ITC 4/48

Information Technology
and Control
Vol. 48 / No. 4 / 2019
pp. 538-544
DOI 10.5755/j01.itc.48.4.23939

Tracking Video Target via Particle Filtering on Manifold

Received 2019/07/31 | Accepted after revision 2019/09/23

http://dx.doi.org/10.5755/j01.itc.48.4.23939

Tracking Video Target via Particle Filtering on Manifold

Huilin Ge, Zhiyu Zhu

School of Electronic Information, Jiangsu University of Science and Technology, 2 Mengxi Road, Jingkou District, Zhenjiang City, Jiangsu Province, China; phone: +86159 5281 2797; e-mail:164005388@qq.com

Zhiyu Zhu

School of Electronic Information, Jiangsu University of Science and Technology, 2 Mengxi Road, Jingkou District, Zhenjiang City, Jiangsu Province, China; phone:+86139 5280 8837; e-mail:zzydzz@163.com

Kang Lou

School of Electronic Information, Jiangsu University of Science and Technology, 2 Mengxi Road, Jingkou District, Zhenjiang City, Jiangsu Province, China; phone: +86157 5177 6196; e-mail:172030045@stu.just.edu.cn

Corresponding author: zzydzz@163.com

Most of existing particle filtering-based video target tracking algorithms are in Euclidean space, when object posture and scale size changes, and to track high dimensional system, it is difficult to guarantee the tracking effect. This paper describes the covariance descriptor to represent the object image region, the geometric deformation of the object image region can be realized by an affine transformation, and the affine transformation matrix is one element of the Lie group. Then particle filter algorithm based on lie group of manifold is proposed, the video tracking system state lies directly on a low dimensional manifold, state samples are drawn moving on the manifold geodesics, thus state space of intrinsic geometrical characteristic can be in full use, which provides a new idea for improving the tracking efficiency and robustness. Simulation results show that object in the case of geometric deformation including scale size changes, rotating, etc. The proposed manifold particle filtering algorithm can still realize target tracking well and improve the real-time performance.

KEYWORDS: Target tracking, particle filtering, covariance matrix, Lie group.

1. Introduction

Video tracking combines advanced technologies in several areas such as image processing, pattern recognition, artificial intelligence, automatic control, and computer science. Video tracking is a key technology for intelligent monitoring. It is widely used in military vision guidance, video surveillance, robot vision navigation, medical diagnosis and meteorological analysis [16, 10].

At present, common video tracking methods can be generally divided into five types: tracking based on region, tracking based on dynamic contour, tracking based on feature, tracking based on model, and tracking based on motion estimation. The most common method is to establish geometric parameter models for motion of target in image, such as affine models and projection models. After establishing the model, the Lucas-Kanade tracker [1] and the mean drift tracker [15] are applied to obtain the model parameters by minimizing the deviation between the template and the current image region. The process of minimization is mostly achieved by the gradient descent method. However, these methods tend to converge to local minimum values and are sensitive to background interference, clutter, occlusion, fast moving and other factors.

Target tracking based on motion estimation can transform the target tracking problem into a Bayesian estimation problem. Because it is not limited by the prior distribution and the state transition model, the target scale can be easily estimated. Therefore, particle filter has been widely used in visual tracking in recent years, and it has become one of the important researches in video target tracking [17, 2]. Particle filtering can be applied to all nonlinear non-Gaussian systems and is not limited by the nature of the noise. However, when it is used in a high-dimensional system, it will also encounter the "dimensionality disaster" problem.

In video target tracking, the covariance of the observed noise is likely to be unknown and will change according to time. When the covariance matrix is used to represent the target area [3, 7, 12, 8] in the image or to match the image, it is necessary to calculate the difference between covariance matrices of two images. Since the covariance is a positive definite matrix, all positive definite matrices form a Riemannian manifold, so the tracking method in Euclidean space is no longer suitable [5]. A more efficient algorithm must be constructed by using the spatial differential geometry of positive definite matrices. Therefore, how to make full use of the intrinsic geometry of the target motion feature. How to apply the manifold method in differential geometry to the particle filter tracking algorithm [4, 11], and how to continuously improve the speed and robustness of the algorithm are a topic worthy of further study.

