propagated according to the state transition model. Then each element of the set is weighted with probability $\pi^{(n)} = p(o_t|S_t = s_t^{(n)})$, which is calculated from the observation model. Finally, the state estimate $\hat{s_t}$ can be either be the minimum mean square error (MMSE) estimate or the maximum a posterior (MAP) estimate.

Feature cues and prior information are two basic issues to be considered as the Bayesian inference is adopted for tracking. In literature, many Bayesian based tracking algorithms were proposed using different cues, including contour [2], color [13], and their fusion [6], [7]. As for the prior information, most of the existing approaches take the previous system states as the prior information [2], usually containing little information about the tracking direction, and thus involving a quite large computational load since many hypotheses need to be randomly generated to cover the target. Here we propose a new cues-fusion based Bayesian tracker characterized by: 1) the two feature cues, i.e., the appearance and shape models, are adapted during the tracking process, compared with the fixed models used in [6], and 2) the prior information is given by a mean shift iteration, which provides a very instructional tracking direction, while maintaining a low computational complexity.

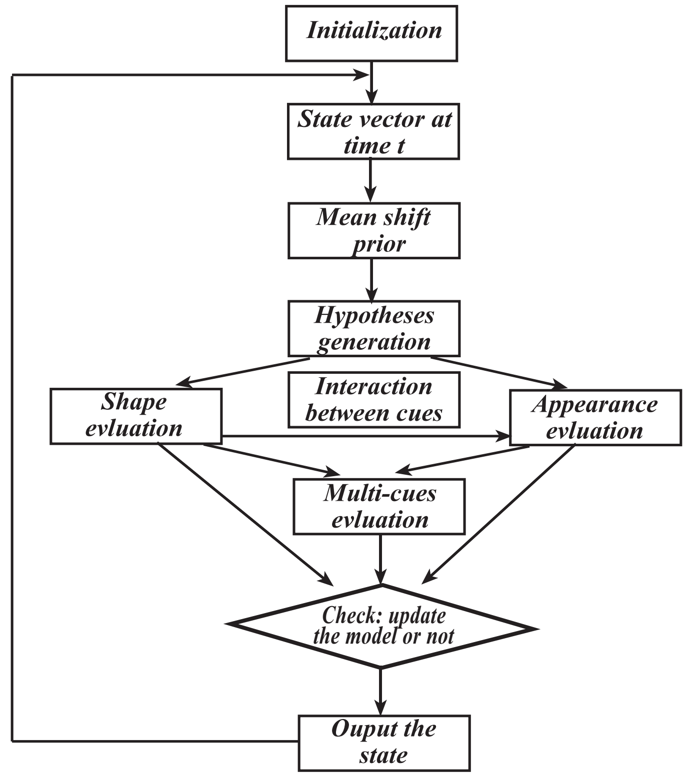


Fig. 1. The flow chart of our multiple cues integration tracking algorithm

III. TRACKING ALGORITHM BASED ON MULTIPLE CUES

The flow chart of the proposed algorithm is outlined in Fig. 1. As shown in Fig. 1, we first apply mean shift iterations to obtain a heuristic prediction for the hypotheses generation process. Each hypotheses is first evaluated by the shape model, only a small number of hypotheses are propagated and evaluated by the appearance model with highest probabilities in shape measure. A maximum a posterior (MAP) estimate of state is obtained based on combined probability of hypotheses. Meanwhile, a selective updating scheme is employed to update parameters of the appearance model and the histogram employed in the mean shift procedure to accommodate the changes of object and environment. Below we give a detailed description about each component of the algorithm, and the algorithm is summarized finally.

A. MOG Based Appearance Model

The appearance of the target is modeled efficiently by a mixture of Gaussians, with the parameters estimated by an on-line EM algorithm [4].

1) Appearance Model: Similar to [4], [10], the appearance model consists of three components S, W, F, where S component captures temporally stable images, W component characterizes the two-frame variations, F component is a fixed template of the target to prevent the model from drifting over time. Thus the likelihood function of appearance model can be formulated as follows,

$$p(o_t^a|s_t) = \prod_{j=1}^{d}\{\sum_{i=w,s,f} w_{i,t}(j)N(o_t^a(j);\mu_{i,t}(j),\sigma_{i,t}^2(j))\} \quad (3)$$

where $N(x;\mu,\sigma^2)$ is a Gaussian density

$$N(x;\mu,\sigma^2) = (2\pi\sigma^2)^{-1/2}exp\{-\frac{(x-\mu)^2}{2\sigma^2}\} \quad (4)$$

and $\{w_{i,t},\mu_{i,t},\sigma_{i,t}, i = s,w,f\}$ represent mixture probabilities, mixture centers and mixture variances respectively, and d is the number of pixels inside the target. The observation of appearance is denoted as o_t^a for short, and the same rule is applied to the observation of shape o_t^s and the observation of multi-cue combined o_t^c later.

