

## UPF Tracking Method Based on Color and SIFT Features Adaptive Fusion

Yibo Li, Xuezheng Zhuang and Yanmei Liu

*School of Automation, Shenyang Aerospace University, Liao Ning, China  
liyibo\_sau@163.com, z\_snow5z@163.com, lymcml@126.com*

### **Abstract**

*Based on the problems that target appears rotation and noise interference in complex environment, an improved multi-feature adaptive fusion tracking method is proposed. The algorithm adopts unscented Kalman particle filter (UPF) to update the measurement information in the sample particles, better overcome the problem of the particle weight degradation. In addition, in order to overcome the defects of additive and multiplicative fusion algorithm in the feature selection, the multiple adaptive fusion characteristics method that target color distribution and scale invariance feature (SIFT) are used as complementary information. Experimental results show that the proposed method is superior to the traditional methods which are based on fixed weight or standard particle filter.*

**Keywords:** UPF; color histogram; scale invariant; adaptive fusion

### **1. Introduction**

Video moving objects tracking has been widely used in the field of video surveillance, robot navigation, human-computer interaction, *etc.* [1]. Compared with the method based on model, feature-based approach can make simple filtering and tracking stability, which shows a good real-time performance and robustness. Currently, there are common characteristics of the target which mainly includes color, texture, corners, edges and contours [2-4], of which the color characteristics due to the rotation and attitude change for the target has strong robustness. SIFT (scale invariant feature transform) can better respond to changes in illumination, partial occlusion case, noise interference. With the development of multi-source information fusion technology, feature extraction makes a number of different features to additive, multiplicative or DS fusion, which experimental results show better tracking effect. But the simple integration produces a large amount of redundant information to increase the computational load. Meanwhile, it is crucial for tracking target how to adaptively fuse [5-6] the different characteristics by the size of contribution value quantify and the information content.

With the development of the nonlinear filtering techniques, especially the proposed particle filter algorithm, the method can theoretically be achieved on approximation for any system of nonlinear, non-Gaussian, and a series of improved algorithms are proposed [7-9]. Standard particle filter algorithm does not integrate the latest measurement information, which brings greater mismatch errors to system model and ultimately produces particle weight degradation, resulting in filtering estimation accuracy reduces even failure. To solve the problem, paper [10] proposes the improve EPF [11], that is UPF (Unscented Kalman Particle Filter). Due to the UPF makes third-order approximation for the mean and variance of the random variable in system, EPF

can weaken the approximate linear error, but it greatly increases the number of iterations of the particles.

Based on the above analysis, the paper uses the target color histogram and SIFT feature as complementary advantages to establish appropriate adaptive fusion method based on the amount of information, which can better overcome the target mutations in light, attitude change, partial occlusion and noise interference. The method not only overcomes the particle degradation, reduces the number of particles, but makes use of the additive and multiplicative adaptive fusion to improve the recognition of the probability density of the system and weakens the impact of noise on the system estimates. Experiments show that the method for target tracking under complex conditions has high accuracy and robustness.

## 2. Filtering Algorithm Principle

### 2.1. Standard Particle Filter Algorithm

The algorithm of standard particle filter (PF) uses a series of random samples  $\chi_k^{(i)}$  and their corresponding weights  $\omega_k^{(i)}$  to represent the posterior probability density or probability density filter of target location, namely,

$$p(\chi_k | Y_k) = \sum_{i=1}^N \omega_k^{(i)} \delta(\chi_k - \chi_k^{(i)}) \quad (2.1)$$

The state transfer function of the system  $p(\chi_k | \chi_{k-1})$  are used as importance probability density function, and combined with the system state equation to generate random sampling particle  $\{\chi_k^i\}_{i=1}^N$  to make output energy  $p(\chi_k^{(i)})$  as the likelihood function, the recursive formula for particle weights is denoted as:

$$\omega_k^{(i)} = \omega_{k-1}^{(i)} \frac{p(Y_k | \chi_k^{(i)}) p(\chi_k^{(i)} | \chi_{k-1})}{q(\chi_k^{(i)} | \chi_{1:k-1}, Y_k)} = \omega_{k-1}^{(i)} p(\chi_k^{(i)}) \quad (2.2)$$

Then the minimum variance estimation of the target location  $\chi_k$  can be expressed as:

$$\hat{\chi}_k = \sum_{i=1}^N \chi_k^{(i)} \omega_k^{(i)} \quad (2.3)$$

The standard PF algorithm without the proposal distribution integrated the latest measurement information, produces the particle degeneracy phenomenon, so the improved UPF algorithm can weak the particle degradation.