2. Manifold Particle Filter Algorithm on Lie Group

In this paper, the projective transformation is used to represent the expansion, translation, deformation and other changes of the image in visual target tracking. This paper also applies the differential geometry math tool to construct the projective transformation of the image into a matrix Lie group [5, 13]. The projective transformation parameters of the target are used as state variables, and the state transition model on the Lie group is established. The Particle Filter algorithm based on the Lie group is analyzed, and the state samples are extracted along the geodesic. The Particle Filter algorithm needs to estimate the mean of the weighted particles when estimating the state. Due to the change of the spatial geometry and the metric, the method of solving the mean value in the European space is no longer suitable for Lie group [9]. The process of solving the problem can be transformed into a constrained optimization problem on the manifold. The intrinsic mean of the manifold is obtained by applying the optimization algorithm on the manifold, and finally the state estimation is realized.

2.1. Representing Projective Transformation as Lie Group

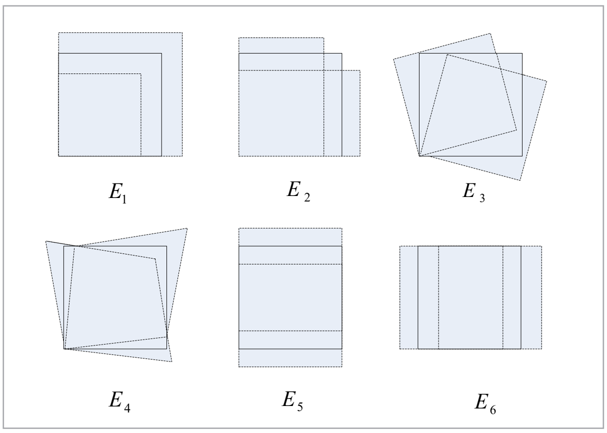
In visual target tracking, the target template is usually used to represent the target of interest. If the target in the image frame is tracked by matching the target template, the geometric deformation of the target image region can be represented as a projective transformation. The 2-dimensional projective transformation matrix is an element of the Lie group, instead of a vector space. Figure 1 shows the geometric deformation of the target in the video image corresponding to the basic elements of each Lie algebra in the 2-dimensional affine group. E_1 represents compression or stretching of the image, E_2 represents image stretching, E_3 represents image rotation left and right, E_4 represents image deformation, E_5 represents image translation up and down, E_6 represents image translation left and right, and the base vector of Lie algebra $SL(3, R)$ is

$$E_1=\begin{bmatrix}1&0&0\\0&1&0\\0&0&0\end{bmatrix},\quad E_2=\begin{bmatrix}1&0&0\\0&-1&0\\0&0&0\end{bmatrix}$$
$$E_3=\begin{bmatrix}0&-1&0\\1&0&0\\0&0&0\end{bmatrix},\quad E_4=\begin{bmatrix}0&1&0\\1&0&0\\0&0&0\end{bmatrix} \quad (1)$$
$$E_5=\begin{bmatrix}0&0&1\\0&0&0\\0&0&0\end{bmatrix},\quad E_6=\begin{bmatrix}0&0&0\\0&0&1\\0&0&0\end{bmatrix}.$$

The matrix expression of the projective transformation model of the moving target is $\begin{bmatrix}\boldsymbol{A}&\boldsymbol{t}\\v&1\end{bmatrix}$. There are 8 parameters in total, where $\boldsymbol{A}$ is a second-order reversible matrix, which represents the deformation of the target E_1 -- E_6. $\boldsymbol{t}$ is the translation vector. $[v\quad 1]$ is the projection of the infinity line. In this paper, the projective transformation group is regularized so that its determinant value is unity, and a special linear group $SL(3,R)$ or a subgroup of Lie group G can be obtained.