2) Parameter Estimation: In order to make the model parameters depend more heavily on the most recent observation, we assume that the past appearance is exponentially forgotten and new information is gradually added to the appearance model. To avoid having to store all the data from previous frames, a on-line EM algorithm is used to estimate the parameters.

The on-line EM process of the parameter estimation can be described as follows:
Step1: During the E-step, the posterior probabilities of each components are computed as

$$m_{i,t}(j) \propto w_{i,t}(j)N(o_{i,t}(j);\mu_{i,t}(j),\sigma_{i,t}^2(j)) \quad (5)$$

which fulfills $\sum_{i=s,w,f} m_{i,t} = 1$.
Step2: The mixing probabilities of each components are estimated as

$$w_{i,t+1}(j) = \alpha m_{i,t}(j) + (1-\alpha)w_{i,t}(j);\ i = s,w,f \quad (6)$$

and a recursive form for moments $\{M_{i,t+1}; i = 1,2\}$ are evaluated as

$$M_{i,t+1}(j) = \alpha o_t^i(j)m_{i,t}(j) + (1-\alpha)M_{i,t}(j);\ i = 1,2 \quad (7)$$

where α acts as an adaption factor.
Step3: Finally, the mixture centers and the variances are estimated in the M-step

$$\mu_{s,t+1}(j) = \frac{M_{1,t+1}(j)}{w_{s,t+1}(j)}$$

$$\sigma_{s,t+1}^2 = \frac{M_{2,t+1}(j)}{w_{s,t+1}(j)} - \mu_{s,t+1}^2(j)$$

$$\mu_{w,t+1}(j) = o_t(j)$$

$$\sigma_{w,t+1}^2(j) = \sigma_{w,1}^2(j)$$

$$\mu_{f,t+1}(j) = \mu_{f,1}(j)$$

$$\sigma_{f,t+1}^2(j) = \sigma_{f,1}^2(j)$$

In practice, however, updating of the appearance model may be dangerous in case that, for instance, some backgrounds are misplaced into the object or the object is occluded. Thus, we developed a selective adaptation scheme to tackle such cases, which is described in subsection D.

B. Distance Transform Based Shape Matching

Object edges are also commonly used as the cue for the visual tracking [2], [5]. In our system, we employ the chamfer matching [14] based shape model as the edge cue.

1) Chamfer Matching: The core of chamfer matching is chamfer distance transform, which was first proposed by Barrow et al. [14] for object recognition and shape alignment, and was improved by Borgefors [15] who introduced a coarse-to-fine scheme to accelerate the process of chamfer matching. In literature chamfer matching has been used frequently for shape matching, by making the model template correlated with a distance transformed edge image [14].

To formalize the idea of chamfer matching in our case, the shape template—-projected model contour of ellipse is represented by a set of points $A = \{a_i\}_{i=1}^{N_a}$. The image edge map is represented as a set of feature points $B = \{b_i\}_{i=1}^{N_b}$, the chamfer transform assigns each location u the distance to the nearest feature in its neighborhood(see Fig. 2).

$$u = \min_{b\in B} ||u-b||^2 \quad (8)$$

A number of similarity functions between two point sets can be defined based on the chamfer transform. In this paper, we choose the average of the distances between each point of A and its closest point in B, which describes the degree of mismatch between point set A and B:

$$MisMatch(A,B) = \frac{1}{N_a}\sum_{a\in A}\min_{b\in B}||a-b|| \quad (9)$$

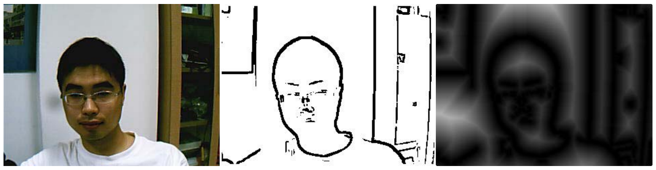
This similarity function can vary smoothly when the point locations have small changes, thus it can be tolerant to image noise and small shape variations.

