ICIC Express Letters
Part B: Applications
Volume **7**, Number **6**, June **2016**

ICIC International ©2016 ISSN 2185-2766
pp. 1291–1297

AN IMPROVED MOVING OBJECT TRACKING METHOD BASED ON MEANSHIFT ALGORITHM

Shiqing Ren and Yang Li

School of Information Science and Engineering
Shenyang Ligong University
No. 6, Nanping Middle Road, Hunnan District, Shenyang 110159, P. R. China
renshiqingcn@126.com; ly2015sy@hotmail.com

Received November 2015; accepted February 2016

Abstract. *MeanShift algorithm is the classic target tracking method, which is widely used in many fields. However, it has some limitations when the target's scale changes. In a tracking scenario, it is not uncommon to observe objects with complex shapes whose scale and orientation constantly change due to the camera and object motions. In this paper, an object tracking method was presented based on MeanShift combined with elliptical polar coordinates transformation and Kalman filter, in which the scale and orientation of the kernel adaptively change depending on the observations at each iteration. According to test of the target of pedestrians, the algorithm's performance is improved obviously compared to the traditional algorithm.*

Keywords: MeanShift algorithm, Elliptical polar coordinates, Kalman filter

1. **Introduction.** In computer vision, tracking refers to the task of generating the trajectories of the moving objects by computing its motion in a sequence of images [1]. Object motion recorded in the form of a trajectory commonly contains translational motion. Numerous approaches have been dedicated to computing the translation of an object in consecutive frames, among which the MeanShift method is one of the most common methods which is also used in the commercial applications. The popularity of the MeanShift method is due to its ease of implementation, real time response and robust tracking performance.

MeanShift is a nonparametric density estimator which iteratively computes the nearest mode of a sample distribution. After its introduction in the literature, it has been adopted to solve various computer vision problems, such as line fitting, segmentation [2] and object tracking [3,4]. Despite its promising performance, as discussed in various papers [5] and [6], the traditional MeanShift method has two main limitations, the first of which is the constancy of the kernel bandwidth. The changes in the object scale require an adjustment of the kernel bandwidth in order to consistently track the object. An intuitive approach to estimate the object scale is to search for the best scale by testing different kernel bandwidths and selecting the bandwidth which maximizes the appearance similarity. Alternatively, after the object center is estimated, a MeanShift procedure can compute the bandwidth of the kernel in the scale space, which is formed by convolving the image with a set of Gaussian kernels at various scales [6].

In this paper, the traditional MeanShift method was extended by introducing the use of elliptical polar coordinates and Kalman filter.

2. **MeanShift Tracking.** The MeanShift method iteratively computes the closest mode of a sample distribution starting from a hypothesized mode. In the context of tracking, a sample corresponds to the color observed at a pixel x and has an associated sample weight $w(x)$, which defines how likely the pixel color $I(x)$ belongs to an object model q.

The traditional MeanShift tracking method evaluates the displacement (translation) of the object centroid by computing the MeanShift vector Δx. Let the initial object position be $\hat{x}$, and compute the new object position $\hat{x}'$, such that $\hat{x}' = \hat{x} + \Delta x$. The MeanShift vector is computed using the following:

$$\Delta x = \frac{\sum_i K(x_i - \hat{x}) w(x_i)(x_i - \hat{x})}{\sum_i K(x_i - \hat{x}) w(x_i)} \tag{1}$$

where the denominator serves as a normalization term, and $K(.)$ is a radially symmetric kernel with bandwidth h defining the tracked object region. Given the color distribution functions q and p generated from the model and candidate object regions, the weight at pixel x is derived from the Bhattacharyya measure and is given by:

