Moving Target Tracking Algorithm Based on Improved Optical Flow Technology

Cheng Yi, Cui Liyun* and Luo Chunguang

School of Electronic Engineering and Automation, Tianjin Polytechnic University, Tianjin, 300387, China

Abstract: A method of moving target tracking based on optical flow technology is proposed in this paper. Extract the feature points of the moving target by the SIFT algorithm, and replace the interesting points in traditional optical flow method with the extracted feature points, then track the moving target through these points. The proposed moving target tracking method has better tracking precision for the SIFT algorithm is invariant to rotation and scale change of moving target .At the same time, the optical flow estimation is obtained using the pyramid LK optical flow algorithm so that the computation of the tracking process is reduced, the accuracy is improved and the movement target tracking has a real-time performance .In this paper, three groups of experiments was conducted, a set of contrast experiments under the proposed method and the traditional optical flow method, a set of validation experiments of single moving target with some shelter and another set of validation experiments of multiple moving target with irregular shelters .The experimental results verify that the the proposed method has a good accuracy, robustness and real-time performance.

Keywords: Moving target, optical flow technology, pyramid LK, SIFT algorithm, target tracking.

1. INTRODUCTION

The research and application of target tracking algorithm is an important branch of computer vision. The goal of moving target tracking is to find and locate the mobile object that we are interested in each frame from the video by the registration of the relevant continuous frame [1]. In the process of movement target tracking, the motion parameters of the target such as speed, acceleration, position, trajectory can be acquired, a further analysis and processing of these information can advance the understanding of the moving target behavior, which is the basis of the advanced application. These advanced applications such as video monitoring, virtual reality, robot navigation, traffic detection and missile guidance, human-computer interaction system, remote video conferencing, and virtual reality technology, are widely used in many fields of civilian and military [2].

The research method of moving target tracking can be divided into two kinds:the methods based on motion analysis and registration. Registration methods such as reference [3] based on SIFT feature matching method can gain higher positioning accuracy, but the amount of calculation is too large to meet the real-time requirements .The most classic method based on movement analysis is the optical flow method .The moving target tracking based ontraditional optical flow method has a good real-time performance but bad positioning accuracy because of the limits of the method itself [4]. So in this paper, SIFT features extracted from the moving targets is to improve the performance of the method based on optical flow method, because the SIFT features is invariant to the translation, rotation, scale zoom, illumination change, also maintains stability in some degree when there exist perspective change, affine transform and noise. It is designed To describe the target with its SIFT features, and calculate optical flow field of these characteristic points and finally track them, so that the problems that the traditional optical flow sensitive to illumination changes, and that of tracking fails caused by the suddenly change of the motion target's position or the serious deformation of the motion target can be effectively resolved. The optical flow target tracking algorithms proposed in reference [5] based on scale invariant features (SIFT features) need to manually choose the target area in the first frame of video, it can't automatically complete moving target detection and tracking of uncertainty number in actual application .In this paper the tracking strategy is improved, first of all, moving targets are identified in the view by optical flow method, and their SIFT features are extracted, then, these targets are described with their SIFT features in subsequent video frames, at last, the optical flow estimation of these feature points of target tracking is calculated .In order to improve the tracking speed the pyramid Lucas – Kanade (LK) is adopted when calculate the optical flow estimation , the proposed algorithm has been improved greatly both in tracking speed and accuracy.

2. THE PROPOSED ALGORITHM

The proposed moving target tracking algorithm is shown in Fig. (**1**). In which, the optical flow field of the moving target in the video is calculated firstly, then the moving object is detected utilizing the threshold filtering, next SIFT feature points of the moving target are extracted and tracked, last coordinate positions of these points in the frame are located.

