SVD Based Kalman Particle Filter for Robust Visual Tracking

Xiaoqin Zhang[1],Weiming Hu[1],Zixiang Zhao[2],Yan-guo Wang[1],Xi Li[1],Qingdi Wei[1]
[1]National Laboratory of Pattern Recognition, Institute of Automation, Beijing, China
{xqzhang,wmhu,ygwang,lixi,qdwei}@nlpr.ia.ac.cn
[2]Beijing University of Aeronautics & Astronautics
zhaozx531@yahoo.cn

Abstract

Object tracking is one of the most important tasks in computer vision. The unscented particle filter algorithm has been extensively used to tackle this problem and achieved a great success, because it uses the UKF (unscented Kalman filter) to generate a sophisticated proposal distributions which incorporates the newest observations into the state transition distribution and thus overcomes the sample impoverishment problem suffered by the particle filter. However, UKF often encounters the ill-conditioned problem when solving the square root of the covariance matrix in practice. In this paper, we propose a novel Kalman particle filter based on SVD (singular value decomposition), and apply it for visual tracking. Experimental results demonstrate that, compared with the particle filter and the unscented particle filter, the proposed algorithm is more robust in tracking performance.

1. Introduction

Object tracking has received significant attention due to its crucial value in visual applications including surveillance, human-computer interaction, intelligent transportation, augmented reality and video compression.

The particle filter [1, 2] has been extensively studied in the tracking literature due to its effectiveness and flexibility. From a Bayesian view, particle filter is essentially a sequential Monte Carlo approach to solve the recursive Bayesian filtering problem, which combines the Monte Carlo sampling techniques with Bayesian inference. It relaxes the linearity and Gaussianity constraints of the Kalman filter and provides a tractable solution to non-linear and non-Gaussian systems. The basic idea of particle filter is to use a number of independent random variables called particles, sampled directly from a proposal distribution, to represent the posterior probability, and update the posterior by involving the new observations. Although it has achieved a considerable success in the tracking literature, it is faced with a fatal problem-sample impoverishment due to its 'suboptimal sampling' mechanism. For the conventional particle filter, the particles are directly sampled from state transition distribution. However, it is not the 'optimal' proposal sampling distribution. When the state transition distribution lies in the tail of the observation likelihood distribution, the weights of most particles are low, leading to the poor performance in practice.

Much effort has been expended to overcome this problem and improve the performance of particle filter in recent years [3, 4, 5, 6, 7, 8]. Among them, the unscented particle filter [4] is the successful one. In the unscented particle filter, the UKF based proposal distribution is introduced as follows. Firstly, a set the sigma samples are generated by UT (unscented transformation) with corresponding weights, and then are propagated through the state transition model, finally the weighted mean and covariance are further calculated to form a better proposal distribution. Compared with the EKF (extended Kalman filter) which approximates to the first-order accuracy for non-Gaussian data, the estimation accuracy of UKF is improved to at least second-order. However, UKF often encounters the ill-conditioned problem when solving the square root of the covariance matrix in practice. To overcome this problem, we propose a SVD based Kalman particle filter, where the sigma samples are generated by SVD of the eigen-covariance matrix. While maintaining the same computational complexity, the proposed tracking algorithm performs quite robustly in tracking performance.

The following paper is arranged as follows. Section 2 presents the unscented particle filtering framework. The detail of the proposed SVD based Kalman filter is described in Section 3. Section 4 introduces the incremental subspace leaning based appearance model. Experimental results are shown in Section 5, and Section 6 is devoted to conclusion.

2. Unscented Particle Filtering Framework

To make this paper self-contained, we first briefly review the particle filter and its major limitation, and then present the unscented particle filter in detail.

