http://ansinet.com/itj

ITJ

ISSN 1812-5638

INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 13 (11): 1863-1867, 2014
ISSN 1812-5638 / DOI: 10.3923/itj.2014.1863.1867

A Moving Objects Detection Algorithm Based on Three-Frame Difference and Sparse Optical Flow

Guo-Wu Yuan, Jian Gong, Mei-Ni Deng, Hao Zhou and Dan Xu
College of Information Science and Engineering, Yunnan University, Kunming 650091, China

Abstract: Moving objects detection is an important research of computer vision. Optical flow method is a main way, but it is limited to use because of its complexity. A moving objects detection algorithm based on three-frame difference and optical flow is proposed. The calculation of optical flow is simplified. Harris corners are detected and then only the corners are selected to compute optical flow information, which reduce the algorithm's complexity. Because the detected moving target area is not complete, three-frame difference method is introduced as a supplement. The experimental results show that the algorithm can achieve real-time and has better results than anyone of these two separate algorithms.

Key words: Moving objects detection, optical flow, frame difference, corner detection

INTRODUCTION

Moving objects detection extracts changed regions from a video image sequence, which is a key research of computer vision. The technology is widely used in visual surveillance, intelligent transportation system, motion analysis, industrial vision, etc.

There are three methods in the moving objects detection and the methods are frame difference, background subtraction and optical flow. Optical flow gives the link between the change of image gray and two-dimensional velocity field and optical flow field reflects the direction and speed of moving pixels. According to the distribution characteristics of optical flow field, moving objects can be extracted. The method has high detection accuracy, but can not obtain the accurate outline of the moving objects. It is so complex to calculate an optical flow field that it is difficult to achieve real-time detection (Huang *et al*., 2009).

However, optical flow method can detect moving objects without any video information and can be preferably combined with moving objects tracking, so it is a good algorithm in practical application. If all pixels' optical flows in an image are calculated, the amount of computation will be very large; if a small number of pixels with certain characteristics are selected to calculate optical flow, the complexity can effectively reduced. At the same time, as long as these selected pixels are representative, it can be guaranteed to obtain relatively complete motion information of moving objects.

Only some representative pixels are selected, so the detected moving area is not complete. Three-frame difference, which is simple, is employed as a supplement to obtain relatively complete moving object areas.

A moving objects detection method is proposed in this study and it combines three-frame difference and optical flow. The method selects Harris corners as representative pixels to calculate sparse optical flow using Lucas-Kanade algorithm and then three-frame difference method is introduced as a supplement. The experiments indicate the method enables optical flow method reach real-time and is better than the results obtained by one of the two alone methods.

HARRIS CORNER DETECTION

Corners are very important points with local feature. Harris gave a definition of the most widely used corner (Harris and Stephens, 1988) and then this corner was called Harris corner.

Harris and Stephens (1988) calculated the second derivative of all pixels' gradation in an image. Because the second derivative image was obtained from the two-dimensional Hessian matrix, this image was called Hessian image:

$$H(p) = \begin{bmatrix} \frac{\partial^2 I}{\partial x^2} & \frac{\partial^2 I}{\partial x \partial y} \\ \frac{\partial^2 I}{\partial y \partial x} & \frac{\partial^2 I}{\partial y^2} \end{bmatrix} \tag{1}$$

Harris and Stephens (1988) corners were detected by the Hessian image's auto-correlation matrix M (x, y) of the small window around each pixel.

Corresponding Author: Guo-Wu Yuan, College of Information Science and Engineering, Yunnan University, Kunming 650091, China

$$M(x,y)=\begin{bmatrix} M_{00} & M_{01} \\ M_{10} & M_{11} \end{bmatrix} \quad (2)$$

Where, in, the values M_{00}, M_{01}, M_{10} and M_{11}, were:

$$M_{00}=\sum_{-K\le i,j\le K} w_{i,j} I_x^2(x+i, y+j) \quad (3)$$

$$M_{01}=\sum_{-K\le i,j\le K} w_{i,j} I_x(x+i, y+j) I_y(x+i, y+j) \quad (4)$$

$$M_{10}=\sum_{-K\le i,j\le K} w_{i,j} I_x(x+i, y+j) I_y(x+i, y+j) \quad (5)$$

$$M_{11}=\sum_{-K\le i,j\le K} w_{i,j} I_y^2(x+i, y+j) \quad (6)$$

Usually, $w_{i,j}$ in the above formula was a weight in the circular window with a radius of K or in a Gaussian distribution with the center of (x, y).