*Address correspondence to this author at the School of Electronic Engineering and Automation, Tianjin Polytechnic University, Tianjin, 300387, P.R. China; Tel: +8618722513086; E-mail: 641294365@qq.com

1874-4443/15 2015 Bentham Open

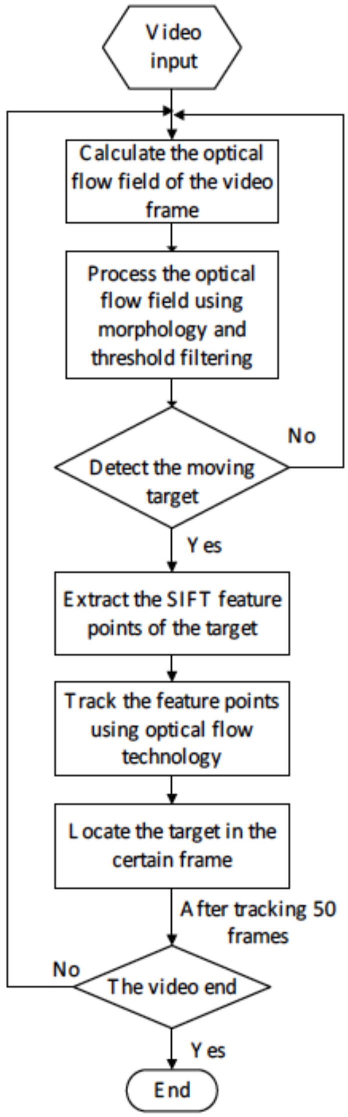


Fig. (1). The algorithm flow chart.

In the proposed algorithm, the pyramid LK method is adopted to calculate the optical flow field of the frames in order to reduce calculating time. Closing operation is used in morphology process. Mean filtering is utilized in threshold filtering, the area whose optical flow vector is larger than the mean value of the whole optical flow vectors is identified as the moving target.

2.1. Pyramid Lucas – Kanade (LK)

(x, y) and $(x+u, y+v)$ are defined as position coordinates of a point in the two adjacent images. In the optical flow estimation, there are two assumptions, the color constancy and little movement. The color constancy make the assumption that a point in two adjacent images have the save color or brightness, its mathematics format is show in Equation (1), H and I are two adjacent images. The little movement means that the position transformation is little, that is u and v are less than 1 pixel.

$$H(x,y) = I(x+u, y+v) \tag{1}$$

Let's take the Taylor series expansion of $I(x+u,y+v)$,the higher order terms is abandoned for it goes to zero:

$$I(x+u, y+v) \approx I(x,y) + \frac{\partial I}{\partial x}u + \frac{\partial I}{\partial y}v \tag{2}$$

Combine Equation (1) and Equation (2):

$$\begin{aligned} 0 &= I(x+u,y+v) - H(x,y) \\ &\approx I(x,y) + \frac{\partial I}{\partial x}u + \frac{\partial I}{\partial y}v - H(x,y) \\ &\approx I_t + \nabla I \cdot [\frac{\partial x}{\partial t}, \frac{\partial y}{\partial t}] \end{aligned} \tag{3}$$

Where,

$$I_t = I(x,y) - H(x,y) \tag{4}$$

$$\nabla I = [\ \frac{\partial I}{\partial x} \quad \frac{\partial I}{\partial y}\ ] \tag{5}$$

$$u = \frac{\partial x}{\partial t} \tag{6}$$

$$v = \frac{\partial y}{\partial t} \tag{7}$$

In the limit, as u and v go to zero, the Equation (3) becomes exact:

$$I_t + +\nabla I \cdot [\frac{\partial x}{\partial t}, \frac{\partial y}{\partial t}] = 0 \tag{8}$$

In Equation (4), it containes two unknowns in one equation , so additional constraints must be added to solve it.

