2010 Second International Conference on Computer Modeling and Simulation

Occlusion Handling Based on Particle Filter in Surveillance System

Xinting Pan*, Xiaobo Chen, Aidong Men
Multimedia Technology Center
Beijing University of Posts and Telecommunications
Beijing, China
*makeawish0283@gmail.com, cchenxiaobo@gmail.com, menad@bupt.edu.cn

***Abstract*—Object tracking with occlusion handling is a challenging problem in intelligent video surveillance system. Among various tracking algorithms, particle filter (PF) is a robust and accurate one for different applications. In this paper, a new approach based on particle filter is presented for tracking object accurately and steadily when the target encountering occlusion in video sequences. First, the object pixels are classified as foreground and background for each frame using background subtraction. Our approach combines the foreground region with the particle initialization and similarity measure step to lower the background distraction. Second, a set of cues including a motion estimation model, an elliptical shape model, a spatial-color mixture of Gaussians appearance model, and an edge orientation histogram(EOH) model is fused and modeled by a data likelihood function. Then, a particle filter algorithm is used for tracking and the particles are weighted and re-sampled based on the fusion of the cues. Results from simulations and experiments with real video sequences show the effectiveness and robustness of our approach for tracking people under occlusion conditions.**

Keywords- Object tracking; occlusion handling; particle filter; data fusion

I. INTRODUCTION

Visual tracking entails the detection of objects and modeling of the features in the video sequence. Object tracking has become an important task in many computer vision applications including vehicle tracking, surveillance network, medical imaging, etc[1]. The tracking is a challenging problem due to the presence of occlusion, severe clutter and illumination changes in the scene. Most of the models in these conditions are nonlinear, non-Gaussian, multi-model or any combination of these. So maintaining the features of objects over time and using them to deal with complex interactions is a challenging task in tracking.

There are also a set of numerous algorithms for different tracking cases; these proposed algorithms can be roughly divided into two categories: deterministic way and stochastic way. Deterministic methods localize the tracked object in each frame by iteratively searching for a region which maximizes the similarity between this region and the target window. For example, Comaniciu et al. [2] employed the Mean Shift algorithm for object tracking. These methods are computationally efficient. However, these methods may converge to a local maximum. They are sensitive to background distraction, clutter, occlusions, and quick moving objects[3]. Stochastic method can handle these problems by maintaining multiple hypotheses in the state space and result in more robustness compared to the mean-shift algorithm. The particle filter, also known as sequential Monte Carlo methods, Condensation algorithm[4] and bootstrap filter[5], has been proven to be a robust algorithm to deal with the nonlinear, non-Gaussian problems. It also allows fusion of different data source or different cues of the same object due to its inherent property. Particle filter keeps tracking of the state through sample based representation of probability density functions (PDF), and construct the posterior PDF of the state space using Monte Carlo integration.

To survive in occlusion conditions, one should take advantage of multiple cues[6-8], like color, motion, edge, etc., as none of these features alone can provide universal result to different environments. With a collection of cues, the tracking system performance can be enhanced and made more robust result to the real world objects, and be able to handle occlusion, splitting, merging and other complicated events related to multiple moving objects. The color histogram is robust against the partial occlusion, but sensitive to the illumination changes in the scene or the confusing colors in the background. And it pays no attention to the local spatial correlation of color pixels, thus the color histogram is only a global measure of the color distribution. Motion cues are important to estimate the position of the targets, but it is intermittent cue and therefore cannot always be reliable. In [9] color cues are combined with motion and sound cues to provide a better result. Color and shape cues are also used in [10], where shape is described using a parameterized rectangle or ellipse. In [11] the color, shape and edge cues are combined under a particle filter framework to provide robust tracking result, it also involves an adoption scheme to select most effective cues in different conditions.

This paper focuses on developing an effective similarity measure to get an optimal likelihood

978-0-7695-3941-6/10 $26.00 © 2010 IEEE
DOI 10.1109/ICCMS.2010.74

function according to the observation probability to decrease the number of low weighted particles and increase the ones with more potential. We also propose a reinforced motion model to ensure uniform spread of particles along the image with motion estimation and thus guarantee robustness. The paper is organized as follows. Section 2 states the modeling of the cues and the likelihood used for occluded objects. Section 3 states the particle filter framework based on the fusion of cues. Section 4 investigates the performance of the proposed algorithm in different conditions. Finally, section 5 summarizes our method and presents the future work.