2.1. Particle Filter

Particle filter [2] is an online Bayesian inference process for estimating the unknown state x_t at time t from a sequential observations $y_{1:t}$ perturbed by noises. A dynamic state-space form employed in the Bayesian inference framework is shown as follows,

$$x_t = f(x_{t-1}, \epsilon_t) \leftrightarrow p(x_t|x_{t-1}) \tag{1}$$

$$y_t = h(x_t, \nu_t) \leftrightarrow p(y_t|x_t) \tag{2}$$

where x_t, y_t represent system state and observation, ϵ_t, ν_t are the system noise and observation noise. $f(.,.)$ and $h(.,.)$ are the state transition and observation models, which characterize the state transition distribution $p(x_t|x_{t-1})$ and the observation distribution $p(y_t|x_t)$ respectively. The key idea of particle filter is to approximate the posterior probability distribution $p(x_t|y_{1:t})$ by a set of weighted samples $\{x_t^i, w_t^i\}_{i=1}^N$, which are sampled from a proposal distribution $q(\cdot)$, *i.e.* $x_t^i \sim q(x_t|x_{t-1}^i, y_{1:t}), (i = 1, \cdot\cdot\cdot, N)$, and then each particle's weight is set to

$$w_t^i \propto \frac{p(y_t|x_t^i)p(x_t^i|x_{t-1}^i)}{q(x_t|x_{t-1}^i, y_{1:t})} \tag{3}$$

Finally, the posterior probability distribution is approximated as $p(x_t|y_{1:t}) = \sum_{i=1}^N w_t^i \delta(x_t - x_t^i)$, where $\delta(\cdot)$ is the Dirac function.

Doucet et al. [9] prove that the 'optimal' proposal distribution is $p(x_t|x_{t-1}^i, y_t)$ in the sense of minimizing the variance of the importance weights. So the question is, how to incorporate the current observation y_t into the transition model $p(x_t|x_{t-1})$ to form an effective proposal distribution.

2.2. Unscented Particle Filter

In order to utilize the current observations, Freitas et al. [4] propose a high-performance unscented particle filter (UPF) by using UKF to generate the proposal distribution.

In the implementation, the state space is expanded as: $x_{t-1}^a = [x_{t-1}^T \epsilon_{t-1}^T \nu_{t-1}^T]$, whose dimension and covariance matrix are $N_a = N_x + N_\epsilon + N_\nu$ and P_{t-1}^a respectively. Consider the nonlinear tracking problem modeled by the state-space equations (1) and (2), the pseudo-code of the unscented Kalman filter is presented as follows.

1. Calculate $2N_a$ sigma points as in [4]

$$\mathcal{X}_{0,t-1}^{(i)a} = \bar{x}_{t-1}^{(i)a}$$

$$\mathcal{X}_{j,t-1}^{(i)a} = [\bar{x}_{t-1}^{(i)a}\ \bar{x}_{t-1}^{(i)a} \pm \sqrt{(n_a + \lambda)P_{t-1}^a}_j]$$

$$W_0^{(m)} = \frac{\lambda}{N_a + \lambda},\ W_0^{(c)} = \frac{\lambda}{N_a + \lambda} + (1 - \alpha^2 + \beta)$$

$$W_j^{(m)} = W_j^{(c)} = \frac{1}{2(N_a + \lambda)}, \lambda = \alpha^2(N_a + \kappa) - N_a$$

$$j = 1, \ldots, 2N_a$$

2. Time update:

$$\mathcal{X}_{j,t|t-1}^{(i)x} = f(\mathcal{X}_{j,t-1}^{(i)x}, \mathcal{X}_{j,t-1}^{(i)\epsilon}), \bar{x}_{t|t-1}^{(i)} = \sum_{j=0}^{2N_a} W_j^{(m)} \mathcal{X}_{j,t|t-1}^{(i)x}$$

$$P_{t|t-1}^{(i)} = \sum_{j=0}^{2N_a} W_j^{(c)} [\mathcal{X}_{j,t|t-1}^{(i)x} - \bar{x}_{t|t-1}^{(i)}][\mathcal{X}_{j,t|t-1}^{(i)x} - \bar{x}_{t|t-1}^{(i)}]^T$$

$$\mathcal{Y}_{j,t|t-1}^{(i)} = h(\mathcal{X}_{j,t-1}^{(i)x}, \mathcal{X}_{j,t-1}^{(i)\nu}), \bar{y}_{t|t-1}^{(i)} = \sum_{j=0}^{2N_a} W_j^{(m)} \mathcal{Y}_{j,t|t-1}^{(i)}$$