LK algorithm [6, 7] is based on the local constraints with the assumption that the pixel's neighbors have the same (u,v), and a neighborhood of 3×3 pixels is usually adopted, which will give 9 equations per pixel. Then the calculation of optical flow is converted to the following equation:

$$\begin{bmatrix} I_x(p_1) & I_y(p_1) \\ I_x(p_2) & I_2(p_2) \\ \vdots & \vdots \\ I_x(p_9) & I_y(p_9) \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = - \begin{bmatrix} I_t(p_1) \\ I_t(p_2) \\ \vdots \\ I_t(p_1) \end{bmatrix} \tag{9}$$

We can write Equation (9) as Equation (10):

$$Ad = b \tag{10}$$

Finally, least squares method is used to solve the equation :

$$A^T Ad = A^T b \tag{11}$$

$A^T A$ is a 2×2 matrix:

$$A^T A = \begin{bmatrix} \sum I_x^2 & \sum I_x I_y \\ \sum I_y I_x & \sum I_y^2 \end{bmatrix} \tag{12}$$

$A^T b$ is a 2×1 matrix:

$$A^T b = -\begin{bmatrix} \sum I_x I_t \\ \sum I_y I_t \end{bmatrix} \quad (13)$$

All sum computation in Equation (12) and Equation (13) include all points in the neighborhood of 3×3. Finally the optical flow vector can be obtained from Equation (13).

In order to further enhance the accuracy of Lucas - Kanade method and computing speed, in the practical application, image gaussian pyramid is combined with Lucas - Kanade algorithm, with the hierarchical strategy, the image is decomposed into different resolutions from the coarse to the fine, the resolution becomes lower and lower with the increase of level of resolution, and the result in coarse scales will be used as the initial value of the next scale to calculate the flow rate of the image sequence on the different resolution, this is the effective method to the calculation of large movement rate, as shown in Fig. (**2**).

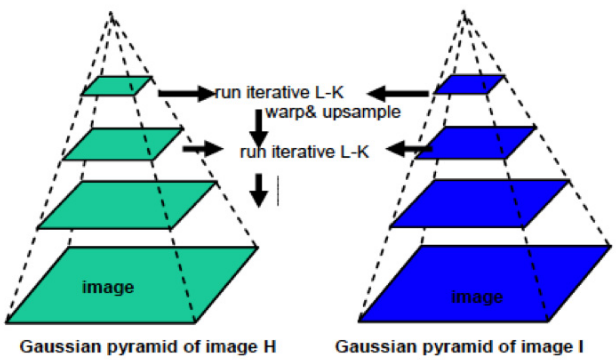


Fig. (2). LK pyramid optical flow.

2.2. SIFT Feature Extraction of Moving Objects

Optical flow method can perform better tracking robustness than other algorithms when the target motion speed is not stable, and SIFT feature tracking algorithm has higher accuracy than other tracking algorithm when there are large displacement between adjacent frames [8]. So in order to enhance tracking accuracy and robustness of the detected moving target, to extract SIFT feature points of moving targets instead of the whole target, in the subsequent frames, it is only to calculate the optical flow of these feature points,thus, the computing speed can be improved.

2.2.1. Detection Extreme Points of Scale Space

A scale space function of a two-dimensional image is defined based on scale space theory as follows:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (14)$$

There, $G(x, y, \sigma)$ is the variable scale gaussian function:

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad (15)$$

(x, y) is a spatial coordinate representing the pixel position on the image, symbol * denotes convolution computation, σ is a scale space factor, the smaller the σ is the less smooth of the image is. Large scale corresponds to the overview features of the image and small scale corresponds to the detail features of the image.

In order to detect stable key points effectively in scale space, the differential gaussian scale space (DOG scale - space) is presented [9]. It is generated by the convolution of gaussian difference nuclear of different scales and images.

$$D(x, y, \sigma) = \big(G(x, y, k\sigma) - G(x, y, \sigma)\big) * I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (16)$$

2.2.2. Location of the Extreme Value Point

The positions of the key points and scales (achieve subpixel accuracy) is to be determined accurately by fitting three-dimensional quadratic function. Unstable Points of low contrast or edge response is to be removed [10]. Fitting function at key points is showed in Equation (17):

$$D(X) = D + \frac{\partial D^T}{\partial X} X + \frac{1}{2} X^T \frac{\partial^2 D}{\partial X^2} X \quad (17)$$

Derivate and set the Equation (17) to zero, extreme value points can be obtained：

$$\hat{X} = -\frac{\partial^2 D^{-1}}{\partial X^2} \frac{\partial D}{\partial X} \quad (18)$$

Corresponding to the extreme value point:

$$D(\hat{X}) = D + \frac{1}{2} \frac{\partial D^T}{\partial X} \hat{X} \quad (19)$$

The value of $D(\hat{X})$ is useful to eliminate the low contrast and unstable feature points, the points with value $D(\hat{X}) < 0.03$ will be treated as unstable points of low contrast and will be eliminated. At the same time, precise location and scale of the feature points can be obtained in this procedure.

