

A Target Tracking Algorithm Based on Improved Camshift and UKF

Zouguo Yan*, Weiguo Liang, Haidong lv

School of Optical-Electrical and Computer Engineering, University of Shanghai for Science & Technology, Shanghai, China Email: ^{*}zouguoyan000@126.com

Received 29 September 2014; revised 25 October 2014; accepted 20 November 2014

Academic Editor: Yashwant K. Malaiya, Colorado State University, USA

Copyright © 2014 by authors and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY). <u>http://creativecommons.org/licenses/by/4.0/</u>

😨 🛈 Open Access

Abstract

A tracking algorithm based on improved Camshift and UKF is proposed in this paper to deal with the problems which exist in traditional Camshift algorithm, such as artificial orientation and tracking failure under color interference as well as object's changed illumination occlusion. Meanwhile, in order to solve the sheltered problem, the UKF is combined with improved Camshift algorithm to predict the position of the target effectively. Experiment results show that the proposed algorithm can avoid the interference of the background color and solve the sheltered problem of the object, so that achieving a precise and timely tracking of moving objects. Also it has better robustness to color noises and occlusion when the object's scale changes and deformation occurs.

Keywords

Camshift, UKF, Tracking Algorithm, Sheltered Problem, Interference of the Background Color

1. Introduction

Target tracking in video sequence is an important research subject in computer vision field; it has been widely applied in video surveillance system, aerospace system, intelligent traffic management, medical diagnosis, military and so on. In recent years, the domestic and foreign scholars have conducted a lot of research on this field and come up with a lot of different tracking methods. Meanshift [1] [2] algorithm is a popular nonparametric and fast matching algorithm based on density estimation which has been widely used. However, the lack of template updating and fixed kernel can make target lost [3] when the scale of target is changed. Bradski [4]

How to cite this paper: Yan, Z.G., Liang, W.G. and Iv, H.D. (2014) A Target Tracking Algorithm Based on Improved Camshift and UKF. *Journal of Software Engineering and Applications*, **7**, 1065-1073. <u>http://dx.doi.org/10.4236/jsea.2014.713094</u>

^{*}Corresponding author.

proposed Camshift algorithm and did Meanshift tracking for each frame. The algorithm can adapt to size change automatically, but when the light is stronger, the interference of similar color background or occlusion occurs and the tracking fails. In order to reduce the influence of illumination on tracking, Camshift algorithm firstly transforms the image from RGB color space to HSV color space [5], builds color histogram with H component of the image, and then replaces the pixel value of the original image with corresponding pixel's statistics, requantizes and gets distribution image of color probability. Color histogram can only reflect the statistical information of target's color, but fails to reflect the spatial information of the target. When the light outside is not stable, Camshift can't go on tracking [6]. A common method of target tracking is feature matching, Speeded-Up Robust Features (SURF) keeps its invariant feature locally on translation, rotation, scaling and so on based on scale space. Also it is better than Scale-invariant feature transform (SIFT) [7] and other algorithms on reproducibility, robustness and resolution. It has characteristics of high speed and high precision. Ta et al. studied a continuous fast target tracking and recognition algorithm named SURFTrack [8] [9] using a local feature description. Experiments prove that it has good tracking performance. Reference [10] combines color feature and SURF feature to position target precisely and updates the result according to the Bhattacharyva coefficient between candidate region and template region. The problems of target deformation, scale change and color interference can be solved effectively, but SURF tracks targets based on matching of feature points, it fails to track easily [11] when the target is small or the feature is monotonous. Also it can't meet the real-time requirement when it is applied to large size image. As in [12], Kalman Filter (KF) is introduced into Camshift algorithm to effectively predict the position of the target and improve the accuracy and real-time of tracking. However, KF is only suitable for linear systems.

According to the above problems, improved Camshift algorithm combined with Unscented Kalman Filter (UKF) [13] [14] is proposed. Firstly the color model of target image is converted from RGB to HSV, then the H component of HSV color space is used to establish color model of target. Because of SURF's invariant feature, the influence of illumination and similar background on Camshift can be reduced when the Camshift's result is added with SURFTrack algorithm by weight; in order to avoid the influence of occlusion, UKF algorithm is introduced into the improved Camshift to predict target's position effectively and make tracking accurate and real-time.