II. OBSERVATION MODELS

This section describes the likelihood function $p(Z|X)$, which expresses the similarity between the hypothesized configuration X of the objects and the observed image Z. The likelihood function is constructed based on the fusion of several cues. The observation models include the shape model, the motion model and the appearance likelihood models include the SMOG and EOH model.

A. Shape Model

We aim to track the position of the object when occlusion happens, without considering the pose of the object in this process. So we apply a simple elliptical model to represent the object as it contains few parameters for computation. In this paper, the object is modeled as a 2-D vertical ellipse with the parameter (x, y, h, φ), where (x, y) is the center of the ellipse, h is the length of the major axis of ellipse, and φ is the eccentricity of the ellipse.

B. Motion Model

One order motion model with constant velocity is often used in Particle Filter. The model is formulated as follows:

$$\boxed{\begin{cases} x_t = x_{t-1} + v_x + u_t \\ y_t = y_{t-1} + v_y + v_t \end{cases}} \tag{1}$$

Where v_x and v_y are both constant. u_t and v_t are assumed to be a zero-mean white Gaussian noise.

However, people usually move irregularly. Translation, rotation, occlusion of human body often occurs in actual situation. The constant velocity model is difficult to adapt to such situation, which causes the failure of tracking.

We modify the constant model with motion estimation. When the object is occluded, this model ensures the uniform spread of particles along the previous velocity vector. Then the new position is recovered from the likelihood information in the measurement update step.

Given the center of the ellipse (x, y) and the $(\vec{x}, \vec{y})$ is the respective velocity vector, we can estimate the sequential state in time t as follows

$$\boxed{\begin{cases} x_t = x_{t-1} + \vec{x} + u_t \\ y_t = y_{t-1} + \vec{y} + v_t \end{cases}} \tag{2}$$

Where u_t and v_t are zero-mean white Gaussian noise.

C. Likelihood Models

Color cues are flexible in the type of objects that they can be used to track. However, it is not discriminative to the change of illuminations and other similar colored regions. Edges are robust to the change of illuminations. But the problem is that images with heavily cluttered backgrounds can lead to a high rate of false alarm. In this part, fusion of color and edge cues is used to construct the likelihood models for particle filter, which increases the accuracy of feature matching.

1) SMOG

The particle filter based tracking algorithms usually use color models. The color histogram is robust against noise and partial occlusion. However, it considers no spatial distribution of color. We use the Spatial-Color Mixture of Gaussians (SMOG)[11] instead of the popular color histogram based model because it considers both the color feature and the spatial layout information of colors.

Assume a mixture of Gaussian with k modes in a joint spatial-color state and w_i is the normalized weight of the lth mode. Let μ_l^c and $\sum_l^c$ be the mean and the covariance of the color feature in the lth mode of the mixture of Gaussians, μ_l^s and $\sum_l^s$ be the mean and the covariance of the spatial feature.

$\phi_{l,t}^s$ and $\phi_{l,t}^c$ is defined to be the spatial match measure and color match measure between the lth mode of the original object O_τ and the current object O_ψ at time t.

$$\phi_{l,t}^S = \exp\left\{-\frac{1}{2}(\mu_{l,t}^{S,O_\psi} - \mu_{l,t}^{S,O_\tau})^T (\textstyle\sum_{l,t}^{S*})^{-1} (\mu_{l,t}^{S,O_\psi} - \mu_{l,t}^{S,O_\tau})\right\} \tag{3}$$

$$\phi_{l,t}^C = \min(w_{l,t}^{O_\psi}, w_{l,t}^{O\tau}) \tag{4}$$

Where:

$$(\textstyle\sum_{l,t}^{S*})^{-1} = (\sum_{l,t}^{S,O_\psi})^{-1} + (\sum_{l,t}^{S,O_\tau})^{-1} \tag{5}$$

In terms of this,

$$\phi_{SMOG}(O_\tau, O_\psi) = \sum_{l=1}^{k} \phi_{l,t}^{S} \phi_{l,t}^{C} \quad (6)$$

The SMOG based likelihood function is:

$$L_{SMOG} \propto \exp\{-\frac{1}{2\sigma_b^2}(1-\phi_{SMOG}(O_\tau, O_\psi))\} \quad (7)$$

σ_b is the observation variance and is set to 0.2 for experiment. See [12] for detail.