2.2.3. Allocating Direction for Key Points

In order to make the descriptor invariant to rotation, the partial feature of the image is needed to allocate a direction for each key point. Using the gradient of neighborhood pixels of key point and distribution features of the direction, modulus gradient and direction can be obtained as follows:

$$m(x, y) = \sqrt{\begin{matrix}(L(x+1, y) - L(x-1, y))^2 \\ + (L(x, y+1) - L(x, y-1))^2\end{matrix}} \quad (20)$$

$$\theta(x, y) = \tan^{-1}\left(\frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)}\right) \quad (21)$$

In which, $m(x, y)$ and $\theta(x, y)$ are the gradient module and orientation of the key point (x,y) respectively. L is the scale of the key point. Sampled within the neighborhood window centered at key points, using histogram to estimate gradient directions of neighbor pixels. The range of gradient histogram is 0 ~ 360 degree, each 10 degrees belongs to one direction. The peak value of the histogram denotes the main direction of the key points in the neighborhood gradient, which is the direction of the key point.

2.2.4. Generate Descriptors for Key Points

Firstly, rotating coordinate to the direction of the key point to ensure the rotation invariance. secondly, selecting 8×8 window centered at the key point, then to calculate gradient direction histogram at eight directions on each small pieces of 4×4, and draw the accumulative value of each gradient direction, a seed point can be formed .In this way a point is composed of 2×2 seed points, every seed point has eight direction vector information. This idea of neighborhood directional information coalition enhances the algorithm's ability to overcome noise, at the same time provides good fault tolerance for the feature matching with positioning error.

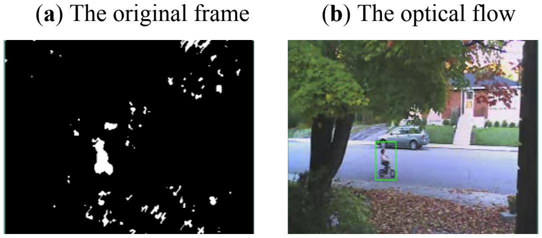
3. EXPERIMENTS OF THE PRESENTED ALGORITHM

In this paper, three groups of experiments were conducted .The first one is a contrast experiment between the algorithm presented in this paper and the traditional optical flow algorithm in the case of a single moving object covered by some shelter .The second experiment is performed to further verify the proposed algorithm about tracking a single moving target with covered cases on the performance of real-time and accuracy .The third experiment is conducted to verify the proposed algorithm about tracking multiple moving targets with covered cases on the performance of real-time and accuracy .In this paper, MATLAB software is used to program the algorithms, the test environment is a computer with a 2.5 GHz CPU, an operating system of Win8 and a Matlab2014 software platform . The image resolution of the video used in experiment 1 is 240×320, and 600×400 in videos used in the second and the third experiment.

3.1. The Results of the First Experiment

The results of key steps in the implementation process of the algorithm proposed in this paper are shown in Fig. (**3**) to Fig. (**6**). Fig. (**5**) is the tracking result of the traditional optical flow method.

(**a**) The original frame (**b**) The optical flow



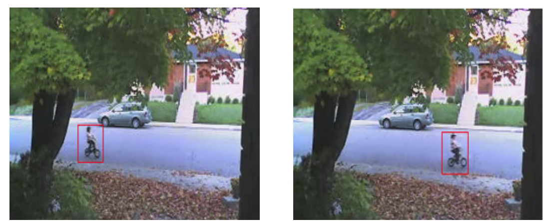
(**c**) Morphological Processing (**d**) Moving target detected

Fig. (3). Target detecting based on optical flow.

