Facial Features Tracking Using Auxiliary Particle Filtering and Observation Model Based on Bhattacharyya Distance

Fatemeh Shirinzadeh, Hadi Seyedarabi, and Ali Aghagolzadeh

***Abstract*—Recently, particle filtering has become an effective algorithm for facial feature tracking. One problem with particle filtering is that as the dimensionality of the state space increases, a large number of particles that are propagated from the previous time are wasted in area where they have low observation probability, hence a very large number of particles are necessary to track the state and, as a result the complexity increases and the speed of algorithm reduces. In this research, auxiliary particle filter with factorized likelihoods is used in order to overcome this problem. In a tracking approach, the estimated state is updated by incorporating the new observations. Therefore an observation model is needed. In this paper a novel color-based observation model that is invariant to changes in illumination intensity is proposed. The proposed observation model employs the Bhattacharyya distance to update a prior distribution calculated by the particle filter. In this paper experimentally is showed that the proposed algorithm clearly outperforms multiple independent template tracking.**

***Index Terms*—Particle filter, facial feature tracking, bhattacharyya distance, color-based observation model.**

I. Introduction

Facial feature tracking is crucial in computer vision applications and is more challenging task than face tracking. It seems the time has come for facial feature tracking applications such as gaze detection, teleconferencing, surveillance, model-based coding, Human-Computer Interaction (HCI) and Facial animation. The facial feature points are the prominent landmarks surrounding facial components: eyebrow, eye, nose and mouth. Since there are many non-rigid motions in facial features, Particle filter works better than Kalman filter and recently, particle filtering has become an effective algorithm for facial feature tracking [1]. In this algorithm, a probabilistic model of state of the object is applied to analyze an image sequence [2], [3].

A posterior density $p(x \mid Z)$ can be defined over the object's state, parameterized by a vector x, given noisy observation $Z=\{..., \bar{z}, z\}$ up to the current time instance.

So the distribution $p(x \mid Z)$ is represented by a set of pairs $\{(s_k, \pi_k)\}$ which contains K particles $s_1, s_2, .., s_k$ and their related weights $\pi_1, \pi_2, ..., \pi_k$.

Manuscript received June 18, 2012; revised August 5, 2012.

Fatemeh Shirinzadeh and Hadi Seyedarabi are with the Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran (e-mail: marjan_shirinzadeh@yahoo.com, seyedarabi@tabrizu.ac.ir).

Ali Aghagolzadeh was with the Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz, Iran. He is now with the Faculty of Electrical and Computer Engineering, Babol Nooshirvani University of Technology, Babol, Iran (e-mail: aghagol@nit.ac.ir).

Particle filter algorithm updates the posterior density recursively from pervious step:

$$\begin{aligned} p(x \mid Z) &\propto p(z \mid x) p(x \mid Z^-) \\ p(x \mid Z) &= \sum_{x^-} p(x \mid x^-) p(x^- \mid Z^-) \end{aligned} \quad (1)$$

Particle filter includes three steps:

1) Sample from $p(x^- \mid Z^-)$, where x^- is the state at previous time instant.
2) Propagate samples via the transition probability $p(x \mid x^-)$.
3) Assign new weight for the samples from the likelihood $p(z \mid x)$, (i.e. $\pi_k = p(z \mid x = s_k)$).

The new set of pairs $\{(s_k, \pi_k)\}$ represents the posterior probability $p(x \mid Z)$.

One problem whit this algorithm is that as the dimensionality of the state space increases, a large number of particles that are propagated from the previous time are wasted in area where they have low observation probability, hence a very large number of particles are necessary to track the state and, as a result the complexity increases and the speed of algorithm reduces [4], [5].

In this research, auxiliary particle filter with factorized likelihoods is used in order to remove this effect. The method is one of the improvements of particle filtering.

The method is explained in the next section and, it is described how it is applied to the template-based tracking of multiple facial features and, finally the experimental result present.

II. Auxiliary Particle Filter with Factorized Likelihoods

In auxiliary particle filter with factorized likelihoods, the state of x is represented in groups of random variable x_i, such that $x = \{x_i\}, i = 1,2,..., N$.

Each x_i represents the state of a facial feature and N is the number of features. In this method, the interdependencies between the different states are ignored, so each partition is propagated and evaluated independently:

$$p(x_i \mid Z) = p(z \mid x_i) \sum_{x^-} p(x_i \mid x^-) p(x^- \mid Z^-) \quad (2)$$

Hence $p(x_i \mid Z)$ is represented by a set of sub-particles and x_i, the samples $s_k = \langle s_{1k}, s_{2k}, ..., s_{Nk} \rangle$ are constructed. The procedure, which repeated for each partition *i*, is summarized their weights $\{(s_{ik}, \pi_{ik})\}$, after propagating the set, a proposal function is made from individual posteriors:

$$g(x) = \prod_i p(x_i \mid z) \tag{3}$$

As foregoing, by ignoring the interdependencies between different as below:

1) Propagate all particles s_k^- via the transition probability $p(x \mid x^-)$ in order to arrive at a collection of K sub-particles μ_{ik}.
2) Evaluate the likelihood associated with each sub-particle μ_{ik}, that is $\lambda_{ik} = p(z \mid \mu_{ik})$.
3) Sample K particles from the collection $\{(s_k^-, \lambda_{ik}\pi_{ik}^-)\}$. In this way it favors particles with high λ_{ik}, i.e. particles which end up at area with high likelihood when propagated with the transition probability.
4) Propagate each chosen particle s_k^- via the transition probability $p(x_i \mid x^-)$ in order to arrive at a collection of K particles s_{ik}^-.