3. Measurement update:

$$P_{y_t,y_t} = \sum_{j=0}^{2N_a} W_j^{(c)} [\mathcal{Y}_{j,t|t-1}^{(i)} - \bar{y}_{t|t-1}^{(i)}][\mathcal{Y}_{j,t|t-1}^{(i)} - \bar{y}_{t|t-1}^{(i)}]^T$$

$$P_{x_t,y_t} = \sum_{j=0}^{2N_a} W_j^{(c)} [\mathcal{X}_{j,t|t-1}^{(i)x} - \bar{x}_{t|t-1}^{(i)}][\mathcal{Y}_{j,t|t-1}^{(i)} - \bar{y}_{t|t-1}^{(i)}]^T$$

$$K_t = P_{x_t,y_t} P_{y_t,y_t}^{-1},\ \bar{x}_t^{(i)} = \bar{x}_{t|t-1}^{(i)} + K_t(y_t - \bar{y}_{t|t-1}^{(i)})$$

$$\bar{P}_t^{(i)} = P_{t|t-1}^{(i)} - K_t P_{y_t,y_t} K_t^T$$

As a result, the proposal distribution is obtained as $q(x_t^i|x_{t-1}^i, y_{1:t}) = \mathcal{N}(\bar{x}_t^{(i)}, \bar{P}_t^{(i)})$, and the unscented particle filter is a natural combination of the UKF proposal distribution and traditional particle filter as presented in Section 2.1.

3. SVD Based Kalman Filter

However, UKF often encounters the ill-conditioned problem when solving the square root of the covariance matrix P_{t-1}^a in practice. Therefore, we propose an SVD based Kalman filter to overcome this problem.

To give a clear view, the flowchart of the SVD based Kalman filter framework is schematically shown in Fig. 1. The SVD based KF shares a close spirit to UKF, firstly, the mean state and eigen-covariance matrix of the sigma samples at time $t-1$ are calculated, and we apply the SVD to the eigen-covariance matrix to obtain

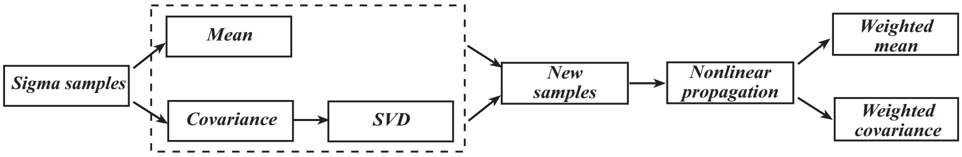

Figure 1. Overview of the SVD based Kalman filter

its eigenvectors. Then the obtained mean and eigenvectors are combined to generate new sigma samples. Finally, the new samples are filtered by the standard Kalman filter. The the detail SVD-based Kalman filter process is presented as follows.

1. Compute the SVD of the eigen-point covariance matrix

$$P^a_{t-1} = U_{t-1} S_{t-1} V^T_{t-1}$$

2. Calculate new sigma samples:

$$\mathcal{X}^{(i)a}_{0,t-1} = \bar{x}^{(i)a}_{t-1}$$

$$\mathcal{X}^{(i)a}_{j,t-1} = [\bar{x}^{(i)a}_{t-1}\ \bar{x}^{(i)a}_{t-1} \pm \rho U_{j,t-1} \sqrt{s_{j,t-1}}]$$

where $U_{j,t-1}, s_{j,t-1}$ are the *jth* eigenvector and eigenvalue respectively, and ρ is the scale parameter.

3. The following step is the same as the standard Kalman filtering.

The basic motivation behind SVD-KF is that the covariance matrix can be characterized by its eigenvectors, and SVD is more numerically robust than Cholesky factorization in the unscented transformation.

4. Incremental Subspace Leaning Based Appearance Model

In our paper, we adopt a subspace based appearance model [10] for observation evaluation, which models the appearance of an object by incrementally learning a low-order eigenspace representation.