$$w(x) = \sqrt{q(I(x))/p(I(x))} \tag{2}$$

3. The Improved MeanShift.

3.1. Elliptical polar coordinates transformation.

For pedestrians target, using elliptical polar coordinates transformation can turn the calculation of two dimensional space to the translation of coordinate axes. Firstly, make the image in Cartesian coordinates to polar coordinates, then do the logarithmic transformation and get the elliptical polar coordinates transformation. Because there is decimal in logarithmic computing, which also will narrow the scope of the value, enlarge the result by multiplying parameters k_1 and k_2. Suppose the length of an image is a and the width is b. k_1 is equal to $a/\log \rho_{\max}$ ($\rho_{\max}$ is the maximum value of radius vector), and k_2 is equal to $b/2\pi$, because the two parameters make the size of the result image in polar coordinates equal to size of the original image in Cartesian coordinates. Transform formulas are as follows:

$$r = k_1 \cdot \log \sqrt{(x - x_0)^2 + a^2(y - y_0)^2/b^2} \tag{3}$$

$$\theta = k_2 \cdot \alpha \tag{4}$$

(x_0, y_0) is the center. The value of α is as follows:

$$\alpha = \begin{cases} \pi/2 - \arctan(\frac{y-y_0}{x-x_0}), & \theta \in (0, \pi) \\ 3\pi/2 - \arctan(\frac{y-y_0}{x-x_0}), & \theta \in (\pi, 2\pi) \end{cases} \tag{5}$$

3.2. Kalman tracker.

A Kalman tracker is realized here in line with Kalman filter theory. In such a case, the Kalman filter is an estimation and prediction tool and is used here for the estimation of an object's location in the current frame, in accordance with the object's locations in the previous frames. By using the initial state of the object in the scene, the Kalman filter can be performed on it. At first, vectors including the location (position) and speed of the object are constituted, and then their initial values are calculated according to the available frames. The part of the object used as the initial value of its location vector is arbitrary, but it is usually chosen as the object's center of gravity. Moreover, the initial velocity can be determined in association with the location changes and the time difference between two frames.

3.3. The improved MeanShift.

In order to improve the stability of tracking, the Kalman filter was used to estimate the motion state of the target. This method can get an estimation parameter and reduce the scope of searching for matching target. Using Kalman filter to predict location, then replace the center of the target area of the original algorithm. The process combines the target movement information into the MeanShift tracking algorithm, then using the MeanShift algorithm to search the final location of target in the neighborhood of the center of the position which is got in the previous step. The method improves the stability of the original algorithm in the tracking of target with deformation.

After getting the center, the similarity measure in logarithmic coordinates transformation space was used to obtain the change of the scale of the target. Specific methods are described below.

Make polar coordinates transformation to the image of two consecutive frames such as T_n and T_{n+1} in the target window, and receive $I_t(\rho,\theta)$ and $I_{t+1}(\rho,\theta)$. Then build the mapping from image to the rotation invariant space. Because in logarithmic coordinates space, pixel scale and rotation changes are along its two orthogonal axes, a one dimensional mapping function can be got by the integral on θ axis:

$$S_k(r) = \sum_{\theta=0}^{2\pi} I_r(r,\theta) \tag{6}$$

Adopt the normalized correlation function as a mapping function S_k and S_{k-1} to measure the similarity:

$$< S_k \cdot S_{k-1} >= \frac{Cov(S_k, S_{k-1})}{\sqrt{D(S_k)}\sqrt{D(S_{k-1})}} = \frac{\sum(S_k-\mu_k)(S_{k-1}-\mu_{k-1})}{\sqrt{\sum(S_k-\mu_k)^2\sum(S_{k-1}-\mu_{k-1})^2}} \tag{7}$$

S_k is the mapping of the template image's r axis, and μ_k is its mean value.

According to the maximum of the correlation coefficient of target position, axis's translation can be used to calculate the scale of the current frame.

