

A novel particle filter based people tracking method through occlusion

Yuru Wang Wei Zhao Jiafeng Liu Xianglong Tang Peng Liu

School of Computer, Harbin Institute of Technology, Harbin 150001, China

Abstract

A novel multi-regions based particle filters that effectively deals with occlusion problem in people tracking has been proposed in this paper. After locating multiple key regions, the algorithm uses several nearly independent particle filters (NIPF) to track each region which will be influenced by the proximity and/or behavior of other regions. Comparing with the tracking method based on independent particle filter, the proposed algorithm is more effective in solving long-time partial or total occlusion problem. And the performance advantage the nearly independent particle filters in comparison with a joint filter is significant. This is demonstrated by the experimental results.

Keywords: Object tracking, Particle Filter, Occlusion.

1. Introduction

Recently, video tracking has received wide attention from researchers. And it becomes significantly more challenging when occlusion happens, especially in people tracking problem, which is concerned in this work. The people tracking literature contains two classical approaches to the problem of occlusion. One way is training a classifier which can differentiate the target and the noise, and the other is the object segmentation and matching scheme. Besides, Gennari[1] deals with occlusion using JPFD (Joint

Probabilistic Data Associative Filter). Zhou [2] judges the occlusion by Huber estimation and search the target by particle filter. However, these methods fail under difficult and complex conditions, and the occlusion problem is still a challenging task in people tracking.

In this paper, we address the problem of occlusion, an issue not handled by traditional methods. Our approach is inspired by the observation that, body can be divided into multiple sub-regions, thus, when occlusion happens, state of the body part being occluded can be implied by the other parts. Based on this consideration, this paper applied nearly independent particle filters (NIPF) in tracking multiple body regions to deal with the occlusion problem. The proposed algorithm is based on the well-known particle filter [3-6], which offers a degree of robustness to unpredictable motion and can correctly handle complicated, nonlinear measurement models. When dealing with multiple body regions, simply running one individual particle filter for each region is not a viable option, because the regions of the body have inter relations which play an important role in tracking multiple region. While the joint particle filter suffers from exponential complexity in the number of tracked targets, n . We have found that using n nearly independent particle filters for n regions, each running in a low-dimensional space, can approximate the optimal joint tracker by judging the occlusion based on single-target predictions from the previous time

step. The result is a set of trackers that are coupled only when necessary.

The rest of this paper is organized as follows: Section 2 presents the appearance model adopted. And the proposed nearly independent particle filters based on multi-regions is detailed in Section 3. Section 4 summaries the proposed algorithm. And experimental results are provided in Section 5. Finally, conclusions are given to highlight major merits and current limitations of the proposed approach.

2. Appearance Model

Color tracking of objects is a well known problem that has been tackled using several approaches. A color model can be represented by a histogram where each bin represents a color region. The HSV color space [7] is a very popular approach for this problem because it is relatively invariable to illumination changes. In this paper, the hue, saturation and value information are quantized as $n_h = 64$, $n_s = 8$, and $n_v = 1$. So, the resulting histogram is composed by $m = n_h \times n_s \times n_v = 512$ bins.

Let the region of the image employed to create the color model be a rectangle region centered at x_c and of size s . $\{x_i\}_{i=1,\dots,n}$ are the locations of the interior pixels of the region. We define a function $b: R^2 \rightarrow \{1..m\}$ which associates to the pixel at location x_i the index $b(x_i)$ of the histogram bin corresponding to the color u of that pixel. Then, the color density distribution for each bin $p(u)$ of the region is calculated in the following way:

$p(u) = c \sum_{i=1}^n \kappa[b(x_i) - u]$ where the parameter κ is the Kronecker delta function.

We found that the probability one pixel belongs to the object decrease when it

becomes far away from the region center. Considering this space information of the pixels, we weight the pixel less as it becomes far away from the center by the function:

$$\pi(r) = \begin{cases} 1 - r^2, & 0 < r < 1 \\ 0, & \text{others} \end{cases}$$

Where r is the ratio of the distance from the pixel to the region center and s . Thus, the color model $p = \{p(u)\}_{u=1,\dots,m}$ of the region is:

$$p(u) = \frac{1}{\sum_{i=1}^n \pi\left(\frac{\|x_c - x_i\|}{s}\right)} \sum_{i=1}^n \pi\left(\frac{\|x_c - x_i\|}{s}\right) \kappa[b(x_i) - u] \quad (1)$$