Extract SIFT feature point sets of the target area after the moving target is detected and the result is shown in Fig. (**4**), the feature points extracted using SIFT algorithm are denoted by blue circles in the figure.

Fig. (4). SIFT feature points of the moving target.

After the moving target is detected and its SIFT feature points are extracted, track the feature points of the target instead of optical flow point using optical flow technology . Part of tracking results of the proposed algorithm are shown in Fig. (**6**). Fig. (**5**) shows parts of the tracking results of the traditional optical flow method.



(**a**) The 30th frame (**b**) The 85th frame

(**c**) The 98th frame (**d**) The 102nd frame

Fig. (5). Tracking results of the traditional optical flow method.

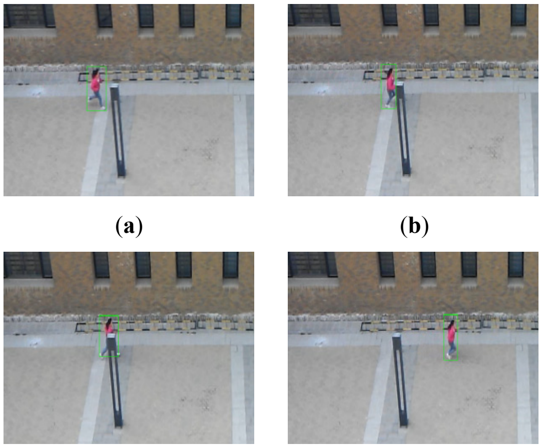
(**a**) The 30th frame (**b**) The 85th frame

(**c**)The 98th frame (**d**) The 102nd frame

Fig. (6). Tracking results of the modified optical flow method.

3.2. The Results of the Second Experiment

In order to verify the tracking stability of the proposed algorithm for non rigid objects with covered conditions, the algorithm was tested using another video .The tracking results are shown in Fig. (**7**).

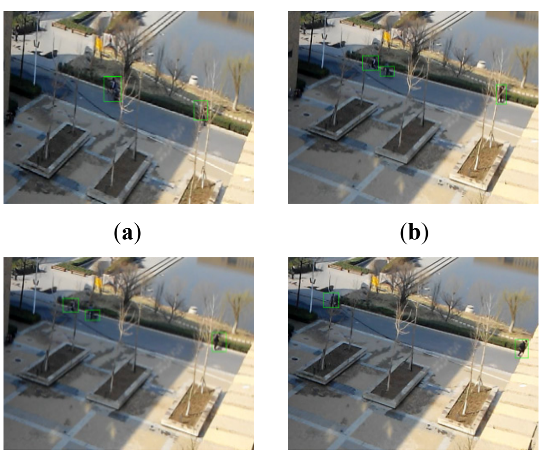


(**a**) (**b**) (**c**) (**d**)

Fig. (7). Tracking result of the second experiment.

3.3. The Results of the Third Experiment

There usually exist multiple moving targets in a natural environment, and the environment will be relatively complicated, which makes it more difficult to track targets and the requirement to the algorithm will be higher. So that such environment is more persuasive to verify the reliability of the algorithm. Fig. (**8**) shows tracking results of the presented algorithm in multi-target with covered case.



(**a**) (**b**) (**c**) (**d**)

Fig. (8). Tracking results of the third experiment.

4. ANALYSIS OF THE RESULTS

It can be seen from the tracking results of the 98th frame and the 102nd frame in Fig. (**6**) that the traditional optical flow method is not stable when the target occlusion happens .In the 98th frame targets has small covered area, tracking box can't pinpoint the target area, in 102nd frame targets have large covered block tracking failed .The contrast is visible in this paper, the proposed algorithm has been improved on the accuracy of target tracking. Fig. (**7**) verified the algorithm again that it can achieve the tracking completely when there exists occlusion. Tracking results in Fig. (**8**) show that the proposed algorithm can complete all of the moving targets tracking in the vision scope, and can still correctly track targets even if they are sheltered by shelters .The accuracy and computing time of the video used in the first experiment based on the improvement algorithm and the traditional optical flow tracking method has been calculate respectively, and compared their accuracy and time consumption in Table **1**. The accuracy and time consumption of the second and the third verified experiment are listed in Table **2**.