2. Algorithm Principle

2.1. Camshift Algorithm

Camshift is an improved algorithm based on Meanshift, mainly composed by reverse projection, Meanshift algorithm and Camshift algorithm [15]. The Camshift algorithm firstly converses color model of image from RGB to HSV color model, then use H component to establish target's color model, then replace the original image pixel value with pixel statistics of the corresponding histogram, re quantize and get distribution image of the color probability.

Camshift algorithm does continuous Meanshift operation for all frames of the video sequence, refers to the center position and the size of the search window by calculating as initial value for next frame's Meanshift search window. Iteration continues like this, target can be tracked.

2.2. UKF Algorithm

According to the nonlinear problems in maneuvering tracking, improving the filtering effect, KF filter was presented based on the Unscented transform namely UKF by Julier [16] which is used for video tracking [17]. on the premise of invariant mean and covariance of stochastic vector, a set of sample points are chosen to go through Unscented transform for pre estimating the mean and covariance with statistics of transformed sigma points after non linearization to avoid error caused by the linearization.

2.2.1. Establish UKF Model

UKF used for target tracking can be divided into two parts, namely, state transform model and state observation model. The state is the state of target, observation is for sequence image; target state includes center position's coordinates and speed of target. Target's speed changes randomly, assuming that its acceleration is *a* and a(k) obeys the Gauss distribution $N(0, \sigma_w^2)$. Set the target state vector: $X(k) = (x, y, \dot{x}, \dot{y})_k$. *x* and *y*

represent the coordinates of the target's center position, \dot{x} and \dot{y} are respectively derivative of x and y, representing the target's speed. Observation variable $y(k) = [x_c(k), y_c(k)]_k^T$ ($x_c(k)$ and $y_c(k)$ represent observation value of target location). UKF state transition model and observation model are respectively:

$$X(k) = \Phi X(k-1) + \Gamma w_k \tag{1}$$

$$y(k) = RX(k) + Ev_k \tag{2}$$

$$\Phi = \begin{bmatrix} 1, 0, t, 0; 0, 1, 0, t; 0, 0, 1, 0; 0, 0, 0, 1 \end{bmatrix}; \quad \Gamma = \begin{bmatrix} t^2/2, t^2/2, t, t \end{bmatrix}^T; \quad E = \begin{bmatrix} 1, 1 \end{bmatrix}^T; \quad R = \begin{bmatrix} 1, 0, 0, 0; 0, 1, 0, 0 \end{bmatrix}$$

where w_k and v_k respectively represent Gauss white noise of state transition model and observation model; t is interval of two adjacent frames, t = 1 s;

Assuming x_c and y_c are coordinates of target's center which is initialized by manually positioning the target area, X(0) = (437, 535, 0, 0).

2.2.2. UKF Filter

Unscented transform is the core of UKF algorithm, also is an important method for nonlinear state estimation. Assuming state mean and variance of state vector with L dimensions at time k-1 are \overline{x}_{k-1} and P(k-1), The state vector of UKF is a amplified vector $x_k^a = (x_k^T, \eta_k^T, \xi_k^T)$ composed of X and noise, the corresponding vector of sigma points is $\chi^a = ((\chi^x)^T, (\chi^{\eta})^T, (\chi^{\xi})^T)^T$, the concrete implementation steps of UKF filter estimation are as follows:

1) Initialization

$$\overline{X}_0 = E[X_0] \tag{3}$$

$$P_0 = E\left(X_0 - \overline{X}_0\right) \left(X_0 - \overline{X}_0\right)^{\mathrm{T}}$$

$$\tag{4}$$

$$\overline{X}_{0}^{a} = E\left[X^{a}\right] = \left[X_{0}^{\mathrm{T}}, 0, 0\right]$$
(5)

$$P_0^a = E\left(x_0^a - \overline{x}_0^a\right)\left(x_0^a - \overline{x}_0^a\right)^{\mathrm{T}} = \mathrm{diag}\left\{P_0, P_y, P_n\right\}$$
(6)

where P_{y} is noise covariance of state transition matrix, P_{n} is the observation noise.