$\|x_c - x_i\|$ represents the distance from the pixel to the center. The resulting histogram is normalized as $\sum_{u=1}^m p(u) = 1$ by

$$1 / \sum_{i=1}^n \pi\left(\frac{\|x_c - x_i\|}{s}\right).$$

Once the color model p is created, it can be compared with other color model q using the Bhattacharyya coefficient [8] which has an important feature that both color models can be compared even if they have been created using regions of different sizes. In the case of a discrete distribution it can be expressed as follows:

$$\rho[p, q] = \sum_{u=1}^m \sqrt{p(u)q(u)} \quad (2)$$

The value $\rho[p, q]$ gives a similarity measure of two color models in the range $[0, 1]$, where 1 means that both color models are identical, and it decreases as they differ.

3. The proposed method

To deal with the occlusion problem in people tracking, we locate n regions on

the body, and the state of the whole body can be represented by the joint state of these regions. Thus, when occlusion happens, the region being occluded can be predicted correctly by other regions. And by this way, we can solve the occlusion problem. Our method is based on Particle filter (PF) which use multiple discrete “particles” to represent the belief distribution over the location of a tracked target. When dealing with the multiple regions tracking problem which can be looked as a multiple targets tracking problem, the independent or joint particle filters can be used. However, the body parts don't move independently, so the independent particle filters can not work, while the joint particle filter suffers from the computation complexity. So, this paper proposes a new particle filter method which runs nearly independent particle filters in body parts.

3.1. particle filter

We review particle filter to introduce notation and to facilitate understanding of the new components.

In the Bayes filtering paradigm, we recursively update the posterior distribution over the current state X_t given all observations $Z_{1:t}$ $@\{Z_1, Z_2, \dots, Z_t\}$ up to and including time t , as follows:

$$P(X_t | Z_{1:t}) = cP(Z_t | X_t) \int P(X_t | X_{t-1}) P(X_{t-1} | Z_{1:t-1}) dX_{t-1}$$

where the likelihood $P(Z_t | X_t)$ expresses the measurement model and $P(X_t | X_{t-1})$ is the motion model. In a particle filter we approximate the posterior $P(X_{t-1} | Z_{1:t-1})$ recursively as a set of weighted N samples $\{X_{t-1}^r, \omega_{t-1}^r\}_{r=1}^N$, where ω_{t-1}^r is the weight for particle. Given this, we can use a Monte Carlo approximation of the integral and get: $P(X_t | Z_{1:t}) \approx cP(Z_t | X_t)$

$$\sum_r \omega_{t-1}^r P(X_t | X_{t-1}^r) \quad (3)$$

One way to view a particle filter is as an importance sampler for this distribution. Specifically, N samples are drawn from the proposal distribution

$$X_t^s : q(X_t) @ \sum_r \omega_{t-1}^r P(X_t | X_{t-1}^r) \quad (4)$$

And then weighted by the likelihood $\omega_t^s = P(Z_t | X_t^s)$. Where r and s are the index of particle at $t-1$ and t , respectively.

3.2. Multi-Regions tracking

3.2.1 Undirected Graph based model

We describe the joint of multiple body regions at frame t as an undirected graph (V, E) , where the nodes V represent the regions and the edges E specify a neighborhood relationship. As shown in fig. 1a, $V = \{X_{ii}\}_{i=1, \dots, n}$ and $E = \{\phi(X_{ii}, X_{jj})\}$, where $\phi(X_{ii}, X_{jj})$ describes the relationship between the state X_{ii} and X_{jj} of regions i, j if the corresponding edge exists. To describe the undirected graph at frame t , we firstly construct the relationship model $R = \{R^{i,j} = (\theta^{i,j}, d^{i,j})\}_{i,j \in E}$,

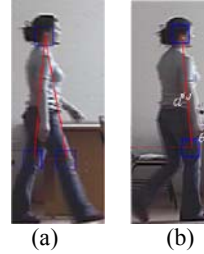


Fig.1: (a)The undirected graph,(b) The definition of ϕ .

where $d^{i,j}$ and $\theta^{i,j}$ are the distance and angle from $\overline{X_i X_j}$ to the motion direction (Fig. 1b). Then, the function ϕ is computed as follows:

$$\phi(X_{ii}, X_{jj}) = k_d f_d + k_\theta f_\theta \quad (5)$$

Where, f_d, f_θ are the distance and angle factor defined as below, and k_d, k_θ are their weights, respectively.

