Multi-target tracking algorithm by corner feature for video stream

Zhaoxia Fu

Party School of Shanxi Provincial Committee of the C.P.C, Taiyuan, 030006, China

email: fzx2005@163.com

Keywords: Moving Target Tracking; Corner Feature; Matching Optimization; Block State

Abstract. In the field of computer vision research, moving target tracking of sequence images is not well solved, and is an old, important and challenging task. Fast and accurate completion of moving target tracking is very an important and difficult problem. In a dynamic environment, moving target tracking is often influenced by various factors, such as weather changes, shadows, occlusion, background confusion. This paper presents a multi-target tracking algorithm by corner feature for video stream, which can solve the problems of video multi-objective motion tracking. The improved Harris operator can extract some even and stable corners, and the matching optimization reduces the amount of false matching points, so they improve the accuracy of target tracking. Experimental results show that the algorithm can complete a stable matching under a variety of complex conditions, and achieve a steady target tracking under a small part of the block state.

Introduction

The technology of moving target tracking as one of the core technologies in the field of computer vision has an increasingly important role in a number of key technologies, such as virtual reality, robotics, intelligent weapon systems [1-4]. Target tracking is obtained by motion parameters of moving objects, such as position, velocity, acceleration and trajectory, etc., for further processing, analysis, and understanding the behavior of moving objects in order to complete higher level tasks. In recent years, the problems of target tracking have been the climax of the study. Ref. [5] divides the target tracking problem into many types, such as single-camera visual tracking and multi-camera visual tracking, rigid tracking and non-rigid tracking, target tracking when the camera is at rest and target tracking when the camera is moving, etc. The practical methods of target tracking are mainly based on the area, contour, centroid, correlation, feature recognition. Feature-based tracking method is the research focus of the current tracking method.

Corner information in the target feature is a comprehensive description of the target, which extraction and tracking of the corners reflect static characteristics and dynamic characteristics of the target. Corners generally are understood to be the larger curvature points of the object edge in an image, and they have two key characteristics, which are located in the larger curvature points and the points on the edge of the object. Corner tracking has two main key: corner extraction and corner matching. Corner extraction in general is mainly based on the edge and the grayscale, and corner matching is based on relevant corner template and corner feature. Experimental results show that target tracking using a certain amount of corners is effective in general. But too many corners will increase computation time, and the accuracy and reliability of corner extraction will affect corner matching in the multi-frame images. In order to accurately extract the corners and reduce false matching corners, and improve operational efficiency, and also achieve effectively target tracking, motion analysis and recognition, we use the improved Harris operator [6] to extract feature points of moving target, and complete the tracking of multiple moving targets by feature matching and matching optimization in the video images.

Moving target region detection and multi-target labeling

The region detection algorithm of moving target includes three steps, such as initial background extraction, motion region extraction, background updating, as follows:

(1) Initial background extraction

Select the initial five frame of a video sequence, and use the background modeling algorithm

based on the mean to extract the initial background of video image sequence.

(2) Motion region extraction

Use the background subtraction algorithm [7] to extract the moving object, which is that the current frame and the extracted background make the differential. The operation can be described as:

$$f_b(x,y,t) = \begin{cases} 1, & if \left| f(x,y,t) - B(x,y,0) \right| > T \\ 0, & others \end{cases} \tag{1}$$

where T is the threshold value, $f(x,y,t)$ and $B(x,y,0)$ are respectively the current frame at t time and the background image.

(3) Background Updating

To reduce the amount of computation and save memory, you can update the background by the following equation:

$$B(x,y,t+1) = \lambda f(x,y,t) + (1-\lambda)B(x,y,t) \tag{2}$$

where $f(x,y,t)$ and $B(x,y,t)$ are respectively the current frame and its background image at t time, λ is called refresh rate, usually taken 0.05.

The most common algorithms of the connected components are the recursive algorithm and the sequential algorithm. The recursive algorithm is inefficient, especially for multiple targets in the binary image, and you need to repeatedly scan the image. The sequential algorithm usually only scans the image twice, and its processing speed is much faster than the recursive algorithm. So the sequential algorithm usually is used to complete the labeling, the process is as follows:

(1) An image is scanned from left to right, top to bottom.