2) Calculate sigma points using Equation (7)

$$\chi_{k-1}^{a} = \left[\overline{X}_{k-1}^{a}, \overline{X}_{k-1}^{a} + \sqrt{(L+\lambda)P_{k-1}^{a}}, \overline{X}_{k-1}^{a} - \sqrt{(L+\lambda)P_{k-1}^{a}} \right]$$
(7)

where $\lambda = a^2 (L+k) - L$; k = 3 - L, *a* is candidate parameter, $0 < a \le 10^{-4}$.

3) Time updating. Take sigma points into state transition Equation (8) and observation Equation (9), calculate the average value of state vector at time k by Equation (10)

$$\chi_{k|k-1}^{x} = F\left[\chi_{k-1}^{a}, \chi_{k-1}^{\eta}\right]$$
(8)

$$y_{k|k-1} = H\left[\chi_{k|k-1}^{x}\right] \tag{9}$$

$$\overline{X}_{k}^{-} = \sum_{i=0}^{2L} w_{i}^{m} \chi_{k|k-1}^{x}$$
(10)

where F is state transition equation, H is observation equation, w_i^m is the weight coefficient for average value, $w_0^m = \lambda/L + \lambda$; $w_i^m = 1/[2(L+\lambda)]$; $i = 1, \dots, 2L$.

4) Observation update equation. Take Equation (8) and Equation (9) into Equation (11) and Equation (12), calculate gain by Equation (13)

$$P_{\overline{y}_k \overline{y}_k} = \sum_{i=0}^{2L} w_i^{(c)} \left[y_{i,k|k-1} - \overline{y}_k^- \right] \left[y_{i,k|k-1} - \overline{y}_k^- \right]^{\mathrm{T}}$$
(11)

$$P_{k}^{-} = \sum_{i=0}^{2L} w_{i}^{c} \left[\chi_{i,k|k-1}^{x} - \overline{x}_{k}^{-} \right] \left[\chi_{i,k|k-1}^{x} - \overline{x}_{k}^{-} \right]^{\mathrm{T}}$$
(12)

$$k = P_{x_k y_k} P_{y_k y_k} \tag{13}$$

$$\beta \ge 0$$
, here $\beta = 0$; $w_0^c = \lambda/(L+\lambda) + (1-a^2+\beta)$ $w_i^c = w_i^m (i=1,\dots,2L)$.
Take gain into Equation (14) and Equation (15) undetermore and variance of state

Take gain into Equation (14) and Equation (15), update mean and variance of state vector.

.

$$\overline{x}_{k} = \overline{x}_{k}^{-} + k\left(y_{k} - y_{k}^{-}\right) \tag{14}$$

$$P_k = P_k^- - k P_{\nu_k \nu_k} k^{\mathrm{T}}$$
⁽¹⁵⁾

Predict position of fast moving target by UKF filter, due to uncertainty of moving object and moving model, feedback Camshift's tracking result to UKF for updating and correcting its state model each time.

3. Tracking Algorithm Based on SURF Matching

3.1. SURF Feature Extraction and Target Description

SURF feature can better describe the target texture and spatial information, its performance is roughly the same as SIFT. Due to the adoption of box shaped filtering and integral image, its computing speed is nearly four times faster than SIFT, also it has good real-time performance. Process of SURF feature extraction is shown in Figure 1.

Process of SURF feature extraction: firstly, establish integral image for each frame image, increase window size of box shaped filter gradually, do fast convolution for integral image, thus construct the Pyramid scale space; then subtract two adjacent images on each layer in Pyramid to obtain the differential scale space, compare each point of this space with all the points on the adjacent scale and the same scale of 3 * 3 * 3 Stereo neighborhood, obtain extreme points by the Hessian matrix, next do interpolation calculation and optimization to get the stable feature points; find the main direction around the feature points, construct the region, do Hart wavelet transform for points on the region, extract SURF description vector; at last, calculate Euclidean distance of the two images' feature vector respectively and match with method of nearest neighbor matching.