$$\begin{cases} f_d = 1 - \frac{|d_t^{i,j} - d^{i,j}|}{\delta_d}, \text{if } 0 \leq |d_t^{i,j} - d^{i,j}| \leq \delta_d \\ f_\theta = 1 - \frac{|\theta_t^{i,j} - \theta^{i,j}|}{\delta_\theta}, \text{if } 0 \leq |\theta_t^{i,j} - \theta^{i,j}| \leq \delta_\theta \end{cases}$$

The domain of ϕ is $[0,1]$. $|d_t^{i,j} - d^{i,j}|$ is the absolute value of the difference, and δ_d, δ_θ are used to define their permitted maximum, $d_t^{i,j}, \theta_t^{i,j}$ are the distance and angle at frame t .

3.2.2 The proposed nearly independent particle filters

To track multiple associated regions with relationships, we define the nearly independent particle filters as follows. At frame t , the posterior distribution of the state of region i factors as:

$$P(X_{it} | Z_{i(1:t)}) \approx cP(Z_{it} | X_{it}) \times \prod_{i,j \in E} \phi(X_{it}, X_{jt}) \times \sum_r \omega_{i(t-1)}^r P(X_{it} | X_{i(t-1)}^r) \quad (6)$$

we sample from the proposal distribution function

$$X_{it}^s : q(X_{it}) = \sum_r \omega_{i(t-1)}^r P(X_{it} | X_{i(t-1)}^r) \quad (7)$$

by drawing a mixture component at random, and then moving each individual target independently. Then, we weight each of the particles X_{it}^s at region i by

$$\omega_{it}^s \approx P(Z_{it} | X_{it}^s) \prod_{ij \in E} \phi(X_{it}^s, X_{jt}^s) \quad (8)$$

We can see, the trackers are not fully independent, because they consider the locations of other targets when scoring particles. In particular, according to the undirected graph (V, E) at time t , we penalize particles that fulfill the relationship model $R^{i,j}$ less.

3.2.3 The joint estimate scheme

To estimate the location of the target, instead of computing the mean of all the associated particles, we construct a region relationship based estimate scheme using

the defined undirected graph. The state of the region i at frame t is estimated as

$$\mathbb{X}_{it} = \sum_{i,j \in E} k_{jt} \sum_{s=1}^N (l_{jt}^s \times X_{jt}^s \times \phi(X_{jt}^s, X_{it}^s)) \quad (9)$$

Where, the state \mathbb{X}_{it} of region i is estimated jointly by computing sum of all the associated regions' weighted states acting on the region i , and the weight is k_{jt} (see next section). Furthermore, the associated regions' acting states is computed by all acting states $(X_{jt}^s \times \phi(X_{jt}^s, X_{it}^s), s=1, \dots, N)$ of the particles of the region X_{jt}^s weighted l_{jt}^s , where $\phi(X_{jt}^s, X_{it}^s)$ is obtained by $R^{i,j}$ (see next section for detail). As shown in formula above, the state \mathbb{X}_{it} of region i is estimated by all the associated regions' particles, thus, when occlusion happens, the state of the region being occluded can be estimated correctly by this joint estimate scheme.

4. Summary

To give a clear description of the proposed scheme, we describe the flow of the algorithm as follows.

Step 1: Initialization: At the first and second frame of the image sequence, locate the location of n regions as the state $X_{i0}, X_{i1}, i=1, \dots, n$. And construct the motion model $P(X_{it} | X_{i(t-1)})_{i=1, \dots, n}$ for each region, the appearance model $Q = \{q_i = \{q_i(u)\}_{u=1, \dots, m}\}_{i=1, \dots, n}$ and relationship model $R = \{R^{i,j} = (\theta^{i,j}, d^{i,j})\}_{i,j \in E}$.

Step 2: At each time step t , the posterior over the state of region i at time $t-1$ is represented by a set of N weighted samples $\{X_{i(t-1)}^r, \omega_{i(t-1)}^r\}_{r=1}^N$. We then create N new particles $\{X_{it}^s, \omega_{it}^s\}_{s=1}^N$ by importance sampling, using the motion model (7) as the proposal distribution

for region i . Do the follows N times iteratively.

(1) **Resample:** Pick a sample $X_{i(t-1)}^r$ with probability $\omega_{i(t-1)}^r$.

(2) **Apply the motion model:** $P(X_{it} | X_{i(t-1)}^r)$ to obtain a newly state X_{it}^s .