2) Multi-Channel Chamfer Matching: When the chamfer matching is applied to a large amount of image clutter, the value of mismatch function is small with arbitrary shapes, therefore it is insufficient to discriminate between different templates. One way to handle this problem is to use the direction information of gradient as a complement. The idea is to use a chamfer distance measure, which not only considers distance in translation space, but also in gradient orientation space.

One possible way of doing this is as follows. The template shape A and the edge map image B are split into multiple channels according to the orientations of the image edges. Each channel contains a subset of edges from template shape and the original image respectively to a range of edge orientations. As a result, the similarity function should be evaluated on each channel of template shape and the corresponding original image respectively. Thus the ultimate $MisMatch$ between feature A and B is as follows:

$$MisMatch(A,B) = \frac{1}{N_a}\sum_{i=1}^{n}\sum_{a\in A_i}\min(\min_{b\in B_i}||a-b||^2,\tau) \quad (10)$$

where A_i and B_i are the feature points in orientation channel i, and τ is threshold to suppress noise. In our experiments, we take 4-channel split to get a balance between efficiency and



(a) (b) (c)

Fig. 2. Computing the chamfer distance. This figure shows (a) an input image, its (b) edge map and (c) the distance transformed edge map, which contains the distance to the nearest edge point at each pixel, representing by grayscale values.

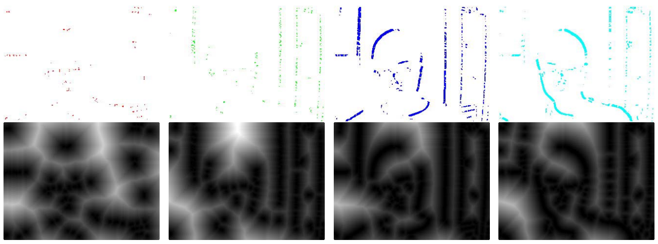


Fig. 3. Computing the multi-channel chamfer distances. The first row of figures shows the multi-channel edge map with four orientations(vertical, horizontal and two diagonal directions), the second row shows the corresponding transformed edge maps.

effectiveness, including vertical, horizontal and two diagonal directions. The 4-channel chamfer distance transform is illustrated in Fig. 3.

3) Chamfer Matching Based Similarity Measure: In order to model the shape matching process in the Bayesian framework, we define a similarity likelihood based on the $MisMatch$ degree in the same metric with appearance model.

$$p(o_t^s|s_t) \propto exp(-\frac{(MisMatch(A, B(s)) - MisMatch_{min})^2}{2\sigma_s^2}) \tag{11}$$

where $\sigma_s = \frac{1}{3}(MisMatch_{max} - MisMatch_{min})$, that can model the $MisMatch$ in the $3\sigma_s$ interval of Gaussian distribution.

4) Multiple Cues Combined Similarity Measure: The same metric shared by the appearance similarity measure and the shape similarity measure makes it easy to define the multi-cue combine similarity measure.

$$p(o_t^c|s_t) \propto p(o_t^a|s_t) \times p(o_t^s|s_t) \tag{12}$$

C. Mean Shift Based Prior

The motivation of embedding the mean shift algorithm into the particle filter framework of our tracking system is to provide a heuristic prediction to the hypotheses generation process, thus to ease the computational burden of huge number of particles.

Suppose the target is well localized at x_{t-1} in frame $t-1$(see first collum of Fig. 4), first we apply mean shift algorithm to the frame t, and the target position is predicted at $\hat{x}_t$. We integrate this information into a fist-order state transition model to form an adaptive state transition model.

$$s_t = \hat{s}_{t-1} + Affine(\hat{x}_t - x_{t-1}) + \epsilon_t \tag{13}$$

As compared with the zero-order transition model (shown in the second collum of Fig. 4), our transition model generates hypotheses more efficiently, since they are tightly centered around the object of interest so that we can easily track the object with less particles.