3.2. Target Locating and Tracking

In this paper, firstly target tracking template is acquired from video frames, then extract SURF feature points of



target template and initialize position and size of the target. Assuming target's Center of last frame is (x_0^{k-1}, y_0^{k-1}) and the scale of tracking window is (h_x^{k-1}, h_y^{k-1}) , relationship between two adjacent frames of images can be described by affine transformation model, establish affine transformation model by feature points of the two adjacent frames through SURF matching, calculate affine transformation matrix of them, thus correct target's position and size. The theory of affine transformation leads to target's center position as follows:

$$\left[x^{k}, y^{k}\right]^{\mathrm{T}} = A_{k}\left[x^{k-1}, y^{k-1}\right]^{\mathrm{T}} + b_{k} = \left[h_{x}, 0; 0, h_{y}\right]\left[x_{k-1}, y_{k-1}\right]^{\mathrm{T}} + \left[b_{x}, b_{y}\right]^{\mathrm{T}}$$
(16)

where h_x and h_y represent telescopic distance of the target tracking window on horizontal and vertical direction respectively, b_x and b_y are the corresponding offset scales, A_k and b_k can be obtained by RANSAC method.

New target's center (x_0^k, y_0^k) and target's window size (h_x^k, h_y^k) in current frame can be obtained by Equation (13). When matched points between the adjacent frames are too many, this method is time consuming, make the centroid of the feature points matched successfully as center position of the current frame, offset value of target's scale is up to maximal change value between horizontal direction and vertical direction.

$$\left(x_{0}^{\prime k}, y_{0}^{\prime k}\right) = \left(\frac{\sum_{i=1}^{N} x_{i}^{k}}{N}, \frac{\sum_{i=1}^{N} y_{i}^{k}}{N}\right)$$
(17)

$$(h_x'^k, h_y'^k) = \left(\max \left| x_i^k - x_0^k \right|, \max \left| y_i^k - y_0^k \right| \right)$$
(18)

where N is number of feature points matched successfully between two adjacent frames.

4. The Algorithm's Improvement and Implementation Steps

The traditional Camshift algorithm transforms color space from RGB to HSV, regards the H component in HSV as the histogram's template information to get distribution of color probability, in order to reduce the influence of illumination on target tracking, but when the background color is similar to target's, accuracy of target tracking will greatly decreases. Because of SURF's rotation and scale invariance, also it adapts to illumination change easily, in some occlusions or chaotic scenes, it can still maintain invariant advantages. An improved target tracking algorithm proposed in this paper with the fusion of Camshift and SURF. The block diagram of the improved Camshift algorithm is shown in Figure 2.

Implementation steps of the improved Camshift algorithm:

1) Initialize the first frame to determine the target template, set size and position of the tracking window, extract SURF feature and histogram of target template;

2) Find the centroid in the search window by Camshift algorithm, obtain new size *s* and center position (x_{c1}, y_{c1}) of search window, window size and Bhattacharyya coefficient between the target image's and the current frame's color histograms;

3) At the same time, use tracking algorithm of SURF's feature matching to find center position (x_c, y_c) and scale (h_x, h_y) in new search window of target in the current frame, calculate Bhattacharyya coefficient ρ_2 between search area's and template's color histograms;

4) After the two above tracking, use ρ_1 and ρ_2 to weight these two tracking results as result of the target in the current frame, it equals to correct Camshift's tracking result with SURF's tracking result since Camshift's tracking result are not very accurate under illumination change or background interference, the tracking result after correction is as the following:

$$(x'_{c}, y'_{c}) = \frac{\rho_{1}}{\rho_{1} + \rho_{2}} (x_{c}, y_{c}) + \frac{\rho_{2}}{\rho_{1} + \rho_{2}} (x_{c1}, y_{c1}), (w', h') = \frac{\rho_{1}}{\rho_{1} + \rho_{2}} (w + w \times h_{x} + h \times h_{y}) + \frac{\rho_{2}}{\rho_{1} + \rho_{2}} (w \times s, h \times s).$$

In the next frame of video image, initialize search window's position and size by value of the fourth step.

Jump to Step 2 and continue to run.