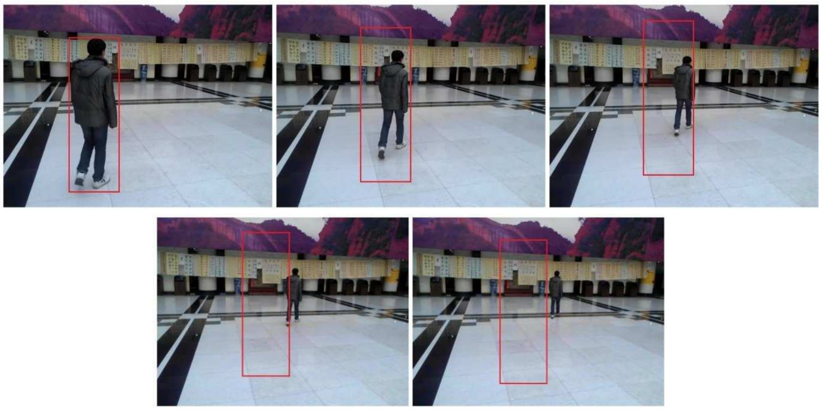
$$W_k = W_{k-1}\exp(dr/k_1) \tag{8}$$

W_k and W_{k-1} are the target scales in two consecutive frames. k_1 is the amplifier parameter, which had been introduced in the formula.

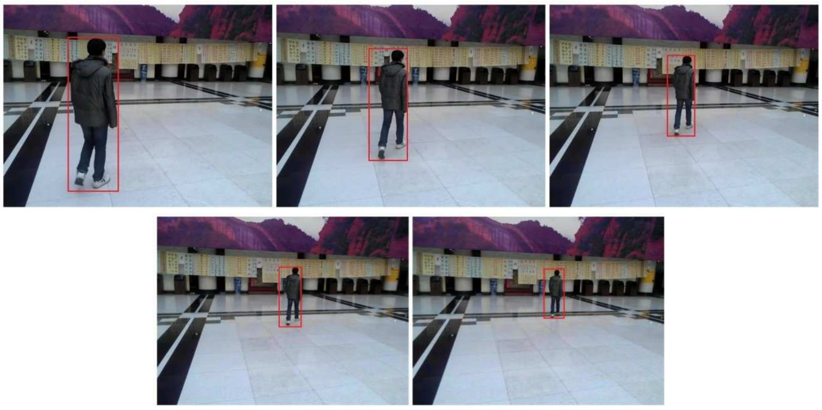
4. **Experimental Process.** For scale change in target tracking, the space location and scale positioning depend on each other. Space positioning is the basis of scale positioning. If space positioning is inaccurate, scale positioning is hard to correct scale. In turn, space positioning is the conditions of the positioning only to extract target feature at the right scale which can achieve effective feature matching.

Tracking processing():
(1) Generate object model q
(2) Compute Kalman initialization parameter
(3) Generate object model polar translation
(4) for (all frames)
(5) Loop until convergence
(6) Generate candidate prior p
(7) Generate candidate model polar translation
(8) Perform MeanShift iterations
(9) Update object centroid
(10) Update object orientation
(11) Update object scale
(12) end Loop