K-means algorithm is used to initialize the parameters of the mixture of Gaussians and Expectation-Maximization (EM) algorithm is used to refine them. See [11] for detail.

2) EOH Model

Edges are detected by the horizontal and vertical Sobel operators K_x and K_y. To detect edges, the color images are converted to grayscale intensity images and then converted into integral image format. And then we can get:

$$G_x(x,y) = K_x * I(x,y), G_y(x,y) = K_y * I(x,y) \quad (8)$$

Then the strength and the orientation of the edges are

$$S(x,y) = \sqrt{G_x^2(x,y) + G_y^2(x,y)} \quad (9)$$

$$\theta = \arctan(G_y(x,y)/G_x(x,y)) \quad (10)$$

The similarity between the target edge and the candidate edge can be computed by the Bhattacharyya distance as follow:

$$\phi_{Edge}(O_\tau, O_\psi) = \sqrt{1 - \sum_i \sqrt{h_i^{O_\tau} h_i^{O_\psi}}} \quad (11)$$

Where the two normalized histogram $h_i^{o_\tau}$ and $h_i^{o_\psi}$ is the target region in current frame and the reference region in the first frame. The edge based likelihood function is :

$$L_{Edge} \propto \exp\{-\frac{1}{2\sigma_b^2}(1-\phi_{Edge}(O_\tau, O_\psi))\} \quad (12)$$

We assume the SMOG and the EOH models are independent, and the joint likelihood function is

$$L(O_\tau, O_\psi) = L_{SMOG} * L_{Edge} \quad (13)$$

As shown in the experiments in section 4, this algorithm is robust during occlusion.

III. PARTICLE FILTER FRAMEWORK

The particle filter is a Bayesian sequential importance sampling technique, which recursively approximates the posterior distribution using a finite set of weighted samples. The posterior probability density function $p(x_k | Z_k)$ of the state vector x_k, given a set of $Z_k = \{z_1, ..., z_k\}$ measurement up to time k. Multiple particles of the object state are generated, and each one with a weight w_k^l, which is the specific particle l $(l = 1,2,...,N)$ in time point k. The particle filter consists of two essential steps: prediction and update. The prediction stage uses the probabilistic transition model $p(x_k | x_{k-1})$ to predict the posterior at time k as:

$$p(x_k | z_{1:k-1}) = \int p(x_k | x_{k-1}) p(x_{k-1} | z_{1:k-1}) dx_{k-1} \quad (14)$$

At time k, the observation z_k is available, the state can be updated by Bayesian rule.

$$p(x_k | z_{1:k-1}) = \frac{p(z_k | x_k) p(x_k | z_{1:k-1})}{p(z_k | z_{1:k-1})} \quad (15)$$

Where $p(z_k | x_k)$ is described by the observation equation. During prediction, each particle is modified by the state model, including the addition of random noise in order to simulate the effect of the noise on the state transition. In the update stage, each particle's weight is re-evaluated according to the new data. The candidate samples $\tilde{x}_k^i$ are drawn from an important distribution $q(x_k | x_{1:k-1}, z_{1:k})$ and the weight of the samples are:

$$\omega_k^i = \omega_{k-1}^i \frac{p(z_k | \tilde{x}_k^i) p(\tilde{x}_k^i | x_{k-1}^i)}{q(\tilde{x}_k | x_{1:k-1}, z_{1:k})} \quad (16)$$

The samples are resampled to generate a new set of particles according to their important weights to avoid degeneracy. An estimate of the measure of degeneracy at time k is given by

$$N_{eff} = \frac{1}{\sum_{l=1}^{N} (\omega_k^l)^2} \quad (17)$$

If the value of N_{eff} is below a user defined threshold, a resampling procedure can avoid degeneracy by eliminating the particles with small weights and replicating particles with larger weights.