Figure 1

Schematic diagram of geometric deformation of a target in a video image corresponding to the basic elements of each Lie algebra in a 2-dimensional affine group

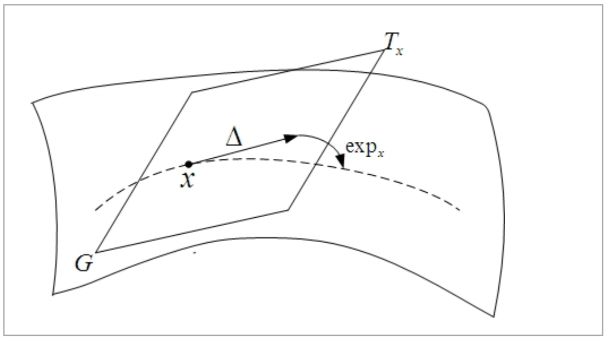


2.2. State Model

Figure 2 shows a 2-dimensional manifold G embedded in a 3-dimensional European space. In the figure, $\boldsymbol{T_x}$ represents the tangent space, which is the tangent plane at point $\boldsymbol{x}$ on the manifold. The cut space can be seen as a set of allowable velocities in which the points on the manifold move over the manifold. The straight arrow Δ represents the tangent at point $\boldsymbol{x}$. The distance between the two points on the manifold is represented by the length of the curve between the two points. The curve with the shortest distance is the geodesic on the manifold, and the length of the geodesic is the intrinsic distance. For each tangent $\Delta\in\boldsymbol{T_x}$, there is a unique geodesic with an initial velocity Δ starting at point $\boldsymbol{x}$. The exponential map $\exp_x$ maps Δ to the endpoint of the geodesic on the manifold.

Figure 2

Manifold and its tangent plane at the point $\boldsymbol{x}$



Select the projective transformation parameter of the target as the state of the system, and use $\boldsymbol{x}_k$ to represent the state vector of the target at time k. $\boldsymbol{Z}_{1:k}=\{\boldsymbol{Z}_1,\boldsymbol{Z}_2,\cdots\boldsymbol{Z}_k\}$ represents the observed sequence of images until time k. In order to obtain the system dynamic model on the Lie group manifold, according to the diagram given in Figure 2, the state vector should be the left invariant vector(the vector in the tangent space) on the Lie group G. $\boldsymbol{v}_k$ represents the moving speed (movement of the image) between the observed image $\boldsymbol{Z}_{k-1}$ at time $k-1$ and the observed image $\boldsymbol{Z}_k$ at time k, and the state transition model can be expressed by the following formula.

$$\begin{cases}\boldsymbol{x}_k=\boldsymbol{x}_{k-1}\exp(\boldsymbol{v}_{k-1})\\ \boldsymbol{v}_k=\boldsymbol{v}_{k-1}+\boldsymbol{\eta}_{k-1}\end{cases}, \quad (2)$$

where $\boldsymbol{\eta}_{k-1}$ represents random noise.

2.3. Observation Model

The measurement process between moving image frames is shown in Figure 3. The projective transformation model of the moving target is represented by $\begin{bmatrix} \boldsymbol{A} & \boldsymbol{t} \\ v & 1 \end{bmatrix}$. The covariance descriptor is used to represent the target area in the image. For a 2-dimensional image, assume that the target area size in the image is $M \times N$, and each pixel point generates a d-dimensional feature vector $\boldsymbol{h}_k = \left[x \ y \ E\right]^T$, $k = 1, 2 \cdots M \times N$. (x, y) is the position coordinate of the pixel. $E = \left(r \ \theta \ I \ I_x \ I_y \ \tan^{-1}\left(\frac{I_y}{I_x}\right) \ I_{xx} \ I_{yy}\right)$, $(r \ \theta)$ represents the polar coordinate of each pixel. I represents the gray level of the pixel, and I_x ,I_y ,I_{xx} ,I_{yy} represents the first-order and second-step degrees of the image respectively. Thus, the covariance matrix of the target region can be determined by the following formula.

$$C = \frac{1}{MN}\sum_{i=1}^{MN}(\boldsymbol{h}_i - \boldsymbol{\mu}_i)(\boldsymbol{h}_i - \boldsymbol{\mu}_i)^T, \tag{3}$$