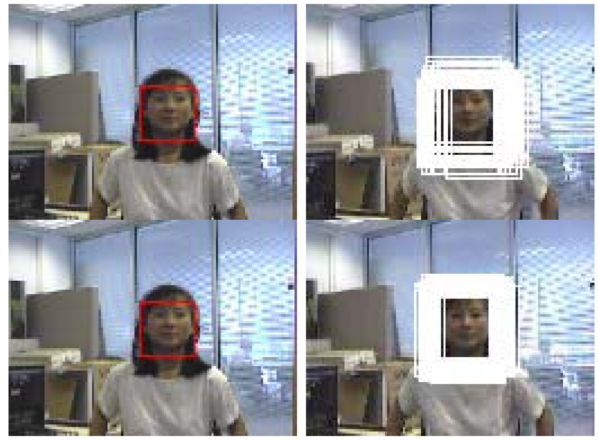


Fig. 4. Hypotheses predictions from Zero-order transition model(top row) and our transition model(bottom row)

TABLE I

SELECTIVE ADAPTATION FOR APPEARANCE MODEL

Updating Strategy
Suppose $\{\pi_a, \pi_s, \pi_c\}$ represent posterior probabilities of appearance measure, shape measure, and the combined feature measure respectively of the estimated state, and $\{T_a, T_s, T_c\}$ represent three thresholds correspondingly;
if $(\pi_a > T_a)\&\&(\pi_s > T_s)\&\&(\pi_c > T_c)$
Update the appearance model and histogram of the target;
else
Keep the appearance model and histogram of the target
end if

D. Selective Adaptation for Appearance Model

The underlying assumption behind tracking algorithm is that the feature of object remains the same between two consecutive frames, which is generally reasonable for a short period of time. However, the model of object will be gradually contaminated due to dynamic changes of appearance and environment, which makes the model inaccurate for tracking in a long run. Thus it has been a important issue to design a proper adaptation scheme for the object model [16]. On the other hand, however, over updating of the models may includes the noise of background into the object model. Thus, a proper updating scheme is of significant importance for the system.

We propose a selective updating scheme based on the confidences of both appearance and shape estimation. This strategy together with the F component in the mixture models can effectively prevent the model from drifting away, and accurately accommodate to the stochastic factors. In our algorithm, the appearance model and the histogram employed in the mean shift procedure are updated if all of three measures i.e., appearance measure, shape measure and the combined measure of the estimated state, are greater than the predefined thresholds, as shown in Table I. Evaluating the updating process on each three similar measure can exclude the wrong cases, when there is only a high similar measure in just one visual cue. For example, a object possessing a high similarity with the shape template but a low similarity with the appearance model will be considered as noise, thus it will not be updated to our model.

E. Summary of Tracking Algorithm

A brief summary of the multiple cues integration tracking algorithm is described as follows.

Algorithm 1 Tracking Algorithm Based On Multiple Cues

Note: π_a, π_s, π_c represent the posterior probability of appearance, shape, and the combined feature respectively at frame t, the histogram and the centroid of the target are denoted by $Hist, x_t$;

Input: Given the available state information $\{\hat{s}_t, x_t, \pi_a, \pi_s, \pi_c, Hist\}$ of frame t, and the observation information $\{o^a_{t+1}, o^s_{t+1}, o^c_{t+1}\}$ of frame $t+1$;

1. Apply mean shift tracking algorithm with fixed kernel bandwidth to the observations of frame $t+1$ to obtain the perdition position of the target's center $\hat{x}_{t+1}$;
2. Generate the hypotheses based on the adaptive transition model:
$s^{(n)}_{t+1} = \hat{s}_t + Affine(\hat{x}_{t+1} - x_t) + \epsilon_{t+1}, n = 1 \cdots N$;
3. Evaluate the hypotheses by the shape measure
$\pi^{(n)}_s = p(o^s_{t+1}|S_{t+1} = s^{(n)}_{t+1}), n = 1 \cdots N$;
Sort $\{\pi^{(n)}_s\}$ in a descending way;
4. The first M sorted hypotheses are evaluated by the appearance model.
$\pi^{(n)}_a = p(o^a_{t+1}|S_{t+1} = s^{(n)}_{t+1}), n = 1 \cdots M$;
The combined measure can be obtained
$\pi^{(n)}_c = \pi^{(n)}_s \times \pi^{(n)}_s$;
5. Maximum a posterior (MAP) estimate of the state
$\hat{s}_{t+1} = \arg\max_{s_{t+1}} p(s_{t+1}|o_{1:t+1}) \approx \arg\max_{s_{t+1}} \pi^{(n)}_c$;
6. Check the three similar measure to make a decision: update the model or not ;

Output: MAP estimation: $\hat{s}_{t+1}$;

IV. EXPERIMENTAL RESULTS

In our experiment, we choose the affine transformations to model the target region, and the proposed algorithm is tested on several different scenes and tracking tasks including out of plane rotation, large illumination changes, agile motions and partial occlusion, as shown in Fig. 5. All of the experiments are carried out on a one-processor Pentium IV 3.0GHz PC with 256M memory, and reach the realtime requirements.