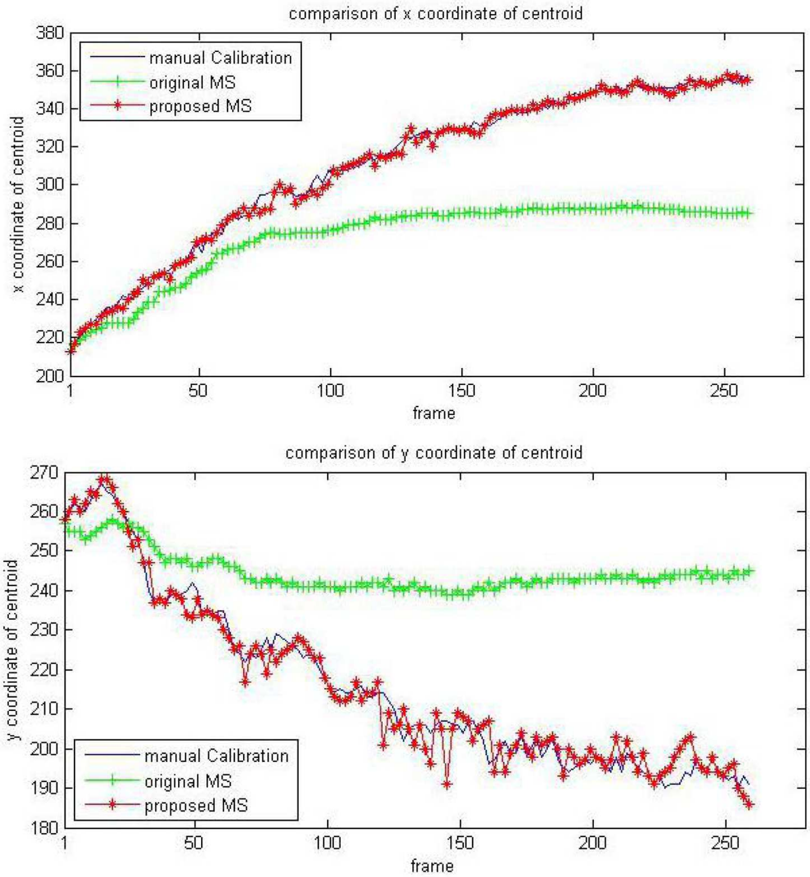
5. **Experimental Results.** Three sets of image sequence were used in this experiment. The first set describes that the target is decreasing and the 51^{st} frame, the 81^{st} frame, the 111^{th} frame, the 141^{st} frame, the 171^{st} frame of them were selected to observe. The second set shows that the target is increasing and the same frames were selected to observe. The last set displays that the target has a change of direction and the 1^{st} frame, the 150^{th} frame, the 290^{th} frame, the 480^{th} frame, the 570^{th} frame were selected at this time. All video sequences are 640×480, and 30 frames in every second. The image depth is 24. All the experiments are in the RGB color space and each color band is equally divided into 16 bins ($16\times16\times16$).