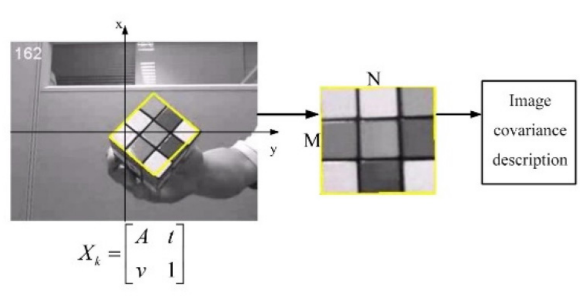
where $\boldsymbol{\mu}_i = \frac{1}{MN}\sum_{i=1}^{MN}\boldsymbol{h}_i$. The likelihood function of the observed model is obtained by calculating the correlation between the covariance matrix of the template region and the covariance matrix of the real-time image frame target region.

$$\rho(y_k \mid X_k^{(i)}) \propto \exp\left(-\frac{\| \log(C_k^{(i)}) - \log(C_0) \|^2}{\sigma^2_{\mathrm{cov}}}\right), \tag{4}$$

where C_0 represents the covariance matrix of the template region, and $C_k^{(i)}$ represents the covariance matrix of the target region of the real-time image frame.

Figure 3

Image measurement process in the video target area



3. Specific Steps of Particle Filter Algorithm on Lie Group Manifold

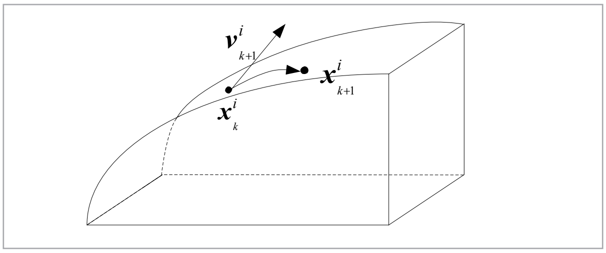
According to the observed image sequence, the particle filter algorithm is applied to estimate the projective parameters (state vector) of the moving target.

Step1: Initialization: The particle set $\{\boldsymbol{x}_0^i\}_{i=1}^{N_s}$ is generated by the prior probability $p(\boldsymbol{x}_0)$, and all the particle weights are $\frac{1}{N_s}$;

Step 2: Prediction: Figure 4 shows a schematic diagram of sampling along the geodesic. $\boldsymbol{v}_k$ is given, the sample $\boldsymbol{v}_{k+1}^i$ is sampled in the left invariant vector space (cut space) of the Lie group manifold. Then according to the dynamic model of the system, the exponential map $\boldsymbol{x}_{k+1}^i = \boldsymbol{x}_k^i \exp(\boldsymbol{v}_{k+1}^i)$ is applied, and the sample $\boldsymbol{v}_{k+1}^i$ is mapped to a state vector sample (particle) $\boldsymbol{x}_{k+1}^i$. Thus, a geodesic on the manifold is obtained. The geodesic starts from $\boldsymbol{x}_k^i$ and $\boldsymbol{x}_{k+1}^i$ is the endpoint of the line. Finally, the samples are obtained from the manifold geodesic.

Figure 4

Schematic diagram of sampling on a geodesic on a manifold



Step 3: Importance weight calculation: Calculate the particle weight according to the likelihood function of the observation model and normalize it;

Step 4: Resample if necessary. In this paper, resample again every 10 filters.

Step 5: Calculate the mean of the weighted particles and obtain the state estimation of the system. According to the differential geometry knowledge, the intrinsic mean calculation of the Lie group manifold can be expressed by the following formula.

$$\bar{\boldsymbol{x}} = \arg\min \sum_{i=1}^{N_s} d^2(\boldsymbol{x}^i, \boldsymbol{x}) \qquad x \in G, \tag{5}$$

where $\boldsymbol{x}^i$ represents the particles sampled in the epidemic, and $\boldsymbol{x}$ is located on the Lie group G. The geodesic distance d between the two Lie algebra elements X and Y can be expressed by the following formula.

$$d(X,Y)=\left\|\log\left(YX^{-1}\right)\right\|. \tag{6}$$

This is a constrained optimization problem on the manifold. The optimization algorithm in the analog European space can construct the intrinsic optimization algorithm. The intrinsic Gauss Newton algorithm is used to solve problem of the intrinsic mean in this paper [14].