From experimental results shown in Fig. 5(a), we can see that the selective updating scheme can adapt to the the illumination changes. Fig. 5(b) shows the result of our algorithm to track a girl's head with out-plane rotation, from which we can notice that both features are complementary in different condition. A tracking result of object with agile motions is shown in Fig. 5(c), it demonstrates that the algorithm has the ability to track the video sequences where large movements exist between two successive frames. The last sequence gives a good example to handle the partial occlusion, the mechanism of selective updating the appearance gives a efficient way to handle the partial occlusion. When occlusion happens, the updating procedure stops because the visual features of the man is different from our target(the girl), it keeps the right model and the target can be tracked well by the multiple cues. In order to obtain a tracking performance above with the girl's sequences, the number of hypothesis possessed by our transition model is present in Fig. 6, of which the average number is only 62.99, while the traditional Bayestian based tracker with zero-order transition model achieves comparable results with 400 hypotheses.

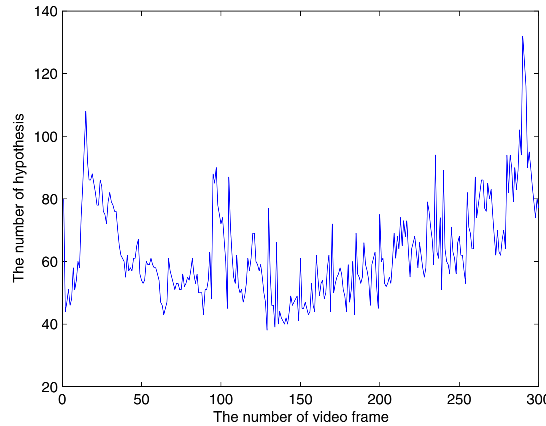


Fig. 6. The number of hypothesis needed to track the object with our transition model in each video frame

Although human head is a major tracking task in the experiments, our algorithm, compare with some exist works [5], [7], can flexibly track objects with arbitrary shapes given the template model. Furthermore, it is easy to extend this framework to multiple object tracking.

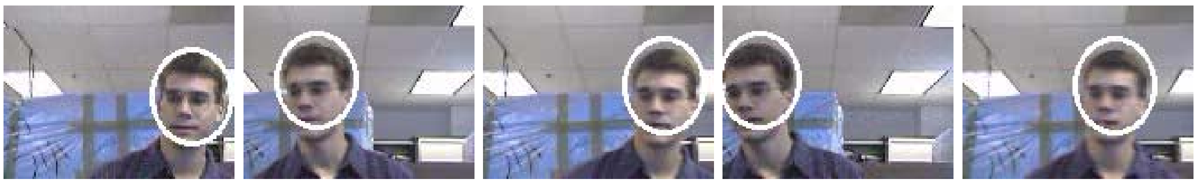
V. CONCLUSIONS AND FUTURE WORKS

A. Conclusions

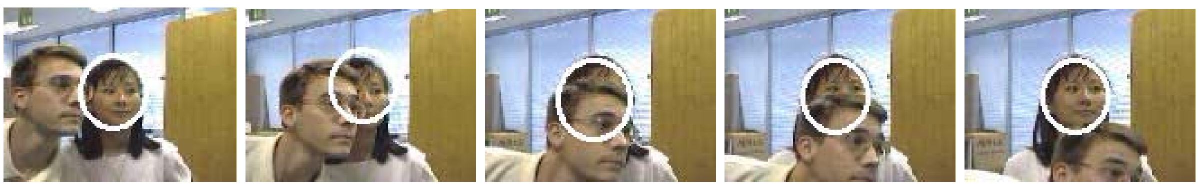
This paper presented a robust multiple cues based tracking algorithm in the Bayesian framework. In our implementation, a MOG(mixture of Gaussian) based appearance model and distance transform based shape model is employed to form a robust multi-cue fusion tracker. Our approach combine the merits of both stochastic and deterministic tracking approaches in a unified Bayesian framework: the mean-shift algorithm is embedded into the particle filter framework seamlessly to give a prior to the hypotheses generation process, which significantly decrease the particle numbers. Moreover, a selective updating scheme is employed to accommodate the changes of appearance and illumination.