(a) Original MS

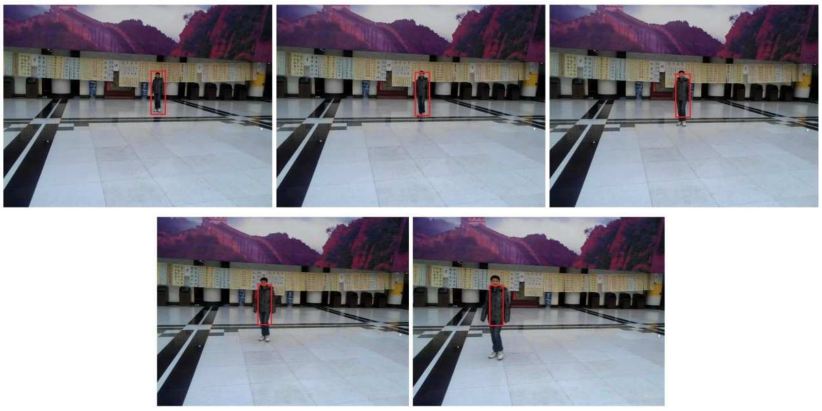


(b) Proposed MS

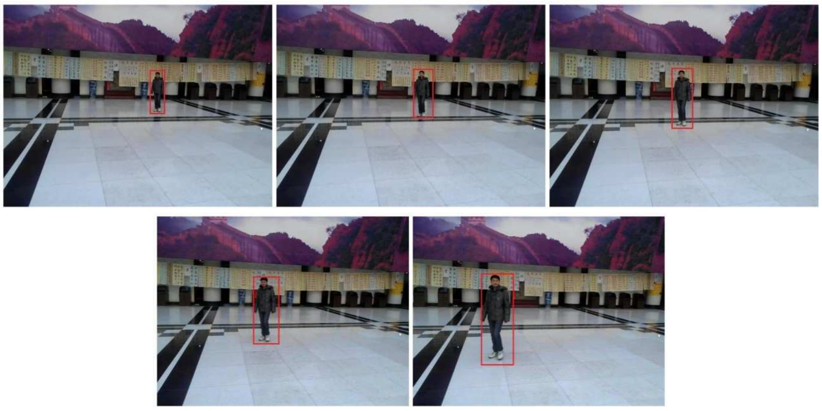


(c) Error comparison of x and y coordinates of centroid in tracking window

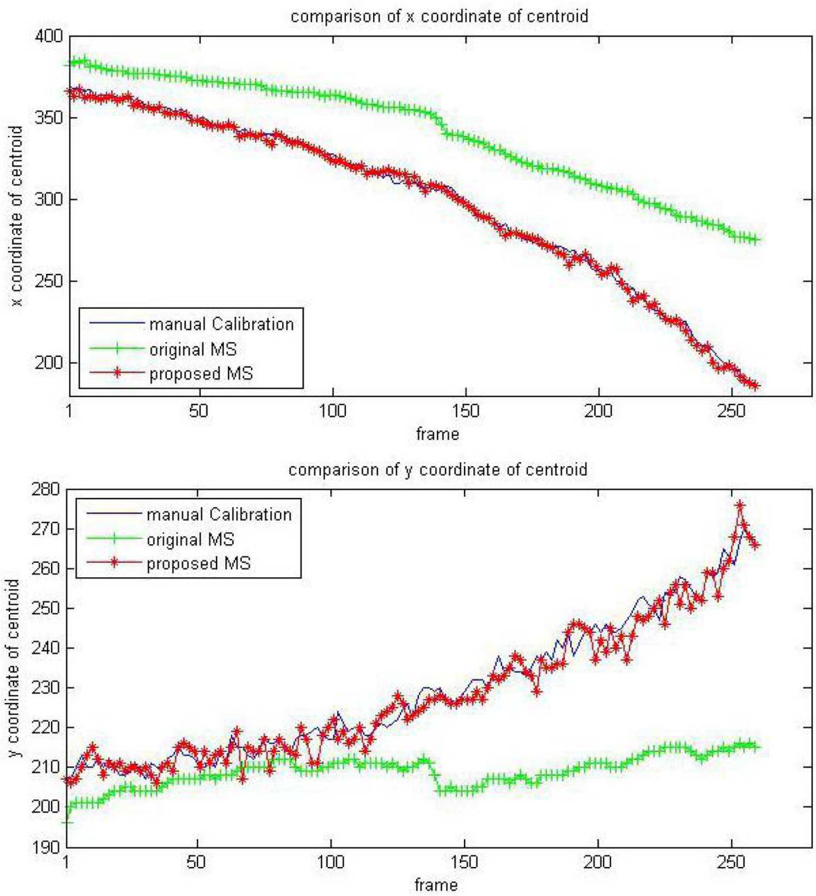
FIGURE 1. Results of tracking decreasing-scale object