Step6: Return to Step2 and perform an iterative operation.

Manifold particle filter algorithm on Lie group can effectively prevent particle degradation. Lie group on manifold can reduce the interference of noise and environmental, promote particles to move to high likelihood region. The Reduction of the number of resampling can also preventing particle degradation.

4. Experiment and Analysis

In order to verify the effectiveness of the algorithm, different algorithms are used in a set of image sequences. This paper records that the Lie group space particle filter algorithm is *LPF* and the Euclidean vector space particle filter tracking algorithm is *VPF*.

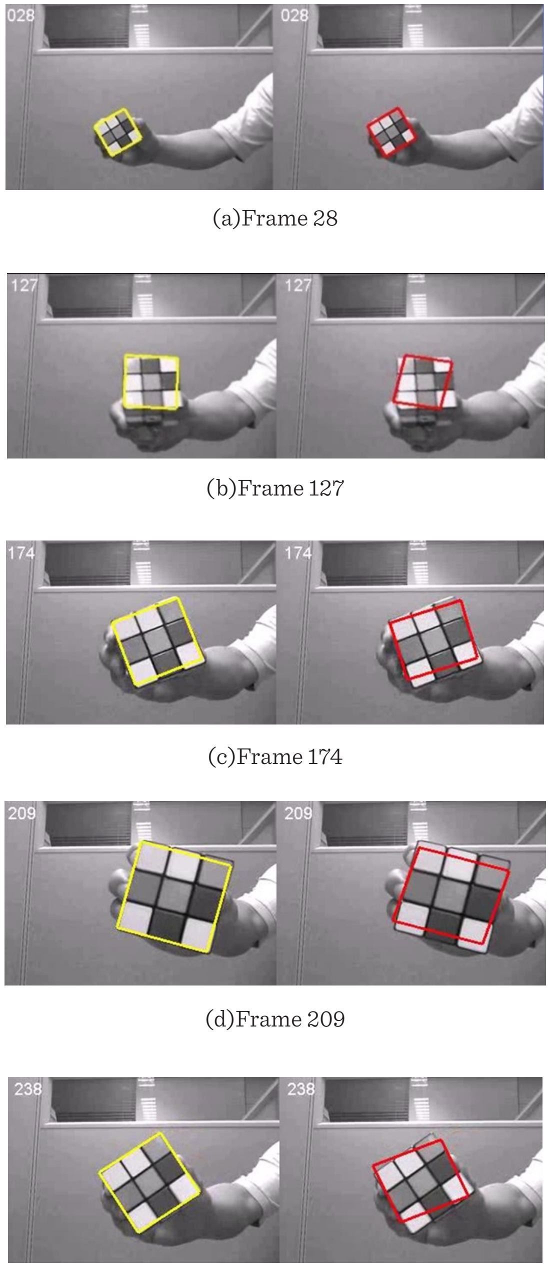
Simulation experiment was carried out by MATLAB. The experiment computer is equipped with 4g ROM and i5 CPU.

The image sequence consists of 270 frames, each frame has a size of 240×320 and the template image size is 51×42. The number of particles set in the two-particle filter tracking comparison test was 400 and 800. The standard deviation of the six Lie group radiological parameters was set to $(0.05\ ,0.02\ ,0.1\ ,0.02\ ,5\ ,5\ )$. In the covariance estimation, the value of σ_{cov} was 0.96. The value of $\boldsymbol{\eta}_{k-1}$ was 5. The results of *LPF* and *VPF* are shown in Figure 5.