(2) If the pixel is "1", those pixels located in its eight neighborhood are given the same reference numerals, and then the "l" pixels in the eight neighborhood are denoted by the same reference numerals, otherwise they are assigned new tags and the tags are input the equivalent table.

(3) If there are not the scanned pixels in the image, it returns step 1.

(4) Find the minimum tag in each equivalent collection of the equivalent table.

(5) Scan the whole image, and substitute the minimum tag in the equivalent table for each tag.

Improved Harris corner algorithm

In this paper, the improved Harris operator extracts some corners, so as to obtain sub-pixel accuracy. Harris corner operator makes corner extraction, as follows:

$$M = \begin{pmatrix} \left(\dfrac{\partial I}{\partial x}\right)^2 & \dfrac{\partial I}{\partial x}\dfrac{\partial I}{\partial y} \\ \dfrac{\partial I}{\partial x}\dfrac{\partial I}{\partial y} & \left(\dfrac{\partial I}{\partial y}\right)^2 \end{pmatrix} \tag{3}$$

where $I(x,y)$ is the gray value of an image. If two eigenvalues of the matrix M have the local maximum, the point is the feature one. In order to avoid calculating the eigenvalues, the actual evaluation function is:

$$R = \det M - k \cdot (traceM)^2 \tag{4}$$

where det expresses matrix determinant. R is positive in the corner region, negative in the edge regions, and small in the flat region. The value of the threshold is determined by the amount of feature points. According to Ref. [8], $k=0.06$, which is for the elimination of noise interference. The calculation of M is used in the differential Gaussian function whose variance is 1.

Feature matching and matching optimization

To enable precise matching, the cross-correlation method of gray window as the center of feature points can be used to match initially. Determining whether two feature points match is based on

Local Area Correlation Coefficient of two images (Local Area Correlation Coefficient, LACC). So LACC is defined as:

$$C(p,q) = \frac{\sum_{i=-n}^{n}\sum_{j=-m}^{m}\left[I_1(u_1+i,v_1+j) - \overline{I_1(u_1,v_1)}\right] \times \left[I_2(u_2+i,v_2+j) - \overline{I_2(u_2,v_2)}\right]}{(2n+1)(2m+1)\sqrt{\sigma^2(I_1) \times \sigma^2(I_2)}}$$

(5)

where $\overline{I_k(u,v)} = \sum_{i=-n}^{n}\sum_{j=-m}^{m}\frac{I_k(u+i,v+j)}{\left[(2n+1)(2m+1)\right]}$ is the gray mean as the center of the point (u,v) and

the size of (2n+1) (2m+1) in the neighborhood region of the image I_k (k=1,2). The correlation coefficient C ranges from $[-1, 1]$. If C is equal to -1, two correlation windows are not similar. If C is equal to 1, two correlation windows are entirely identical. If one pair of feature points are matching, the value of LACC is necessarily bigger by Eq. (5).