Observation Likelihood: As shown in [10], given the learned the subspace U and the new observation y_t, the observation likelihood is based on the reconstruction error of the observation y_i in the object subspace, which is defined as follows.

$$RE = ||y_t - UU^T y_t||^2 \tag{4}$$

As a result, the observation likelihood is naturally formed as

$$p(y_t|x_t) = exp(-RE) \tag{5}$$

Incrementally Subspace Learning: Given the SVD of the previous appearance data $A = \{I_1, \ldots, I_t\}$, i.e. $A = U\Sigma V^T$, where each column I_i is the observation of the object in the *ith* frame. After tracking k frames, we have obtained k newest observations of the object $E = \{I_{t+1}, \ldots, I_{t+k}\}$, the R-SVD algorithm [11] efficiently computes the SVD of the matrix $A' = (A|E) = U'\Sigma' V'^T$ based on the SVD of A as follows:

1. Apply QR decomposition to and get orthonormal basis $\tilde{E}$ of E, and $U' = (U|\tilde{E})$.

2. Let $V' = \begin{pmatrix} V & 0 \\ 0 & I_k \end{pmatrix}$ where I_k is a $k \times k$ identity matrix. It follows then,

$$\Sigma' = U'^T A' V' = \begin{pmatrix} U^T \\ \tilde{E} \end{pmatrix} (A|E) \begin{pmatrix} V & 0 \\ 0 & I_k \end{pmatrix}$$

$$= \begin{pmatrix} U^T AV & U^T E \\ \tilde{E}^T AV & \tilde{E}^T E \end{pmatrix} = \begin{pmatrix} \Sigma & U^T E \\ 0 & \tilde{E}^T E \end{pmatrix}$$

3. Compute the SVD of $\Sigma' = \tilde{U}\tilde{\Sigma}\tilde{V}^T$ and the SVD of A' is

$$A' = U'(\tilde{U}\tilde{\Sigma}\tilde{V}^T)V'^T = (U'\tilde{U})\tilde{\Sigma}(\tilde{V}^T V'^T)$$

In this way, the R-SVD algorithm computes the new eigenbasis efficiently.

5. Experimental Results

In our experiment, the object is initialized manually and affine transformations is considered only. Specifically, the motion is characterized by $s = (t_x, t_y, a_1, a_2, a_3, a_4)$ where $\{t_x, t_y\}$ denote the 2-D translation parameters and $\{a_1, a_2, a_3, a_4\}$ are deformation parameters. Each candidate image is rectified to a 20×20 patch, and the feature is a 400-dimension vector with zero-mean-unit-variance normalization.

In order to demonstrate the effectiveness of our approach, we conduct a comparison experiment among the SVD based KPF (Kalman particle filter), a standard PF (particle filter)[1] and UPF [4] on a video with manually labeled groundtruth.

The David sequence[2] is sampled alternately to form a rapid motion testing sequence. In our implementation, the parameters are set to $\{N = 200, var(\epsilon) =$

[1]Here, a Gaussian transition distribution $x_t \sim \mathcal{N}(x_{t-1}, \Sigma)$ is taken as the proposal distribution

[2]We acknowledge to the author of the source data available at the URL: http://www.cs.toronto.edu/ dross/ivt/

Tracking Method	Frames Tracked	MSE (by pixels)
PF	16/61	26.9481
UPF	61/61	7.1875
SVD based KPF	61/61	3.9868

Table 1. Quantitative results of SVD based KPF tracker and its comparison with PF tracker and UPF tracker

$[5^2, 5^2, 0.01^2, 0.02^2, 0.002^2, 0.001^2]\}$ corresponding to the number of particles and the covariance matrix of the transition distribution respectively. As shown in the first column of Fig.2, the particle filter based tracker fails to track the object at frame 31, because the particles are sampled from the transition distribution to catch the object motion. When the object has rapid and arbitrary motion, the particles drawn from this distribution do not cover a significant region of the likelihood, and thus the weights of most particles are low, leading to the tracking failure. More particles and an enlargement for the diagonal elements of the covariance matrix would improve its performance, but this strategy involves more noises and a heavy computational load. The second column of Fig.2 shows the tracking performance of the unscented particle filter, from which we notice that the tracker follows the object throughout the sequence. However, Cholesky factorization is not numerically robust and often encounters the ill-conditioned problem, thereby resulting to the inaccurate localization and size. In comparison, our method achieves the more accurate results, because the covariance matrix is fully characterized by its eigenvectors, and SVD is more numerically robust than Cholesky factorization. Meanwhile, we have conducted a quantitative evaluation of these algorithms, and have a comparison in the following aspects: frames of successful tracking, MSE (mean square error) between the estimated position and the labeled groundtruth. In table 1, it is clear that the PF tracker fails at frame 31 while the UPF and SVD based KPF trackers succeed in tracking throughout the sequence. Additionally, the SVD based KPF tracker outperforms the UPF tracker in term of accuracy.