In this paper, particle filter with fusion of cues is used as follow:

1) *Object segmentation*: The object in the scene is extracted from the background by background subtraction, and then the object is considered as the region of interest (ROI).
2) *Initialization*: Generate N samples $\{x_0^l\}$ from the prior distribution $p(x_0)$, which is assumed to be a uniform distribution. Set the initial weight $\omega_0^l = 1/N$.
3) *Prediction*: For a set of N samples, $x_{k+1}^l \sim p(x_{k+1} | x_k^l)$, which can be obtained from the motion model. Then we computer the spatial

mixture of Gaussian and the edge orientation histogram of the ROI.

4) *Measurement update*: Compute the likelihood function between the last state and current state. The likelihood function $L(z_{k+1} \mid x_{k+1}^l)$ is based on the multiple models above. Then the weight of the specific particle can be derived by

$$\omega_{k+1}^l \propto \omega_k^l L(z_{k+1} \mid x_{k+1}^l) \tag{18}$$

The normalized weight is:

$$\overline{\omega}_{k+1}^l = \frac{\omega_{k+1}^l}{\sum_{l=1}^{N} \omega_{k+1}^l} \tag{19}$$

5) *Estimation*: The state estimation $\hat{x}_{k+1}$ can be obtained from $\hat{x}_{k+1} = \sum_{l=1}^{N} \overline{\omega}_{k+1}^l x_{k+1}^l$. The effective particles can be selected according to equation (17), and in condition the N_{eff} below the threshold, then perform the resampling step.

6) *Resampling*: The particles with high important weights are duplicated and the low weights below the threshold are eliminated. This process is necessary to avoid the degeneracy.

7) Return to step 3.prediction

The particle filter process is used to estimate the optimal position of the objects based on a fusion of cues to adapt different conditions and increase the robustness of the algorithm.

IV. EXPERIMENTAL RESULTS

The above algorithm is implemented using Visual C++ on Windows XP professional platform and tested on AMD-2.22G processor with 2G RAM. The test video for this example is in the PETS 2004 database, which is an open database for research on visual surveillance. The video is available on http://homepages.inf.ed.ac.uk/rbf/CAVIARDATA1/.

A. Motion Model

One order motion model with constant velocity is often used in Particle Filter. However, it performs poorly to adapt to the change of motion, especially when people stay still or change the direction. As the comparison shown in Fig.1 and Fig.2, from frame 457 to frame 517, the person keep still for a while and then change the direction. One order motion model has some problem in tracking, but the estimation motion model has good performance.

Fig.1 Motion model without motion estimation

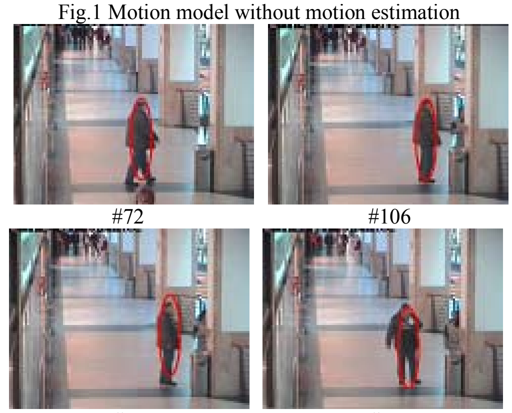
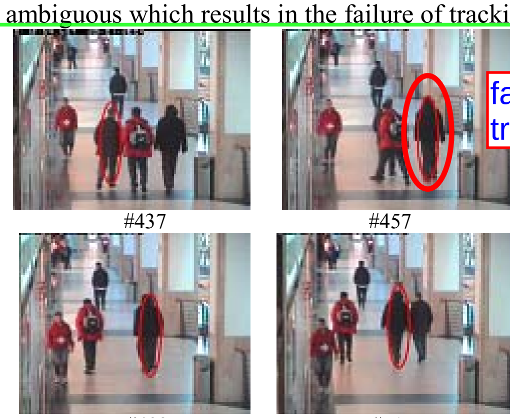


Fig.2 Motion model with motion estimation

B. Fusion of Cues

1) Occlusion handling with single cues

As is seen in this scene, the clothes color of the two people is similar. So the particle filter with only color cues is distracted by the other person. Fig. 3 shows the target person runs out of the tracking area from frame 457 to frame 493 because of the distraction. Fig. 4 shows the target person is tracked with edge cues, and we can see the edge cues suffer from the merge and split process in the scene. Because when the two people overlapped, their edges are ambiguous which results in the failure of tracking.



failure of color cue tracking

Fig.3 The target person is walking from left to right with full occlusion in frame 457 and frame 493. The object is tracked using color cues.