### 2.2. Unscented Kalman Particle Filter

UPF algorithms that use UKF to gain importance probability density on the basis of PF update the sampled particles each time by the UKF algorithm, and the results of sample mean and variance are used as the next new particle. The main idea of unscented filtering is to recursive and update for system model state and the error covariance by an effective nonlinear change method.

Sample points set  $\{\chi_i, W_i\}_{i=1,2,\dots,n}$  represents target state  $\chi$ , detailed in Document[12], the nonlinear function is described by  $y=f(\chi)+\omega$ , where  $\omega$  is white noise,  $Q$  is covariance matrix; Sample points are passed through the system state equations, and the associated sample statistics are calculated, you can get:

$$\gamma_i = f(\chi_i) \quad (2.4)$$

$$\hat{y} = \sum_{i=0}^{2n} W_i \gamma_i \quad (2.5)$$

$$\hat{P}_{yy} = \sum_{i=0}^{2n} W_i (\gamma_i - \hat{y})(\gamma_i - \hat{y})^T + Q \quad (2.6)$$

$$\hat{P}_{xy} = \sum_{i=0}^{2n} W_i (\chi_i - \hat{\chi})(\gamma_i - \hat{y})^T \quad (2.7)$$

When the equation  $y = f(\chi)$  represents state equation of the system, by the formula (2.5)、(2.6) the system new state estimation can be got, they are state predictive value  $\bar{\chi}$  and variance of the predicted value  $\bar{P}_{xx}$ .

When the equation  $y = f(\chi)$  represents observation equation of the system, we get the statistics obtained by the formula (2.5)、(2.7) updated equation of Kalman filter observation to get the improved UPF update equation:

$$\hat{\chi} = \bar{\chi} + K(y - \hat{y}) \quad (2.8)$$

$$P_{xx} = \bar{P}_{xx} - K \hat{P}_{yy} K^T \quad (2.9)$$

$$K = P_{xx} P_{yy}^{-1} \quad (2.10)$$

$$\hat{\chi}_{i+1} = \bar{\chi} + P_{x_i y_i} (P_{y_i y_i})^{-1} (y - \hat{y}) \quad (2.11)$$

$$P_{i+1} = P_i + P_{x_i y_i} (P_{y_i y_i})^{-1} P_{yy} [P_{x_i y_i} (P_{y_i y_i})^{-1}]^T \quad (2.12)$$

### 3. The Implementation of the Adaptive Multi-feature Fusion

#### 3.1. Tracking Model

Filtering method is used to track the target, and firstly motion model is created that is one order recursive model expressed by formula:

$$\chi_k^i = A \chi_{k-1}^i + v_k^i \quad (3.1)$$

Where  $A$  is the state transition matrix of the system.  $v_k^i$  is a Gaussian white noise whose mean is zero.

According to the sample particle weight distribution method, based on Bhattacharyya distance, the likelihood function is defined as the observation model of tracking system:

$$p(z_k | x_k^i) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{l_i^2}{2\sigma^2}\right) \quad (3.2)$$

Where  $l_i$  is the distance between observed value and true value for the  $i$  particle;  $\sigma$  is the variance of the Gaussian distribution; the formula of system status corresponding particle weight distribution can be expressed as:

$$\omega_k^i = \omega_{k-1}^i p(z_k | x_k^i) \quad (3.3)$$

#### 3.2. Feature Fusion

As a result of the attitude change in the target, rotating and blocked or partially blocked, its color feature and objectives of SIFT feature remain better robust. Therefore, color histogram

features and SIFT features in the system are used to estimate as the target characteristic observation information.

**3.2.1. Weights of the Color Histogram Feature:** Assuming the initial position of a target image is  $\chi_0 = (\mathbf{x}, \mathbf{y})^T$ , around the initial position, the characteristics of the target are searched, thus the probability density of characteristics  $u$  of the target feature space can be expressed as:

$$p_u(y) = C_h \sum_{i=1}^m K\left(\frac{\|y - x_i\|}{a}\right) \delta(b(x_i) - u) \quad (3.4)$$