Assign a weight π_{ik} to each sub-particle as follow:

$$\begin{aligned} w_{ik} &= \frac{p(z \mid s_{ik})}{\lambda_{ik}} \\ \pi_{ik} &= \frac{w_{ik}}{\sum_j w_{ij}} \end{aligned} \tag{4}$$

After this procedure, N posteriors are obtained. Then, sampling K particles from the proposal function $g(x)$ is approximately equivalent with constructing the particles by sampling independently from $s_k = \langle s_{1k}, s_{2k}, ..., s_{Nk} \rangle$. Finally, in order to obtain the total posterior, a weight is assigned to each particle [3]:

$$\pi_k = \frac{p(s_k \mid Z^-)}{\prod_i p(s_{ik} \mid Z^-)} \tag{5}$$

In general case, the above equation cannot be evaluated by an appropriate model, the weights need to be estimated. Use of prior information is utilized. After normalizing the sum to one, a collection $\{(s_k, \pi_k)\}$ is obtained as the particle-based presentation of $p(x \mid Z)$.

III. Tracking Facial Features

This research uses proposed particle filter for tracking the 2-D position of N facial micro-features in image sequences. At the first frame, the positions of six feature points are selected manually (Fig. 1).

The state x, $x = \langle x_1, ..., x_i, ..., x_N \rangle$ is a 2N dimensional random variable.

After describing the particle filter, the observation model and the transition probability models must be defined which are explained in the sections *A* and *B*.

Fig. 1. Selected features in the first frame.

A. Observation Model

After the 2D particles are propagated, the weight of each sub-particle needs to be determined. This is done by evaluating the observation likelihood $p(z \mid x_i)$. In a tracking approach, the estimated state is updated at each time step by incorporating the new observations. Therefore, a similarity measure is needed, which is based on color distributions [6]. A popular measure between two distributions is the Bhattacharyya coefficient.

Observation model is initialized at the first frame of the sequence when a set of *N* windows are centered by the user around the facial features that will be tracked. If o_i denote template feature vector, which contains the RGB color information at window *i*. The template feature vector that contains the RGB color information at the window around x_i be q_i. The dimension of each window is $L \times L$ and the Bhattacharyya coefficient is defined by [7], [8]:

$$\rho(o_i, q_i) = \sum_{j=1}^{L^2} \sqrt{\frac{o_i}{E\{o_i\}} \cdot \frac{q_i}{E\{q_i\}}} \tag{6}$$

where $E\{.\}$ is the average intensity of the color template. The larger ρ is, the more similar the templates are. The distance between two windows defined by Bhattacharyya distance [8], [9]:

$$d = \sqrt{1 - \rho(o_i, q_i)} \tag{7}$$

Small Bhattacharyya distances correspond to large weights so the proposed color-based observation model is defined:

$$p(z \mid x_i) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{d^2}{2\sigma^2}} \tag{8}$$

And σ is the parameter of the model.

B. Transition Model

To simplify the evaluation of the transition density, it is assumed that $p(x_i \mid x^-) = p(x_i \mid x_i^-)$, which means that each feature can be propagated individually. A very simple zero order model with Gaussian noise is used, hence:

$$p\left(x_i \mid x_i^-\right) = x_i^- + N(0,\sigma) \tag{9}$$

IV. Experimental Results

The proposed method was applied to Cohn-Kanade data base [8]. Selected facial features are the mouth corners, the upper and lower lip and the corners of eyes. A set of 50 particles are used in all experiments. The results for tracking six facial features are shown in Fig. 3 and Fig. 4.

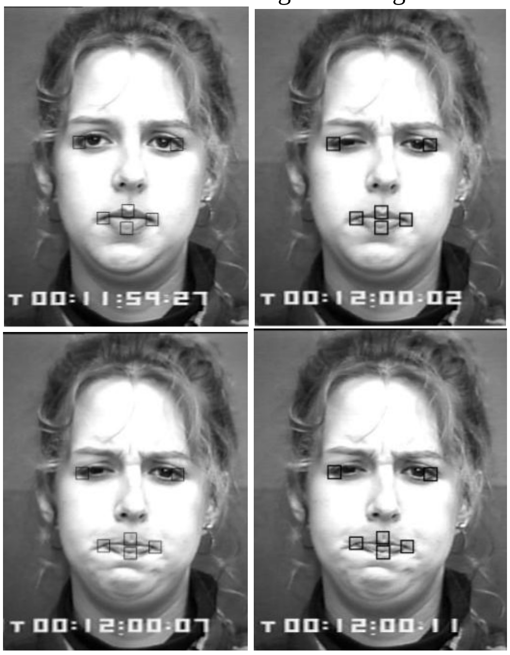

Fig. 2. Proposed method whit 50 particles, Results are presented for frames 5, 10, 15 and 20.