Table 1. Comparison of traditional method and the modified method.

	Traditional optical flow tracking	**The modified optical flow tracking**
Tracking accuracy	86%	99%
Target tracking time consumption (ms/frame)	0.35	0.23

Table 2. Accuracy and time consumption of verified experiments.

	The second experiment	**The third experiment**
Time consumption (ms/frame)	0.32	0.33
Tracking accuracy	99%	96%

It is easy to know from Table **1** that the modified algorithm improved the tracking accuracy by 13%, and reduce 0.12 ms on tracking time consuming than the conventional optical flow tracking algorithm. This is because the goal SIFT feature points was adopted to replace the optical flow point to track the target. Another reason is that the image pyramid was used to reduce time .As video resolution used in the second experiment and the third experiment is higher than that in the first experimental video so the time consumption increased as shown in Table **2**, but is still lower than the traditional optical flow tracking. The tracking accuracy of the third experiment decreased because of the complicated environment, but it is still within the acceptable range of application.

CONCLUSION

Through comparison between the improved algorithm and the traditional optical flow tracking algorithm on tracking accuracy and time consumption proves that the improved algorithm has better robustness and real-time performance .It is proved by the second experiment and the third experiment that the improved algorithm can achieve tracking single moving target and multiple moving targets successfully in case of occlusion.

CONFLICT OF INTEREST

The authors confirm that this article content has no conflict of interest.

ACKNOWLEDGEMENTS

This study is supported by the National Natural Science Foundation of China (Project No. 61203302) and the Tianjin Research Program of Application Foundation and Advanced Technology (14JCYBJC18900).

REFERENCES

[1] L. Ni, and W. Wenchao, "A moving object tracking method based on optical flow computing", *Information Technology and Informatization*, vol. 121, no. 03, pp. 116-117, 2014.

[2] W. Zhenjie, and M. Xiaobo, "Improved optical flow estimation algorithm based on frame difference", *Computer Engineering and Applications*, vol. 49, no. 18, pp. 200-203, 2012.

[3] T. Qiu,"Visual object tracking algorithm based on sift features", *Journal of Shangrao Normal College*, vol. 31, no. 3, pp. 36-41, 2011.

[4] Z. Chi, X. Zhiyu, and L. Jilin, "Monocular vision odometry based on the fusion of optical flow and feature points matching", *Journal of Zhejiang University (Engineering Science)*, vol. 48, no. 2, pp. 279-284, 2014.

[5] W. Yin, L. Liangfu, X. Zhangshu, and L. Shigang, "Optical flow motion tracking algorithm based on sift feature", *Computer Engineering and Applications*, vol. 49, no. 15, pp. 157-161, 2013.

[6] B. D. Lucas, and T. Kanade, "An iterative image registration technique with an application to stereo vision", *International Joint Conference on Artificial Intelligence*, vol. 2, pp. 674-679, 1981.

[7] D. Guo, A. L. van de Ven, and X. Zhou, "Red blood cell tracking using optical flow methods", *IEEE Journal of Biomedical and Health Informatics*, vol. 18, no. 3, pp. 991-998, 2014.

[8] D. G. Lowe, "Distinctive image features from scale-invariant keypoints", *International Journal of Computer Vision (IJCV)*, vol. 60, no. 2, pp. 91-110, 2004.

[9] S. Gauglitz, and T. Höllerer, "Evaluation of interest point detectors and feature descriptors for visual tracking", *International Journal of Computer Vision*, vol. 94, no. 3, pp. 335-360, 2011.

[10] Y. Motai, S. K. Jha, and D. Kruse, "Human tracking from a mobile agent: optical flow and kalman filter arbitration", *Signal Processing: Image Communication*, vol. 27, pp. 83-95, 2012.

Received: May 26, 2015 Revised: July 14, 2015 Accepted: August 10, 2015