Combine the improved Camshift algorithm with UKF filter to avoid the influence of the occlusion, nonlinear motion, fast speed and other factors on tracking. Algorithm's block diagram is shown in Figure 3:

First of all, initialize the initial state X_0 of UKF filter by centroid of search window, according to Equation (1), precast position of the window, correct predicted results by Equation (2), finally regard state obtained by UKF's state correction equation as input of improved Camshift algorithm at the next time.

5. Experimental Results and Analysis

The algorithm in this paper is implemented based on MATLAB2013a software platform. Experimental video sequence is AVI format, video acquisition speed is 20 frame/s and the resolution is 1200*900, the initial state of the car is X(0) = (352,306,0,0) and the initial state of ball is X(0) = (437,535,0,0). Tracking effects of



Figure 3. Block diagram of improved Camshift algorithm combined with UKF.

the algorithm are verified from color interference and occlusion compared with the traditional Camshift algorithm's. The experimental results are showed in **Figures 4-7**.

Figure 4 shows tracking effects of traditional Camshift, larger window offset occurs when the target is more and more close to the black car, moreover tracking has trend of failure when the car turns. **Figure 4** shows the tracking effect of the improved Camshift algorithm, due to Camshift's tracking result are corrected by linear weighted method of improved Camshift's tracking result and SURF's tracking result, the tracking algorithm is considered by not only the color information for tracking effectiveness also texture feature for strengthening tracking accuracy. From the result figures, we can see moving target can be accurately tracked by improved Camshift algorithm under background color interference, comparison of tracking performance between **Figure 4** and **Figure 5** is shown in **Table 1**.





Figure 5. Improved camshift algorithm.



Figure 6. Tracking results of improved camshift algorithm.

Figure 6 and **Figure 7** show the tracking effects of improved Camshift algorithm and algorithm in this paper when the moving ball is blocked. **Figure 6** shows tracking window fails to catch up with ball timely after the ball rolls across box. It can be seen in **Figure 7**, when the ball is shield, algorithm in this paper can keep tracking from occlusion, the tracking window can still accurately and timely track with the ball; Time to locate target is reduced and the accuracy is increased because UKF algorithm is introduced to predict the target, comparison of tracking performance between **Figure 6** and **Figure 7** is shown in **Table 2**.

The experimental results show that, target's tracking effect of improved Camshift combined with UKF is more obvious than the traditional Camshift algorithm's under color interference of the similar background occusion, it can realize the accurate and real-time tracking of moving target. Tracking effects of algorithm in this paper are better than improved Camshift algorithm's, it can track moving target accurately and real-timely under occlusion by moving targets.

6. Conclusion

The algorithm in this paper firstly tracks moving target by Camshift algorithm, then linearly weights Camshift's tracking result with tracking result of SURF to reduce the influence of similar background interference on tracking result, and finally combines the improved Camshift algorithm with UKF to solve the occlusion problem



Figure 7. Tracking results of improved camshift algorithm combined with UKF.

Table 1. Performance comparison of traditional Camshift and improved Camshift.

Algorithm	x Average Error/Pixcel	y Average Error/Pixcel	Average Tracking Error ms/Frame
Traditional Camshift	27.6 ± 6.3	25.5 ± 7.5	6.5
Improved Camshift	14.4 ± 5.2	12.5 ± 5.5	3.3

Table 2. Performance comparison of improved Camshift algorithm and algorithm in this paper.

Algorithm	Average Iteration/Count	Average Location/ms	Average Error/Pixcel
Improved Camshift	2.5	7.8	13.4
Algorithm in this paper	1.6	5.7	7.1

because UKF can predict target even if the target is sheltered or moving nonlinearly. Since the UKF algorithm has a very accurate tracking result even target moves with uncertain direction and strong random, it may be used for multiple targets tracking in the future. Also SURF can't accurately track the target under unitary texture feature, so in multiple targets tracking we should pay attention to research of tracking characteristics.

Acknowledgements

The author would like to thank the student, Lv Haidong, who has helped in implementing the proposed algorithm and making this work possible.