In the above formula, the normalization coefficients for probability density of the target feature is represented as,

$$C_h = 1 / \sum_i^n K(\|x_i^* \parallel) \quad (3.5)$$

The purpose of normalized is to make  $\sum_{u=1}^m p_u(y) = 1$ ; effective area  $a$  is a complete target search area;  $\delta$  is Dirac function;  $y$  represents the center coordinates of the complete search area;  $m$  is the number of pixels in the complete search area;  $b(x_i)$  represents a specific function value that belongs to the target feature corresponding with the position of the  $i$ th particle  $x_i$ ;  $K(\cdot)$  is the feature weights weighting function of color histogram feature. The corresponding weight values are distributed to the pixel values according to different contributions, the specific formula is defined as follows:

$$K(s) = \begin{cases} 1 - s^2, & s < 1 \\ 0, & else \end{cases} \quad (3.6)$$

In the above formula,  $s$  is the distance from a pixel position to the regional center, which determines the weight of current pixel. In order to effectively avoid the influence of the target boundary pixels and interference and occlusion, the pixel should be given a smaller weight value through  $K(\cdot)$  when the pixel position is far away the regional center, the initial template of the whole moving targets is expressed by following formula:

$$p(y) = \{p_u(y)\}_{u=1, \dots, m} \quad (3.7)$$

In the process of tracking system, the matching degree between the candidate models and the target models is measured by Bhattacharyya distance. For the  $k$ th frame, the  $i$ th particle of the positional parameter  $(x_k^i, y_k^i)^T$ , the corresponding matching degree can be showed as:

$$d_i = \sqrt{1 - \rho(p, q)} \quad (3.8)$$

$$\rho(p, q) = \sum_{u=1}^m \sqrt{p_u(y) q_u(y_0)} \quad (3.9)$$

Formula (3.9) is the Bhattacharyya coefficient, then according to the formula (3.7) and (3.8), the update formula of particle weight value can be obtained based on the color histogram:

$$\omega_{CLORk}^i = \omega_{CLORk-1}^i \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{d_i^2}{2\sigma^2}\right) \quad (3.10)$$

**3.2.2. Weight Value of SIFT Feature:** Because the scale invariant features and the transform features have well maintained invariance to brightness and rotation changes, SIFT

features can well meet the requirements of target tracking under the complex conditions. In this paper, Difference of Gaussian (DOG) method is used to calculate the SIFT features.

DOG operator is generated by different scales Difference of Gaussian kernel and image convolution, as shown in the following formula:

$$\begin{aligned} D(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma) \end{aligned} \quad (3.11)$$

In the above formula,  $L$  represents the scale space of images,  $I(x, y)$  represents the pixel value of images at the position  $(x, y)$ , the 2-dimensional Gaussian kernel function is:

$$G(x, y, \sigma) = \frac{1}{2\pi} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (3.12)$$

Where  $\sigma$  represents the mean square error of Gaussian distribution, and which can be called scale space factor. To detect the key point of the target image rectangle after division, the detection parameters are set as follows:

$$O = (\log_2^{\min(M,N)}) - o_{\min} - S, o_{\min} = -1, S = 3$$

Where  $O(Octave)$  represents the number of layers,  $o_{\min}$  represents the first *Octave*,  $S$  represents the number of Scale sub-level,  $M$  and  $N$  are respectively the corresponding rank number of each segmented image blocks, scale factor  $\sigma$  defined as following:

$$\sigma(o, s) = \sigma_0^{2 \frac{o+s}{s}} \quad (3.13)$$

$$s \in [S_{\min}, S_{\max}] = [-1, S], o \in [o_{\min}, O], \sigma = 1.6 \times 2^{1/S}.$$

In a frame image, suppose that the tracked target contains  $m$  SIFT feature points, and  $T$  represents the target model, then  $T = \{f_i\}_{i=1}^m$ , where  $f_i$  represents a feature vector set of SIFT feature points. Assuming at  $k$  time, the candidate target model is represented by  $\hat{T}_k$ , then  $\hat{T}_k = \{T_k^i\}_{i=1}^N$ , where  $T_k^i$  is a feature vector set of SIFT feature points. The number of matching feature points between the target model and the candidate target model is used to describe the similarity of the models, and the observation likelihood function is defined as following according to the observation model formula (3.2) of the tracking system:

$$p(z_k | \chi_k = T_k^i) = 1 - \exp\left(-\sum_{i=1}^m d(f_i, \text{map}(f_i))\right) \quad (3.14)$$

Where  $f \in T$ ,  $\text{map}(f_i)$  is a mapping used to calculate the feature vector and,  $\text{map}(f_i) \in T_k^i$ , the  $d(f_i, \text{map}(f_i))$  is defined as following:

$$d(f_i, \text{map}(f_i)) = \begin{cases} 1, \text{map}(f_i) \neq \emptyset \\ 0, \text{else} \end{cases} \quad (3.15)$$