In order to illustrate the superiority of *LPF*, this paper selects Rubik's cube as the tracking target. The experimenter moves the target cube from far to near in the video. In the process of moving, the Rubik's cube is also rotated. Five representative trace results are selected, which are 28th, 127th, 174th, 209th and 238th frames respectively. The left side of Figure 5 is the tracking result of *LPF* algorithm, and the right side is the tracking result of *VPF* algorithm. The 28th frame

Figure 5

Comparison of *LPF* and *VPF* tracking results



(a)Frame 28

(b)Frame 127

(c)Frame 174

(d)Frame 209

(e)Frame238

of the Rubik's cube has just begun to move, the deformation is small, and both algorithms can achieve accurate tracking. However, in the 127th and 238th frames, the video object rotates, and the yellow box indicates that *LPF* can achieve steady and accurate tracking. The *VPF* indicated by the red box does not reflect the real situation of the target boundary. When the target in the 174th frame and the 209th frame is deformed, the *VPF* indicated by the red box also loses the accuracy of the target size. This shows the fact that the *VPF* algorithm takes the parameter space as a whole.

As can be seen from the tracking results in the left graph, the *LPF* algorithm combines the geometry of the parameter space, and uses the parameters between the previously consecutive frames to establish the Lie group space structure. It improves the tracking accuracy and robustness of the algorithm.

To further illustrate the superiority of the particle filter tracking algorithm on the manifold, Table 1 and Table 2 give a comparison data of the tracking error and tracking time of the two algorithms.

Table 1

Tracking performance comparison between *LPF* and *VPF* in the cases of the same number of particles

N	*Err.*		*Time*	
	LPF	*VPF*	*LPF*	*VPF*
200	5.6268	30.3762	0.2038	0.4193
400	5.4954	21.0041	0.487	0.8643
600	5.0332	14.8350	0.620	1.247

Table 2

Tracking performance comparison between *LPF* and *VPF* in the cases of the same tracking errors

Err.	*N.*		*Time*	
	LPF	*VPF*	*LPF*	*VPF*
3	1100	2300	1.135	4.913
5	600	910	0.620	1.501
10	100	750	0.106	1.251

N is the number of particles, *Err*. represents the mean error, and *Time* represents the average match time per frame, in seconds. It can be seen from Table 1 that when *N*=200, the mean error of *VPF* is six times that of *LPF*. However, as the number of selected particles increases, when *N*=600, the average matching time per frame of *LPF* is twice as slow as that of *VPF*. It can be seen from Table 2 that *LPF* needs fewer particles and less computing time than *VPF* when the *Err.* is the same. The advantage of VPF is that it can achieve higher filtering accuracy with fewer particles. In the case of the same filtering error, *VPF* requires fewer particles and less computing time, so the tracking speed is faster than *LPF*. *VPF* achieves better comprehensive performance between filtering accuracy and tracking speed. *VPF* improves the tracking accuracy and computing speed by controlling the number of particles.

5. Conclusion

In this paper, the application of popular particle filter algorithm in video target tracking is proposed. The sequential Monte Carlo method is used to realize state sampling directly on low-dimensional manifolds by using the affine Lie group and making full use of the Lie group structure of the projective transformation parameters. The dimension of the target tracking system is reduced, and the real-time and robustness of particle filtering is improved. In this paper, the mean value of the sample is calculated on the manifold, and the state estimation of the system is obtained. This can reduce the influence of the noise statistical characteristics of the European space on the weight variance, help solve the particle degradation problem, and improve the tracking accuracy and robustness of the algorithm. By comparing with the particle filter algorithm of European vector space, the theoretical data is used to analyze the superiority of the proposed algorithm. The effectiveness of the method is verified by experiments.

Highlight: This paper considers directly establishing the state model of the video tracking system on the manifold. This paper also introduces metrics related to the geometry of the state space in the equation of state of the system. At the same time, the particle filter algorithm on the manifold is analyzed, and the intrinsic geometric properties of the state space are fully utilized to provide a new idea for improving the efficiency and robustness of the tracking algorithm. The algorithm proposed in this paper can still achieve tracking well under the conditions of target scale change, geometric deformation such as rotation and multi-target.

Acknowledgement

This study was supported by the National Natural Science Foundation of China (No. 61671222).