Harris corners are situated in the two largest eigenvalues of the Hessian image's auto-correlation matrix. If located to the original image, these corners are actually located these centre points around which there are at least two different directions' edges or textures.

First derivative can response to a uniform change gradient, but second derivative cannot response to it, so Harris corners can overcome a uniform change gradient. At the same time, the corners are calculated by the Hessian image's auto-correlation matrix, so they are rotational invariant and have better detecting robustness (Shi and Tomasi, 1994).

LUCAS-KANADE ALGORITHM

Optical flow can be defined as the speed of image pixels. If all pixels' optical flow vectors in an image are calculated, these results are called dense optical flow, which has staggering computation; if only these small number of pixels with certain characteristics are selected to calculate, sparse optical flow is obtained.

Lucas-Kanade algorithm for computing dense optical flow was originally proposed by Lucas and Kanade (1981). The LK algorithm has three following assumed conditions (Bradski and Kaehler, 2011):

- **Constant brightness:** Moving objects' brightness or color remains unchanged between two image scene's adjacent frames
- **Continuous time:** The moving objects' pixels corresponded between two adjacent frames have sufficiently small movement
- **Spatial coherence:** Adjacent pixels have a similar movement

In the assumption 1, if the grayscale value of a pixel (x, y) is I (x, y, t) at the moment t and the pixel moves to a new location (x+Δx, y+Δy) at the moment t+Δt, then its value is referred to I(x+Δx, y+Δy, t+Δt). The assumption 1 shows that each pixel's grayscale value on the moving objects is invariable, there are:

$$\frac{\partial I(x,y,t)}{\partial t}=0 \quad (7)$$

In the assumption 2, the movement changes can be seen as the derivative of the grayscale to time.

According to the assumption 1 and considering that the pixel coordinates of x and y are functions of time t, the brightness I (x, y, t) is replaced with I (x(t), y(t), t) and then the following formulas are derived by the chain rule of partial differential equations:

$$\begin{aligned} &\frac{\partial I}{\partial x}\cdot\Delta x+\frac{\partial I}{\partial y}\cdot\Delta y+\frac{\partial I}{\partial t}\cdot\Delta t=0 \\ &\Rightarrow \frac{\partial I}{\partial x}\cdot u+\frac{\partial I}{\partial y}\cdot v+\frac{\partial I}{\partial t}=0 \end{aligned} \quad (8)$$

Wherein assuming $I_x=\frac{\partial I}{\partial x}, I_y=\frac{\partial I}{\partial y}, I_t=\frac{\partial I}{\partial t}$ that is, I_x and I_y are the partial derivative of an image I to x and y, I_t is the partial derivative of an image I to t. (u, v) is the calculated optical flow vector and then the above equation can be written as:

$$I_x\cdot u+I_y\cdot v+I_t=0 \quad (9)$$

In the above equation, two unknown parameters u and v need to be solved using the assumptions 3 in the LK algorithm. If adjacent pixels move similarly, the equations in the neighboring pixels can be created to solve the center pixel's motion information. For example, let the current pixel as a center, the equations in the 5×5 neighborhood can be established as follows:

$$\underbrace{\begin{bmatrix} I_x(p_1) & I_y(p_1) \\ I_x(p_2) & I_y(p_2) \\ \vdots & \vdots \\ I_x(p_{25}) & I_y(p_{25}) \end{bmatrix}}_{\substack{A \\ 25\times 2}} \underbrace{\begin{bmatrix} u \\ v \end{bmatrix}}_{\substack{d \\ 2\times 1}} = \underbrace{\begin{bmatrix} I_t(p_1) \\ I_t(p_2) \\ \vdots \\ I_t(p_{25}) \end{bmatrix}}_{\substack{b \\ 25\times 1}} \quad (10)$$