Then the corresponding weights updating formula is:

$$\omega_{SIFTk}^i = \omega_{SIFTk-1}^i \left( 1 - \exp \left( - \sum_{i=1}^m d(f_i, map(f_i)) \right) \right) \quad (3.16)$$

**3.2.3. Adaptive Fusion based on Multi-feature Weights:** Multi-feature fusion rule is currently a hot research field in multi-feature fusion based on video tracking. We know that a single additive fusion or multiplicative fusion has its disadvantage, a uncertainty adaptive fusion rule is proposed based on the idea proposed in the reference [13]. For N sampling particles of tracking system, normalize the above feature weights respectively:

$$\omega_{di} = \omega_{CLORk}^i / \sum_{i=1}^N \omega_{CLORk}^i \quad (3.17)$$

$$\omega_{Di} = \omega_{SIFTk}^i / \sum_{i=1}^N \omega_{SIFTk}^i \quad (3.18)$$

In this paper two features are fused, assuming that the observation of these two feature values is in the case of conditional independence, then the fused weights can be calculated as follows:

$$\tilde{\omega}_i = \frac{1}{(1 + \beta^1)(1 + \beta^2)} [\omega_{di}\omega_{Di} + \beta^1\omega_{di} + \beta^2\omega_{Di} + \beta^1\beta^2] \quad (3.19)$$

Whereby  $\beta^1$  and  $\beta^2$  are the adaptive adjustment factors after normalized two feature weights respectively. Fused features are equivalent to adaptively adjust the weight values distribution by the additive fusion and multiplicative fusion. The uncertain adaptive adjustment factor  $\beta^i$  is calculated as follows:

$$\beta_{t+1}^i = \sigma_t H(p_t^i) \quad (3.20)$$

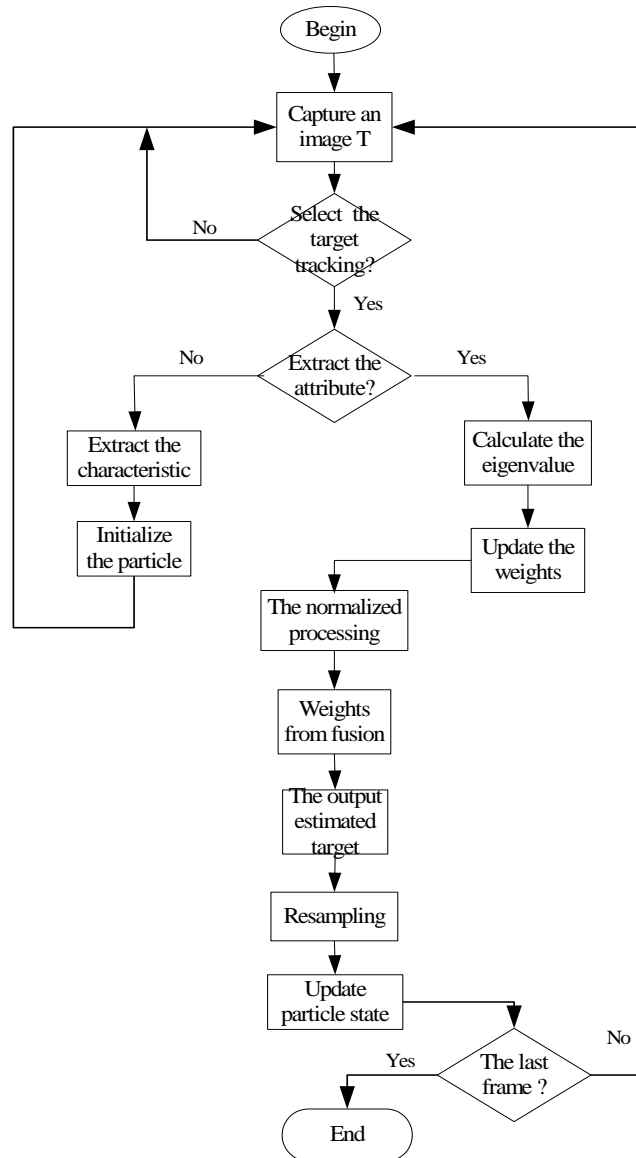
Whereby  $\sigma_t$  is the space position variance of all the sampling particles at  $t$  time, it represents the dispersion degree of the sampling particles in space;  $\sigma_t = tr(\Sigma)$ ,  $\Sigma$  is the position error covariance of two particles at different position. The larger the  $\sigma_t$  is the more dispersing the position of sampling particles is in space.  $H(p_t^i)$  is the information entropy of particle observation probability value of the  $i$ th eigenvalue at  $t$  time, which represents the dispersion degree of the observation probability value to each particles, then:

$$H(p^i) = - \sum_{j=1}^N (p(z^i | x_j) \log_2 p(z^i | x_j)) \quad (3.21)$$

Wherein  $p(z^i | x_j)$  denotes the observation probability of the  $i$ th feature value in the  $j$ th sample particle.  $H(p^i)$  replies the distribution degree of sample particle with the corresponding feature observation values about the system state.