(a) Original MS



(b) Proposed MS

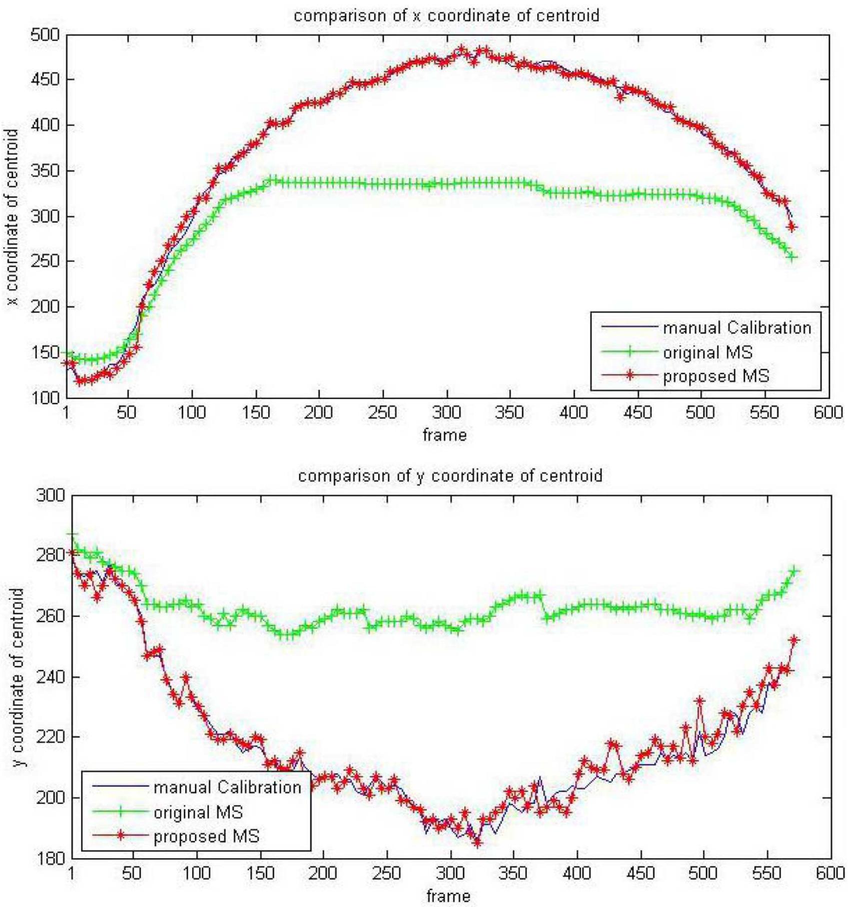


(c) Error comparison of x and y coordinates of centroid in tracking window

FIGURE 2. Results of tracking increasing-scale object

(a) Original MS

(b) Proposed MS



(c) Error comparison of x and y coordinates of centroid in tracking window

Figure 3. Results of tracking with the change of direction

Figure 1 shows the original MeanShift algorithm which is easy to lose the target, when the target is decreasing and the track window has no adjusting accordingly. Relatively the proposed algorithm gets a well tracking result. Figure 2 is the scenario when target is become bigger, two methods all do not lose the target, but the proposed algorithm is more accurate. Figure 1(c) and 2(c) show, the error comparison of x and y coordinates of centroid of tracking window, and the difference can be observed clearly. Figure 3 shows a scenario of the target having a direction change. The experimental results show that original method loses the target soon. While the target is becoming small, many background information entered the track windows. Therefore, it makes the tracking failed. The algorithm in this paper adjusts the tracking window constantly, which can get a better result.

6. **Conclusion.** In this study, an improved MeanShift tracking algorithm was put forward. This method can adjust the tracking window constantly while target size changes. The experimental results show that the algorithm can well track target with scale change and has a certain practical value. In fact, some problems, such as motion blur and target speed, need to be considered in the further research.

REFERENCES

[1] J. Wang and Y. Yagi, Adaptive mean-shift tracking with auxiliary particles, *IEEE Transactions on Systems, Man & Cybernetics, Part B: Cybernetics*, vol.39, no.6, pp.1578-1589, 2009.

[2] X. Yang, H. Li and X. Zhou, Nuclei segmentation using marker-controlled watershed, tracking using Mean-Shift, and Kalman filter in time-lapse microscopy, *IEEE Transactions on Circuits & Systems I: Regular Papers*, vol.53, no.11, pp.2405-2414, 2006.

[3] A. P. Leung and S. Gong, Mean-Shift tracking with random sampling, *Proc. of British Machine Vision Conference*, Edinburgh, England, 2006.

[4] C. Yang, R. Duraiswami and L. Davis, Efficient mean-shift tracking via a new similarity measure, *IEEE Computer Society Conf. on Computer Vision & Pattern Recognition (CVPR 2005)*, San Diego, USA, 2005.

[5] A. Yilmaz, Object tracking by asymmetric kernel mean shift with automatic scale and orientation selection, *IEEE Conf. on Computer Vision & Pattern Recognition*, pp.1-6, 2007.

[6] J. Wang, T. Bo, Y. Xu and M. Cohen, Image and video segmentation by anisotropic kernel mean shift, *European Conf. on Computer Vision*, pp.238-249, 2004.