The above equation is a system equation with excessive constraint condition. Here, the least squares of the equation can be established to solve the minimization of $\|ad-b\|^2$:

$$\underbrace{(A^TA)}_{2\times 2}\underbrace{d}_{2\times 1}=\underbrace{A^Tb}_{2\times 1} \quad (11)$$

By Eq. 9 and 11, the two unknown parameters u and v can be obtained. Eq. 11 has a more detailed representation as follows:

$$\underbrace{\begin{bmatrix} \Sigma I_x I_x & \Sigma I_x I_y \\ \Sigma I_x I_y & \Sigma I_y I_y \end{bmatrix}}_{A^T A} \begin{bmatrix} u \\ v \end{bmatrix} = -\underbrace{\begin{bmatrix} \Sigma I_x I_t \\ \Sigma I_y I_t \end{bmatrix}}_{A^T b} \quad (12)$$

When $(A^T A)$ is reversible, the equation's solution is as follows:

$$\begin{bmatrix} u \\ v \end{bmatrix} = (A^T A)^{-1} A^T b \quad (13)$$

When $(A^T A)$ has full rank (rank = 2), that is, $(A^T A)$ has two larger feature vectors, $(A^T A)$ is reversible. There are two larger feature vectors in the position where both directions' edges in the image are intersecting. A corner is such a position and the characteristic of $(A^T A)$ is better in the corner.

THREE-FRAME DIFFERENCE

Three-frame difference method is an improved algorithm based on two-frame difference. It puts three adjacent frames as a group, subtracts both two adjacent frames and lets two differential results do the logical AND operation. The algorithm can better obtain the moving objects' region if the moving objects' speed and frame rate are suitable.

The detailed algorithm is described as follows:

- Let n-frame sequence images represent as $\{f_0(x, y), .., f_{(k)}(x, y), ... f_{(n-1)}(x, y)\}$. Wherein $f_{(k)}(x, y)$ represents the k-th frame of the video sequence images. Select three adjacent frames $f_{(k-1)}(x, y)$, $f_{(k)}(x, y)$, $f_{(k+1)}(x, y)$ in the video sequence images to calculate respectively the difference between the two adjacent frames:

$$d_{(k-1,k)}(x,y) = |f_{(k}(x,y) - f_{(k-1)}(x,y)| \quad (14)$$

$$d_{(k,k+1)}(x,y) = |f_{(k+1)}(x,y) - f_{(k)}(x,y)| \quad (15)$$

- By setting an appropriate threshold T to the two difference images obtained in the previous step, the binary images $b_{(k-1,\ k)}(x, y)$ and $b_{(k,\ k+1)}(x, y)$ may be obtained:

$$b_{(k-1,k)}(x,y) = \begin{cases} 1 & d_{(k-1,k)}(x,y) \ge T \\ 0 & d_{(k-1,k)}(x,y) < T \end{cases} \quad (16)$$

$$b_{(k,k+1)}(x,y) = \begin{cases} 1 & d_{(k,k+1)}(x,y) \ge T \\ 0 & d_{(k,k+1)}(x,y) < T \end{cases} \quad (17)$$

- Let $b_{(k,\ -1,\ k)}(x, y)$ and $b_{(k,\ k+1)}(x, y)$ do the logical AND operation for each pixel (x, y), the binary image $b_{(k-1,\ k)}(x, y)$ may be obtained:

$$B_{(k-1,k)}(x,y) = b_{(k-1,k)}(x,y) \otimes b_{(k,k+1)}(x,y) \quad (18)$$

- In order to remove the hole inside the target area and the isolation noise in the image, the binary image $b_{(k-1,\ k)}(x, y)$ is filtered using morphology operations:

$$B_{(k-1,k)}(x,y) = \text{Close}(\text{Open}(B_{(k-1,k)}(x,y))) \quad (19)$$

In the above formula, Open(·) and Close(·) denote respectively the opening operation and closing operation in the morphological processing.