### 3.3. Algorithm Implementation

It is in the framework of particle filter that the algorithms update particle information after each sampling using a modified unscented Kalman filter algorithm for the next sampling new particles, that adaptive fusion is used with target color information and SIFT features to achieve the goal of tracking. Specific tracking process is showed in Figure 1.



**Figure 1. Algorithm Flowchart of UPF Tracking Method Based on Color and SIFT Features Adaptive Fusion**

#### 4. Experiments and Results Analysis

Target tracking experiments are implemented in VC++ 6.0 environment and supported by the OpenCV library functions. The analysis of the algorithms is obtained based on MATLAB R2009a platform.

The first Experiment, we select the standard video sequences provided by the network to test. The image size is 468×337, and the frame rate is 20 frames per second. The tracking results are shown in Figure 2 and Figure 3. The first frame is manually selected. Figure 2 is based on UPF tracking method of two features with fixed weights, and Figure 3 shows the tracking method proposed in this paper. Since the background is complex, and with the changing of the goal rotation and size, the tracking error in Figure 2 becomes larger, due to

target block diagram of the tracking is off target, while the method proposed in this paper remains better tracking effects which is showed in Figure 3.



**Figure 2. The Tracking Results based on Multi-feature with Fixed Weights**



**Figure 3. The Tracking Results based on Multi-feature with Adaptive Fusion Weights**

The video sequences are used in the second Experiment, the image size is  $640 \times 480$  shown in Figure 3. We set a white rectangle to track the distance target. Shown in Figure 4, when the target appears blocked, the method based PF algorithm, due to the similar color to track the target, causes the target tracking error, that shifts away from close range target tracking to track the target, while in Figure 5 the distance little target can be always tracked well.



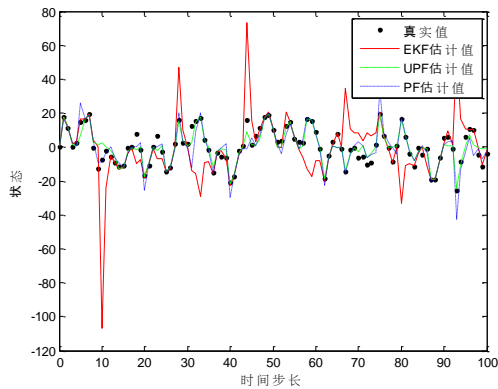
**Figure 4. The Tracking Results based on PF**



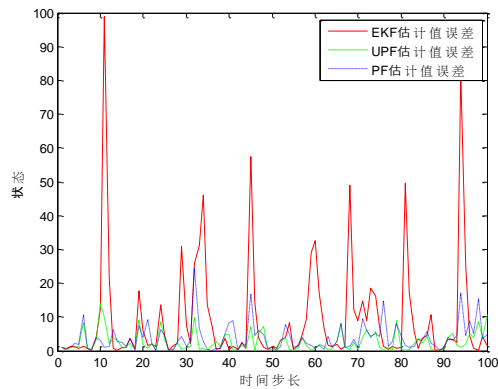


**Figure 5. The Tracking Results based on UPF**

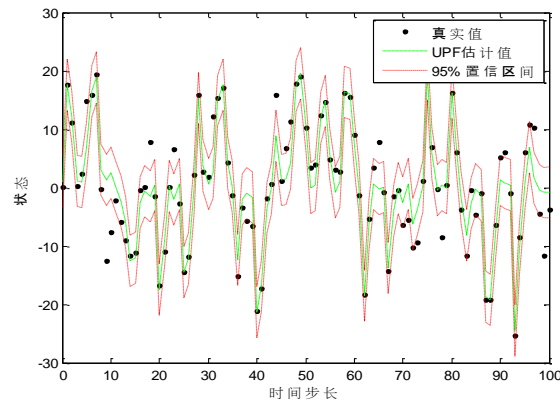
The simulation results of the algorithms of EKF, PF and UPF are shown in Figures 6,7 and Figures 8. In the whole simulation time, the tracking error of UPF filter is far below 5%, which we can see in Figure 8, the algorithm of PF is unstable in tracking, and EKF shows the maximum tracking error. Therefore, it can be concluded that UPF tracking algorithm to track the target also has high accuracy and stability.