Fig. 3. Results are presented for frames 7, 9, 11 and 22.

In comparison, the proposed method clearly outperforms the independent feature tracker and the modified auxiliary particle filtering scheme. The use of the proposed observation model provides a robust tracking of the templates at the presence of appearance changes due to facial expressions. Furthermore, the tracker can successfully recover from long term occlusions. This is achieved with only 50 samples in total, which is an order of magnitude less than the number of particles that have been used for tracking each template independently.

V. Conclusion

Since there are many non-rigid motions in facial features, Particle filter works better than Kalman filter. One problem with particle filtering is that as the dimensionality of the state space increases, a large number of particles that are propagated from the previous time are wasted in area where they have low observation probability, hence a very large number of particles are necessary to track the state and, as a result the complexity increases and the speed of algorithm reduces. In this research, auxiliary particle filter with factorized likelihoods is used in order to overcome this problem. In addition a Color-based observation model is used which is invariant to changes in illumination intensity [9]. The proposed observation model employs the Bhattacharyya distance to update a prior distribution calculated by the particle filter. In this research experimentally is showed that the proposed algorithm clearly outperforms multiple independent template tracking.

References

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

[2] M. Pitt and N. Shephard, “Filtering via simulation: Auxiliary particle filters,” *J. American Statistical Association*, vol. 94, 1999, pp. 590–599.

[3] I. Patras and M. Pantic, “Particle filtering with factorized likelihoods for tracking facial features,” in *Proc. 6th IEEE Intl. Conf. on Automatic Face and Gesture Recognition*, 2004, pp. 1-6.

[4] J. Deutscher, A. Davison, and I. Reid. “Automatic partitioning of high dimensional search spaces associated with articulated body motion capture,” *Int. Conf. Computer Vision and Pattern Recognition*, 2001.

[5] C. Sminchisescu and B. Triggs. “Covariance scaled sampling for monocular 3d body tracking,” in *Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition*, pp. 447–454, 2001.

[6] K. Nummiaro, E. Koller-Meier, and L. V. Gool, “An Adaptive Color-based Particle Filter,” *Image and Vision Computing, Elsevier Science B.V.*, 2003.

[7] A. F. M. Smith and A. E. Gelfand, “Bayesian statistics without tears: A sampling-resampling perspective,” *J. American Statistical Association*, vol. 46, 1992.

[8] T. Kailath, “The divergence and bhattacharyya distance measures in signal selection,” *IEEE Trans. on Communication Technology*, vol. 15, 1967, pp. 52–60.

[9] F. Aherne, N. Thacker, and P. Rockett, “The Bhattacharyya Metric as an Absolute Similarity Measure for Frequency Coded Data,” *TIPR’97*, 1997.

[10] T. Kanade, J. Cohn, and Y. Tian, “Comprehensive Database for Facial Expression Analysis,” 2000.

Fatemeh Shirinzadeh received the B.S. degree from University of Guilan, Rasht, Iran, in 2006 and the M.S. degree from University of Tabriz, Iran in 2009 in Electrical Engineering. She is currently working as an instructor in Ahrar University, Rasht, Iran. Her research interests include image processing, computer vision, pattern recognition and target tracking.

Hadi Seyedarabi received B.S. Degree from University of Tabriz, Iran, in 1993 and M.S. degree in Telecommunication engineering from K.N.T. University of technology, Tehran, Iran in 1996 and Ph.D. degree from University of Tabriz, Iran, in 2006 all in Electrical Engineering. He is currently an assistant professor of Faculty of Electrical and Computer Engineering in University of Tabriz, Tabriz, Iran. His research interests are image processing, computer vision, Human-Computer Interaction, facial expression recognition and facial animation

Ali Aghagolzadeh received B.S. degree from University of Tabriz, Tabriz, Iran, in 1985 and M.Sc. degree from Illinois Institute of Technology, Chicago, IL, USA, in 1988 and Ph.D. degree from Purdue University, West Lafayette, IN, USA, in 1991 all in Electrical Engineering. He is currently a professor of Faculty of Electrical and Computer Engineering in Babol Nooshirvani University of Technology, Babol, Iran. His research interests are Digital Signal Processing, Digital Image Processing, Computer Vision and Human - Computer Interaction. He has published more than 100 journal and conference papers on pattern recognition, facial expression recognition, image fusion, video Compression and image and video coding.