References

- [1] Yuan, X. and Wang, L.P. (2008) Moving Human Tracking Based on Meanshift Algorithm. *Computer Engineering and Science*, **30**, 64-84.
- [2] Comaniciu, D. and Ramesh, V. (2000) Real-Time Tracking of Non-Rigid Objects Using Mean Shift. IEEE Conference on Computer Vision and Pattern Recognition, New York, 142-149.
- [3] Liu, Q., Tang, L.B. and Zhao, B.J. (2012) Meanshift Tracking Algorithm with Adaptive Tracking Window. System Engineering and Electronics, 34, 409-412.
- Bradski, G.R. (1998) Real Time Face Tracking as a Component of a Perceptual User Interface. Proceedings of IEEE Workshop Applications of Computer Vision, Princeton, 214-219.
- [5] Wu, H.M. and Zheng, X.S. (2009) Efficient and Improved Camshift Tracking Algorithm. *Computer Engineering and Application*, 45, 178-180.
- [6] Herbert, B., Andreas, E., Tinne, T., et al. (2006) Speeded-Up Robust Features. Computer Vision and Image Understanding, 404-417.
- [7] Luo, J. and Gwun, O. (2009) A Comparison of SIFT, PCA-SIFY and SURF. *International Journal of Image Processing*, 3, 143-152.
- [8] Ta, D.N., Chen, W.C., Gelfand, N., et al. (2009) SURFTrack: Efficient Tracking and Continuous Object Recognition Using Local Features Descriptors. Computer Vision and Pattern Recognition (CVPR09), Miami, 2937-2943.
- [9] Su, D.Z., Wang, K., Wang, Y.L., et al. (2013) Moving Target Tracking Based on SURF Algorithm and Kalman Prediction. Journal of Naval Aeronautical Engineering Institute, 28, 379-382.
- [10] Ding, P.H., Fan, X.N. and Liu, J.D. (2012) A Tracking Algorithm for Moving Object Based on Mixed Algorithms. *Science Technology and Engineering*, 12, 4188-4190.
- [11] Kloihofer, W. and Kampel, M. (2010) Interest Point Based Tracking. IEEE International Conference on Pattern Recognition, Istanbul, 23-26 August 2010, 3549-3552.
- [12] Dong, H.Y., Chen, S.F. and Zhu, J.Y. (2010) Research of Moving Targets Tracking Algorithm Based on Kalman Filtering. 2010 3rd International Conference on Intelligent Networks and Intelligent Systems, Shenyang, 1-3 November 2010, 20-23.
- [13] Vaccarella, A., De Momi, E., Enquobahrie, A. and Ferrigno, G. (2013) Unscented Kalman Filter Based Sensor Fusion for Robust Optical and Electromagnetic Tracking in Surgical Navigation. *IEEE Transactions on Instrumentation and Measurement*, 62, 2067-2081. <u>http://dx.doi.org/10.1109/TIM.2013.2248304</u>
- [14] Henke, D., Magnard, C., Frioud, M., Small, D., Meier, E. and Schaepman, M.E. (2012) Moving-Target Tracking in Single-Channel Wide-Beam SAR. *IEEE Transactions on Geoscience and Remote Sensing*, 50, 4735-4747. http://dx.doi.org/10.1109/TGRS.2012.2191561
- [15] Liang, J., Xiang, J. and Hou, J.H. (2011) Automatic Tracking Algorithm Based on Camshift and Kalman Filter. *Microcomputer and Application*, 30, 28-31.
- [16] Julier, S.J. and Uhlmann, J.K. (2004) Unscented Filtering and Nonlinear Estimation. Proceedings of the IEEE, 92, 401-422. <u>http://dx.doi.org/10.1109/JPROC.2003.823141</u>
- [17] Xu, H.N., Xiao, H., Hou, H.L., *et al.* (2011) Tracking Algorithm for Maneuvering Objects in Simplified Interacting Video Images with Multiple Models Based on UKF. *Opto Electronic Engineering*, **37**, 15-18.



IIIIII II

 \checkmark

Scientific Research Publishing (SCIRP) is one of the largest Open Access journal publishers. It is currently publishing more than 200 open access, online, peer-reviewed journals covering a wide range of academic disciplines. SCIRP serves the worldwide academic communities and contributes to the progress and application of science with its publication.

Other selected journals from SCIRP are listed as below. Submit your manuscript to us via either submit@scirp.org or Online Submission Portal.