Experimental results and analysis

Choose the length 16s of indoor surveillance video as an experimental material. The tracking algorithm runs on Interl (R) Celeron (R), Main frequency 2.60GHz, Memory 2.00G of Pc machine. In the experiment, there are two targets. According to the proposed algorithm, target detection is executed for video stream by the background subtraction method. After each target is respectively labeled, the method of improved Harris corner extracts some stable corners. The LACC method is used to obtain the initial matching points, and then the RANSAC method is used to eliminate mismatching points, so as to get tracking results. Fig. 1 shows the tracking results of corner feature matching.
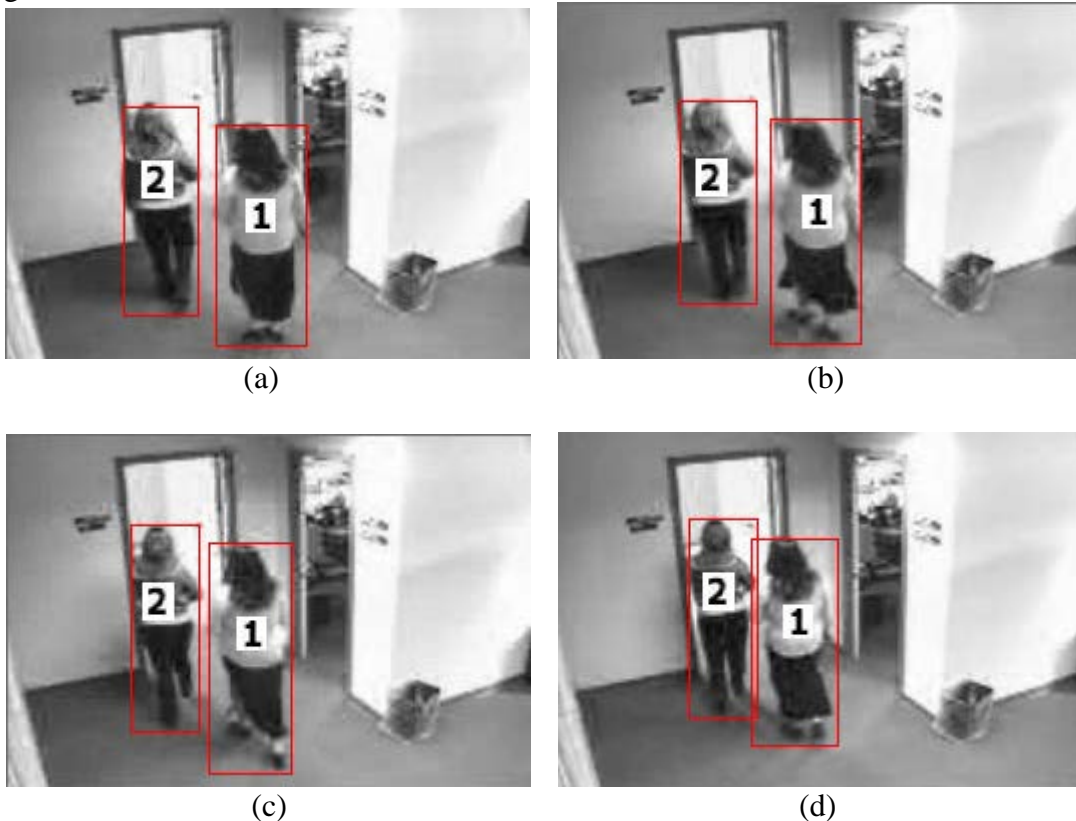

(a)

(b)

(c)

(d)

<div align="center">(e)   (f)</div>

<div align="center">Fig.1. The tracking results</div>

Conclusion

The multi-target tracking algorithm by corner feature is proposed in this paper. The improved Harris operator is used to extract feature points of moving targets, so the distribution of extracted feature points is stable, and its calculation is also simple. By the matching optimization, the amount of false matching points is reduced, and tracking accuracy is also improved. Tracking experiments show that the algorithm can complete a stable matching under a variety of visual angle change, rotation, affine transformation, lighting change, and also achieve a steady target tracking under a small part of the block state.

References

[1] Coifman B, Beymer D, Mclauchlan P, Malik J. A real-time computer vision system for vehicle tracking and traffic surveillance. Transportation Research Part C, 1998, 6(4): 271-288.

[2] Magee D. Tracking multiple vehicles using foreground, background and motion models. Image and Vision Computing, 2004, 22(2): 143-155.

[3] Dubuisson-Jolly M, Lakshmanan S, Jain A K. Vehicle segmentation and classification using deformable templates. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1996, 18(3): 293-308.

[4] Koller D, Daniilidis K, Nagel H. Model-based object tracking in monocular image sequences of road traffic scene. International Journal of Computer Vision, 1993, 10(3): 257-281.

[5] Hou Zhi-Qiang, Han Chong-Zhao. A Survey of Visual Tracking. Acta Automatica Sinica, 2006, 32(4): 603-617.

[6] Harris C, Stephens M. A Combined Corner and Edge Detector. In: Proceedings Fourth Alvey Vision Conference, Manchester, UK, 1988, 147-151.

[7] McKenna S et al. Tracking groups of people. Computer Vision and Image Understanding, 2000, 80(l): 42-56.

[8] Schmid C, Mohr R, Bauckhage C. Comparing and evaluating interest points. The Sixth International Conference on Computer Vision, 1998, 230-235.