Fig.4 The object is tracked using edge cues.

2) Occlusion handling with fusion of cues

From the result presented in Fig.3 and Fig.4, we can see the single cue can hardly provide robust and accurate result under occlusion. We test our algorithm in the same video with the combination of the color cues and edge cues. When the person starts to walk without occlusion, color cues are more accurate and therefore can discriminate the object from the red cloth person. When the person with the same clothes cross and overlapped in the scene, the edge cues become relatively more reliable in tracking. Fig. 5 shows the tracking result during the occlusion and the object is tracked successfully.

Fig.5 The object is tracked using the fusion of color and edge cues.

V. CONCLUSION

This paper proposed an efficient and robust Particle Filter based approach for occlusion handling with fusion of cues. The motion estimation is proposed to make more accurate state estimation in particle filtering than constant velocity model. The appearance models, including SMOG and EOH model, are fused to ensure the particle filter with accuracy and robustness in tracking.

We compared this algorithm with the single cue method, which uses only color cues or edge cues. And its reliability is shown in cluttered and occlusion scenes. However, our algorithm is not effective in [illegible]vere occlusion environment such as non-rigid [illegible]ith fast rotation. Our future research area includes the adaptive data fusion schemes and improving the performance of the algorithm with less time consumption.

failure of edge cue tracking

ACKNOWLEDGMENTS

This research is supported by Special Foundation of International Science and Technology Cooperation and Exchange (20072801).

REFERENCES

[1] R. Collins, A. Lipton, and T. Kanade, "Introduction to the special section on video surveillance," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 22, no. 8, pp. 745-746, (2000).

[2] D. Comaniciu, V. Ramesh, and P. Meer, "Kernel-Based Object Tracking," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 25, no. 5, pp. 564-577, (2003).

[3] D. Comaniciu, V. Ramesh, and P. Meer, "Real-time tracking of non-rigid objects using mean shift," 2000. Los Alamitos, CA, USA: Institute of Electrical and Electronics Engineers Computer Society.

[4] M. Isard and A. Blake, " Condensation-Conditional Density Propagation for Visual Tracking," Int'l J. Computer Vision, vol. 29, no. 1, pp. 5-28, (1998).

[5] N. Gordon, D. Salmond, and A. Smith, "A novel approach to nonlinear/non-Gaussian Bayesian state estimation," IEEE Proceedings-F, vol. 140, pp. 107-113, (1993).

[6] Y. Wu and T.S. Huang, "Robust Visual Tracking by Integrating Multiple Cues Based on Co-Inference Learning, " Int'l J. Computer Vision, vol. 58, no. 1, pp. 55-71, (2004).

[7] Weiming, H., Z. Xue, et al. "Occlusion reasoning for tracking multiple people." IEEE Transactions on Circuits and Systems for Video Technology 19(1): 114-121, (2009).

[8] Meng, B., M. Zhu, et al. "Real time tracking by LOPF algorithm with mixture model", Wuhan, China, SPIE. (2007).

[9] P. Pérez, J. Vermaak, and A. Blake, "Data fusion for tracking with particles," Proceedings of the IEEE, vol. 92, no. 3, pp. 495-513, (2004).

[10] C. Shen, A. van den Hengel, and A. Dick, "Probabilistic multiple cue integration for particle filter based tracking," in Proc. of the VIIth Digital Image Computing: Techniques and Applications. C. Sun, H. Talbot, S. Ourselin, T. Adriansen, Eds., 10-12, (2003).

[11] Wang, H., et al., "Adaptive object tracking based on an effective appearance filter". IEEE Transactions on Pattern Analysis and Machine Intelligence,(2007).

[12] Lichtenauer, J., M. Reinders, and E. Hendriks. "Influence of the observation likelihood function on particle filtering performance in tracking applications."(2004). Seoul, Korea, Republic of: IEEE Computer Society.