**Figure 6. Estimates of Tracking Algorithms**



**Figure 7. Estimates Error of Tracking Algorithms**



**Figure 8. The error range of UPF**

## 5. Conclusion

In this paper, several target tracking methods based on multi-feature fusion are compared. On the one hand, the traditional tracking algorithm is improved based on multi-feature, and the problem of adaptive selection of features fusion is solved, which can better deal with complexity of the environment and the changes of the target size better, and the robustness of target tracking is enhanced; on the other hand, by comparing several tracking algorithms, UPF algorithm better weakens the particles degradation, enhances the filtering accuracy of this algorithm effectively. Experimental results show that the target tracking method proposed in this paper has high accuracy and stability in terms of tracking on targets. However, the computational of UPF tracking algorithm is very complex. The next step is to study binocular stereo vision. The further study need be done to achieve a long time tracking.

## Acknowledgements

Thanks my teachers for guidance to my research. I have learned new views from them, they give me a profound influence on my future research. I am full of enthusiasm and perseverance to this paper. Give sincerely thanks to my friends in this field. Finally, thanks to the members of the expert reviewers to guidance of my papers in their busy.

## References

- [1] Y. Shi and C.-Z. Han, "Adaptive UKF Method with Applications to Target Tracking", *Acta Automatica Sinica*, (2011) June.
- [2] D. Comaniciu and P. Meer, "Image segmentati on using clustering with saddle point detection", *IEEE International Conference on Image Processing*, (2002).
- [3] M. E. Tipping and C. M. Bishop, "Probabilistic Principal Component Analysis", *Journal of Royal Statistical Society*, (1999) January.
- [4] H. T. Nguyen, Q. Ji and AWM. Smeulders, "Spatio-Temporal Context for Robust Multitarget Tracking", *On Pattern Analysis and Machine Intelligence*, (2007) January.
- [5] J. Ma, "Video moving target tracking method. Beijing: Electronic Industry Press", (2013).
- [6] J. Xu, M. Fang and H. Yang, "Computer vision, motion detection and tracking", Beijing: Defense Industry Press, (2012).
- [7] K. J. Watanabe and S. Rittscher, "An HMM-Based Segmentation Method for the Monitoring Movies", *On Pattern Analysis and Machine Intelligence*, (2012) January.
- [8] R. Xiang and J. Li, "Particle Filter Tracking Method of Multiple Features Based Adaptive Fusion", *Journal of Computer-Aided Design & Computer Graphics*, (2012) January.
- [9] R. Zhu, S. Zhang, Q. Sun and S. Liu, "Video Object Tracking Based on Particle Filter with Kalman Prediction Important Proposal Distribution", *Journal of Sichuan Ordnance*, (2013) October.
- [10] S. J. Julier and J. K. Uhlmann, "Unscented Filtering and Nonlinear Estimation", *IEEE Trans. Signal Processing*, (2004) March.
- [11] J.-P. Jia, Y.-N. Zhang and R.-C. Zhao, "Tracking of Objects in Image Sequences using Multi-freeness Mean Shift Algorithm", *Proceeding of 2005 international Conference on Machine Learning and Cybernetics*, (2005).
- [12] A. Doucet, S. J. Godsill and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering", *Statistics and Computing*, (2002) March.
- [13] L. Yang, M. Li and P. Zhang, "New improved particle filter algorithm", *JOURNAL OF XIDIAN UNIVERSITY*, (2010) May.

## Authors



**Yibo Li**, he was born in 1963, he received master degree from Nanjing University of Aeronautics and Astronautics in 1988 and doctor degree from Northeastern University in 2003. He became a teacher in Shenyang Aerospace University since 1988. His main research interests are image processing, pattern recognition, flight control, *etc.*



**Xuezheng Zhuang**, she was born in 1988, she received bachelor degree from Beihua University in 2012. Her main research interests are image processing, pattern recognition, *etc.*



**Yanmei Liu**, she was born in 1974, she received master degree from Shenyang Institute of Automation, Chinese Academy of Sciences. Her main research interests are Pattern Recognition and Intelligent Systems, *etc.*