References

1. Baker, S., Matthews, I. Lucas-Kanade 20 Years On: A Unifying Framework. International Journal of Computer Vision, 2004, 56(3), 221-255. https://doi.org/10.1023/B:VISI.0000011205.11775.fd
2. Chang, C., Rashid, A. Kernel Particle Filter for Visual Tracking. IEEE Signal Processing Letters, 2005, 12(3), 242- 245. https://doi.org/10.1109/LSP.2004.842254
3. Guan, X. W., Liu, X. Z., Luo, Z. B. Target Tracking Method Based on Region Covariance Matrix. Infrared Technology, 2009, 31(2), 99-102.
4. Guogang, W., Yunpeng, L., Hongyan, S. Covariance Tracking via Geometric Particle Filtering. In Second International Conference on Intelligent Computation Technology and Automation 2009, 250-254. https://doi.org/10.1109/ICICTA.2009.68
5. Kwon, J., Choi, M., Park, F. C., Chun, C. Particle Filtering on the Euclidean Group: Framework and Applications. Robotica, 2007, 25(6), 725-737. https://doi.org/10.1017/S0263574707003529
6. Kwon, J., Park, F. C. Visual Tracking via Particle Filtering on the Affine Group. The International Journal of Robotics Research, 2010, 29(2), 198-217. https://doi.org/10.1177/0278364909345167
7. Li, G. W., Liu, Y. P. Target Tracking with Feature Covariance Based on an Improved Lie Group Structure. Chinese Journal of Scientific Instrument, 2010, 31(1), 111-116.
8. Porikli, F., Tuzel, O., Meer, P. Covariance Tracking Using Model Update Based on Lie Algebra. Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2006, 728-735.
9. Rathi, Y., Vaswani, N., Tannenbaum, A., Yezzi, A. Tracking Deforming Objects Using Particle Filtering for Geometric Active Contours. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 29(8), 1470-1475. https://doi.org/10.1109/TPAMI.2007.1081
10. Sheng, X., Hu, Y. Distributed Particle Filters for Wireless Sensor Network Target Tracking. ICASSP'05. Philadelphia, 2005, 4(2), 845-848.
11. Snoussi, H., Mohammad-Djafari, A. Particle Filtering on Riemannian Manifold. Bayesian Inference and Maximum Entropy Methods in Science and Engineering. AIP Conference Proceedings, 2006, 219-226. https://doi.org/10.1063/1.2423278
12. Tuzel, O., Porikli, F., Meer, P. Region Covariance: A Fast Descriptor for Detection and Classification. In: Leonardis, A., Bischof, H., Pinz, A. (Eds.), Computer Vision - ECCV 2006. ECCV 2006. Lecture Notes in Computer Science, 3952. Springer, Berlin, Heidelberg, 589-600. https://doi.org/10.1007/11744047_45
13. Tuzel, O., Porikli. F., Meer, P. Learning on Lie Groups for Invariant Detection and Tracking. IEEE Transactions on Signal Processing, 2008, 56(1), 84-92. https://doi.org/10.1109/CVPR.2008.4587521
14. Tuzel, O., Subbarao, R., Meer, P. Simultaneous Multiple 3D Motion Estimation via Mode Finding on Lie Groups. In Proceedings of 10th International Conference on Computer Vision, Beijing, China, 2005, 1(1), 18-25. https://doi.org/10.1109/ICCV.2005.226
15. Zhang, H. H., Huang, W. M., Huang, Z. Y., Li, L. Y. Affine Object Tracking with Kernel-Based Spatial-Color Representation. IEEE Confrence on Computer Vision and Pattern Recognition, 2005, 1(2), 293-300.
16. Zhou, S. K., Chellappa, R., Moghaddam, B. Visual Tracking and Recognition Using Appearance Adaptive Models in Particle filters.IEEE Transaction on Image Processing, 2004, 13(11), 1491- 1506. https://doi.org/10.1109/TIP.2004.836152
17. Zhu, Z. Y. Particle Filter Algorithm and Its Application. Beijing: Science Press, 2010.