(a) Scene with large illumination changes

(b) Object with out-plane rotation



(c) Object with agile motions



(d) Partial occlusion

Fig. 5. Experimental results in different scenes (illumination change, out plane rotation, agile motion and partial occlusion).

B. Future Works

Future work will address to the more elegant framework for multi-cue integration, and to integrate the multi-cue into just one model, in which the visual cues will be self-modified to the change of conditions. Furthermore, the tracking system will be enhanced by the new visual cues, such as motion, or some 3D information.

ACKNOWLEDGMENT

This work was supported in part by the Chinese Academy of Science Selective Support of 100 Outstanding Scholarship No.60675039, in part by the NSFC (National Natural Science Foundation of China) No.60505003, and in part by the MOST (The Ministry of Science and Technology of the Peoples Republic of China) No.2006AA04Z217.

Some test video streams are downloaded from the following websites: http://vision.stanford.edu/ birch/headtracker/ and http://www.cs.bu.edu/groups/ivc/HeadTracking/. Special acknowledgment is devoted to the respective owner of the video clips.

REFERENCES

[1] M.Kass, A.Witkin, D.Terzopoulos, "Snakes: active contour models," *International Journal of Computer Vision*, 1(4): 321-332, 1988.

[2] M.Isard, A.Blake, "Condensation: conditional density propagation for visual tracking," *International Journal of Computer Vision*, 29(1):5-28, 1998.

[3] D.Comaniciu, V.Ramesh, and P.Meer. "Real-time tracking of non-rigid objects using mean shift," *IEEE Conference on Computer Vision and Pattern Recognition*, pages 142-151, 2000.

[4] A.D.Jepson, D.J.Fleet, and T.F.El-Maraghi, "Robust online appearance models for visual tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 25, No. 10, pp. 1296-1311, 2003.

[5] S.Birchfield, "Elliptical head tracking using intensity gradients and color histograms," *IEEE Conference on Computer Vision and Pattern Recognition*, 232-237, 1998.

[6] M.Isard, A.Blake, "ICondensation: unifying low-level and high-level tracking in a stochastic framework," *In Proceedings of the European Conference on Computer Vision*, volume 1, pages 893-908, 1998.

[7] Y.Wu, T.Huang, "Robust Visual Tracking by Integrating Multiple Cues Based on Co-Inference Learning", *International Journal of Computer Vision*, vol.58, No.1, pp.55-71, June, 2004.

[8] G.Loy, L.Fletcher, N.Apostoloff and A.Zelinsky, "An Adaptive Fusion Architecture for Target Tracking", *Proc. of FG02*, 261-265, 2002

[9] A.Soto and P.Khosla, "Probabilistic Adaptive Agent Based System for Dynamic State Estimation Using Multiple Visual Cues", *ISRR*, 2001.

[10] S.Zhou, R.Chellappa, B.Moghaddam, "Visual Tracking and Recongnition Using Appearance-adaptive Models in Particles Filters," *IEEE Transaction on Image Processing*, 11: 1434-1456, 2004.

[11] M.Arulampalam, S.Maskell, N.Gordon, and T.Clapp, "A tutorial on particles filters for online nonlinear/non-gaussian bayesian tracking," *IEEE Trans. Signal Processing*, vol. 50, pp 174–188, Feb. 2002.

[12] Z.Chen, "Bayesian Filtering: From Kalman Filters to Particle Filters, and Beyond", *Technical report*, 2003.

[13] K.Nummiaro, E.Koller-Meierand, L.Van Gool, "An Adaptive Color-Based Particle Filter", *Image and Vision Computing*, 21, 2003, 99-110.

[14] H.Barrow, J.Tenenbaum, R.Bolles, and H.Wolf, "Parametric correspondence and chamfer matching: Two new techniques for image matching", *In Proc. 5th Int. Joint Conf. Artificial Intelligence*, 659-663, 1977.

[15] G.Borgefors, "Hierarchical chamfer matching: A parametric edge matching algorithm", *IEEE Trans. Pattern Analysis and Machine Intell*, 10(6):849-865, November 1988.

[16] I.Matthews, T.Ishikawa and S.Baker, "The Template Update Problem," *IEEE Trans. Pattern Analysis and Machine Intell*, vol. 26, no. 4, pp. 810–815, 2004.