(3) **Score:** Assign a weight ω_{it}^s to the state X_{it}^s according to (8).

(a) We model the observation $p_{it}^s = \{p_{it}^s(u)\}_{u=1, \dots, m}$ according to formula (2). And Compute the similarity b_{it}^s which is $P(Z_{it} | X_{it}^s)$ in formula (2) between the observation p_{it}^s and model q_i .

(b) Get the distance and angle $d_i^{i,j}$, $\theta_i^{i,j}$ and compute $\phi(X_{it}^s, X_{jt}^s)$ based on the model $R^{i,j} = (\theta^{i,j}, d^{i,j})$ according to formula (5), where $i, j \in E$. The relationship function penalizes the particles that fulfill the relationship model $R^{i,j}$ less.

(c) Apply the weight updating function (8) to obtain the new weight ω_{it}^s of particle X_{it}^s .

Step 3: Estimate: Then, estimate the state \mathbf{K}_{it}^s of region i , $i = 1, \dots, n$, by the joint estimate scheme, according to (9)

(1) Compute the sum of the weight $C_j = \sum_{s=1}^N b_j^s$ of all particles for each region $j, j = 1, \dots, n$, and $C = \sum_{j=1}^n C_j$, then the

weight k_{jt} of region j in estimate the

state of region i is $k_{ij}^s = \frac{C_j}{C}$.

(2) Obtain the function $\varphi(X_{jt}^s, X_{it}^s)$ according to $R^{i,j} = (\theta^{i,j}, d^{i,j})$, $j = 1, \dots, n$.

(3) Estimate the state \mathbf{K}_{it}^s of region i according to (9).

Summary up, the particle filters employed in this paper are nearly independent because they consider the locations of other targets when scoring particles and estimating the state, while sample particles independently. By this way, the multiple regions can be tracked jointly with less computation complexity, and when some region is occluded, its state can be estimated correctly jointly by other regions.

5. Experimental results

We evaluated our approach by tracking through three kinds of video, and present quantitative results as well as a graphical comparison of the different tracker methodologies which are joint and independent particle filters. While a theoretically optimal joint filter is too computationally expensive to track multiple regions, we show that multiple nearly independent trackers generate tracks of similar quality to those generated by a joint tracker in occlusion situation, but at a substantially lower cost.

In the experiment, taking two regions as an example, we compare three methodologies in the effectiveness and computation complexity, which are independent PF, joint PF and the proposed PF, testing on three kinds of video: simulation video, CAVIAR [9] and the video captured by us, and taking two regions as example. In Fig. 2, when occlusion happens, the proposed PF can track the multiple regions successfully. Fig. 3 gives the comparison between the proposed PF with the other two methods testing on simulation video which has the feature of nonlinear and nongaussian, in which we can see that the independent PF fail in dealing with occlusion, while the joint PF and the proposed PF succeed. This is because the independent PF track each region independently, and one region's failure tracking can not be rectified by the other re-

gions as did in joint PF. Furthermore, we compute the time consume at each frame of the proposed method and joint PF in table 1, in which the proposed method spent much less time than the joint PF. As a conclusion, the proposed PF method

can track people successfully as joint particle filter with less time, which demonstrates the effectiveness of the propose method.