MOVING OBJECTS DETECTION BASED ON THREE-FRAME DIFFERENCE AND SPARSE OPTICAL FLOW

Taking into account the characteristics of the optical flow and three-frame difference method, a moving objects detection method based on three-frame difference and LK optical flow is proposed. Its description is as follow:

Input: video sequences images (I_1, I_2, I_3, I_4, ..., I_n,)
Output: the moving objects' area in the video sequence

Algorithm steps:

(1) Initialization steps:

The 1st and 2nd frame (I_1 and I_2) are subtracted and then Harris corners are detected from the results. The optical flow vectors linked to Harris corners between I_1 and I_2 are calculated using LK optical flow method; then the vectors whose length is less than a threshold are removed. Initialize k = 3.

(2) Iterative steps:

(2a) When a new frame I_k is received, it is subtracted from I_{k-1}. Then, binarizing the difference image between I_{k-1} and I_k, the result does logical AND operation with the one between I_{k-2} and I_{k-1}. Thus, the moving object area is gotten.

(2b) Combine the optical flow vector's endpoint of I_{k-1} with the area from Step 2.1, the final moving objects' area is obtained;

(2c) Harris corners are detected in the area from the Step 2b. These corners are used to calculate optical flow vectors associated I_{k-1} with I_k by the LK algorithm. And then with a threshold, the vectors whose length is less than a threshold are removed.

(2d) Let k = k +1 and return to Step 2a.

By analyzing Step 1, each frame's moving object are can be found except the 1st frame when its next frame is gotten. That is to say, the detection of moving object has a delay with one frame. For example, when the frame I_3 enters, only the moving region of the frame I_2 can be obtained and the one of the frame I_3 can not be. In addition, the 2nd frame I_2 has a different detection method compared with its succeeding frames because only I_1 and I_2 are gotten, so that three-frame difference cannot be carried out. Therefore, Harris corners on which the optical flow of I_2 and I_1 depends are obtained directly from the difference images. But the Harris corners in other frames are obtained from the final moving object area in the previous step.

EXPERIMENTS AND ANALYSIS

This algorithm is run in the PC with the Intel i3-350M, 2G memory and Visual C++6.0 installed of OpenCV 1.0. Its video test sets are standard ones provided by the references (CVPR, 2001), (PETS2000, 2000) and (PETS2001, 2001). These videos include room, outdoor, strong shadows and weak shadow fragments. The results of the experiment diagram are shown in Fig. 1.

Figure 1a is test frames of the standard video, which respectively come from the 100th frame in highway I, the 50th one in highway II, the 150th one in PETS2000 and the 600th one in the Camera2 of PETS2001. Figure 1b is obtained by three-frame difference, which can be seen to contain a lot of noise and found that the detection result is not good when moving objects are too slow. Figure 1c is the optical flow vectors calculated by the proposed method. Figure 1d is the method's result and it is better than any one of the two methods.

The above experimental results show that the proposed method can detect relatively complete moving object areas. The experiments find that rapid objects have better result because the short optical flow vector can be moved by a certain large threshold.

The experiments show that the proposed method can achieve an average processing speed of 21.6 frames per second in the above experimental platform for the video resolution of 320×240 pixels.