6. Conclusion

This paper presents an SVD based Kalman particle filter for visual tracking. In our algorithm, a set of sigma samples are generated by SVD of the covariance matrix, and then these sigma points are propagated by the standard Kalman filter to generate a sophisticated proposal distribution. The obtained proposal distribution is incorporated into the particle filter to form a robust tracking algorithm. Experimental results demonstrate the effectiveness and promising of our approach.

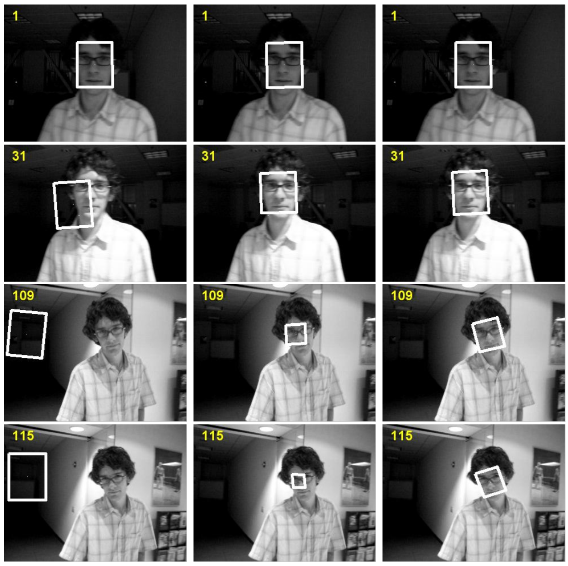

Figure 2. The tracking results (first column: PF, second column: UPF, third column: SVD based KPF)

7. Acknowledgment

This work is partly supported by NSFC (Grant No. 60672040, 60705003) and the National 863 High-Tech R&D Program of China (Grant No. 2006AA01Z453).

References

[1] M. Isard and A. Blake, "Condensation: conditional density propagation for visual tracking", *IJCV.*, 29(1): 5-28, 1998.

[2] M. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particles filters for online nonlinear/non-gaussian bayesian tracking", *IEEE Trans. on Signal Processing*, 50(2): 174-188, 2002.

[3] J. Deutscher, A. Blake, I. Reid, "Articulated body motion capture by annealed particle filtering", *Proc. CVPR'00*, pp. 126-133, 2000.

[4] R. Merwe, A. Doucet, N. Freitas, and E. Wan, "The unscented particle filter", *Technical Report CUED/F-INFENG/TR 380*, Cambridge University Engineering Department, 2000.

[5] Y. Rui and Y. Chen, "Better proposal distributions: object tracking using unscented particle filter", *Proc. CVPR'01*, pp. 786-793 , 2001.

[6] Y.Wu, T.Huang, "Robust Visual Tracking by Integrating Multiple Cues Based on Co-Inference Learning", *IJCV.*, 58(1): 55-71, 2004.

[7] B. Han, Y. Zhu, D. Comaniciu and L. Davis, "Kernel-Based Bayesian Filtering for Object Tracking", *Proc. CVPR'05*, pp. 227-234, 2005.

[8] X. Zhang, W. Hu, G. Luo, and S. Maybank, "Kernel-Bayesian Framework for Object Tracking", *Proc. ACCV'07*, pp. 821-831, 2007.

[9] A. Doucet, S. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering", *Statistics and Computing*, 10(3): 197-208, 2000.

[10] J. Lim, D. Ross, R. S. Lin, and M. H. Yang, "Incremental learning for visual tracking", *In Advances in Neural Information Processing Systems*, pp.793-800, 2004, The MIT Press.

[11] A. Levy and M. Lindenbaum, "Sequential Karhunen-Loeve basis extraction and its application to images", *IEEE Trans. on Image Processing*, 9(8):1371-1374, 2000.