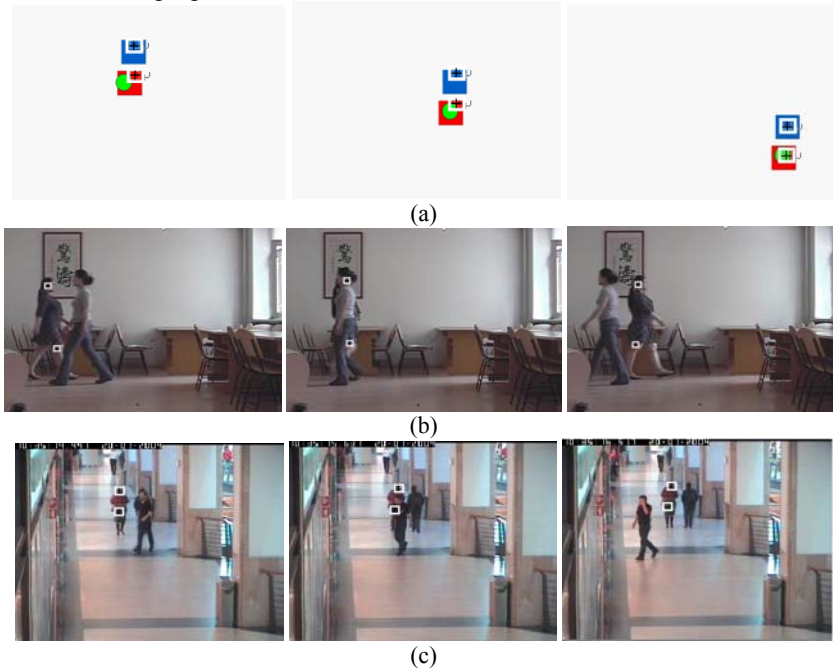


Fig.2. Tracking result of the proposed method: (a) (b) (c) are the result on three kinds of video.

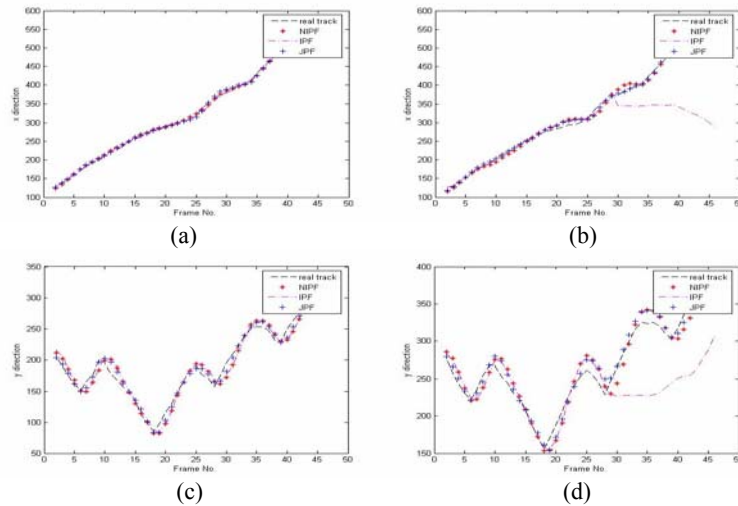


Fig. 3. The comparison between the proposed PF and the other method. (a) the tracking result at x direction in region 1, (b) the tracking result at x direction in region 2, (c) the tracking result at y direction in region 1, (d) the tracking result at y direction in region 2.

Table. 1: The time consuming of the proposed method and joint particle filters.

	Video1(s)	Video2(s)	Video3(s)
The proposed method	1.437	1.547	1.39
Joint particle filters	2.953	3.016	2.937

6. Conclusions

Through locating some key regions on the people, this paper proposed a nearly independent particle filters in tracking multiple regions to deal with the occlusion problem in people tracking. While a theoretically optimal joint filter is too computationally expensive to track multiple targets at once, we show that multiple nearly independent trackers generate tracks of similar quality to those generated by a joint tracker, but at a substantially lower cost.

However, more complex interactions are possible when conditioning on different body parts, and much more complexity undirected graph should be construct to express the interactions of multiple regions, which can be of our further research.

7. References

[1] Gennari G, Hager G.D. Probabilistic data association methods in visual tracking of groups . In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. Vol. 2, II-876 - II-881.

[2] Zhou S K, Chellappa R, Moghaddam B. Appearance tracking using adaptive models in a particle filter. In:

Proceedings of the 6th Asian Conference on Computer Vision, 2004.

[3] N. Gordon, D. Salmond, and A. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," IEE Proceedings F, vol. 140, no. 2, pp. 107–113, 1993.

[4] M. Isard and A. Blake, "Contour tracking by stochastic propagation of conditional density," in Eur. Conf. on Computer Vision (ECCV), pp. 343–356, 1996.

[5] J. Carpenter, P. Clifford, and P. Fernhead, "An improved particle filter for non-linear problems," tech. rep., Department of Statistics, University of Oxford, 1997.

[6] F. Dellaert, D. Fox, W. Burgard, and S. Thrun, "Monte Carlo Localization for mobile robots," in IEEE Int. Conf. on Robotics and Automation (ICRA), 1999.

[7] J.D. Foley and A. van Dam. Fundamentals of Interactive Computer Graphics. Addison Wesley, 1982.

[8] F. Aherne, N. Thacker, and P. Rockett. The Bhattacharyya Metric as an Absolute Similarity Measure for Frequency Coded Data. Kybernetika, 32:1–7, 1997.

[9] Fisher R, et al. CAVIAR test case scenarios. 2003.
<http://homepages.inf.ed.ac.uk/rbf/CAVIAR>.