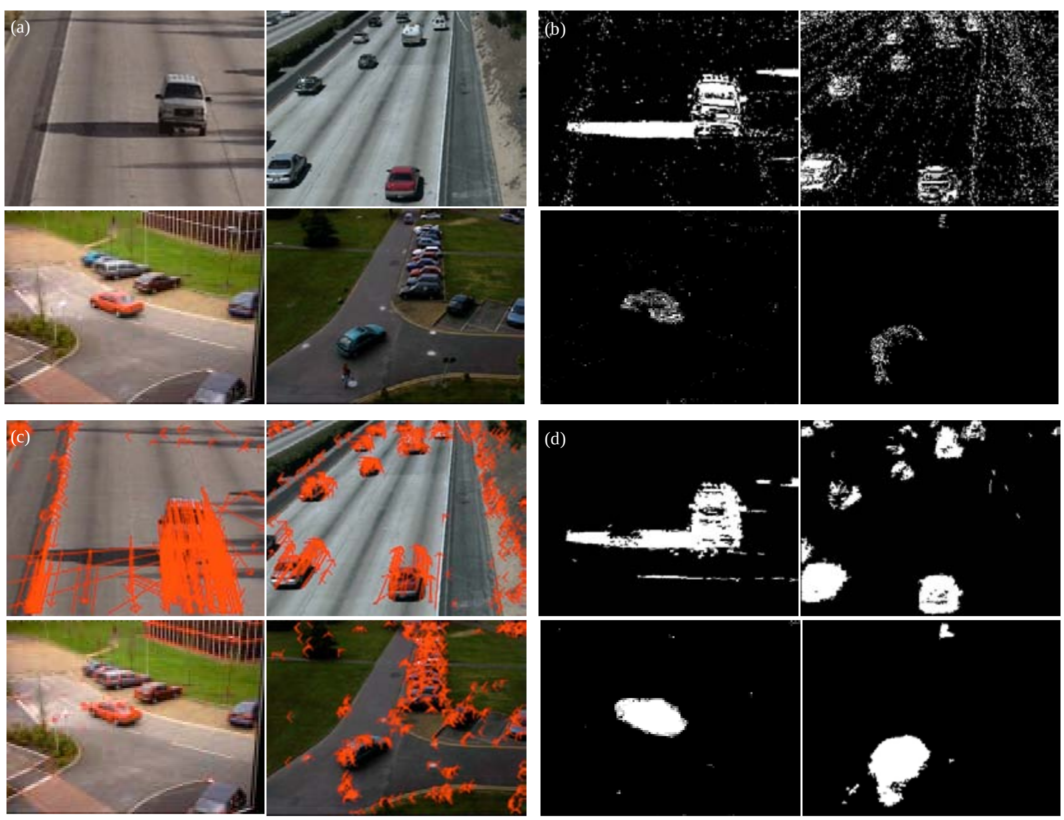

Fig. 1(a-d): The comparison of the moving object detection algorithm based on a combination of three-frame difference and optical flow. (a) Detected image, (b) Results of three-frame difference, (c) Calculated optical flow vectors and (d) Results of the proposed algorithm

CONCLUSION

A moving objects detection method based on three-frame difference and sparse optical flow is proposed in this study. The algorithm simplifies the calculation of optical flow, three-frame difference is employed as a supplement and then the proposed method achieves good results, which has certain reference to moving objects detection.

ACKNOWLEDGMENTS

This study is supported by the Natural Science Foundation of China (No.61163024, 61262067), the Application and Foundation Project of Yunnan Province (No. 2011FB019), the Science Research Foundation of Education Department of Yunnan Province (No. 2012C105), the College Students' Innovation and Entrepreneurship Training Programs of Yunnan Province (No. 2012022), the Undergraduate Research Project of Yunnan University (No. 2012071) and the Young Teachers Training Program of Yunnan University.

REFERENCES

Bradski, G. and A. Kaehler, 2011. Learning OpenCV-Computer Vision with the OpenCV Library. O?Reilly Media, USA.

CVPR, 2001. Computer vision and robotics research laboratory testbed data. http://cvrr.ucsd.edu/aton/shadow/.

Harris, C. and M. Stephens, 1988. A combined corner and edge detector. Proceedings of the 4th Alvey Vision Conference, August 31-September 2, 1988, New York, USA., pp: 147-151.

Huang, C., G. Yuan and D. Xu, 2009. Multi-target detection and tracking by Gaussian mixture model and blob tracking analysis. J. Inform. Comput. Sci., 6: 2403-2410.

Lucas, B.D. and T. Kanade, 1981. An iterative image registration technique with an application to stereo vision. Proceedings of the 7th International Joint Conference on Artificial Intelligence, Volume 2, August 24-28, 1981, Vancouver, BC, Canada, pp: 674-679.

PETS2000, 2000. The first IEEE international workshop on performance evaluation of tracking and surveillance. http://ftp.pets.rdg.ac.uk/pub/PETS2000/.

PETS2001, 2001. The second IEEE international workshop on performance evaluation of tracking and surveillance. http://www.cvg.cs.rdg.ac.uk/PETS2001/.

Shi, J. and C. Tomasi, 1994. Good features to track. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, June 21-23, 1994, Seattle, WA., USA., pp: 593-600.