A Novel Method for Moving Objects Detection Inspired by Frog's Visual Characteristics

Zhiyong Li*, Zhen Jiang, Ahmed Sallam and Junmin Liu

College of Information Science and Engineering,
Hunan University, Changsha, China
zhiyong.li@hnu.edu.cn, 245736419@qq.com, sallam_ah@hnu.edu.cn, 326860194@qq.com

Abstract

The detection of moving objects is the hot topic of computer vision researches now. On this subject, we propose a novel method to detect moving object in surveillance videos based on the visual features of the frog eyes. This method is divided into three steps; first, to extract the moving regions using Frame Difference and Spatial-temporal entropy methods; second, to establish the edge map by the Canny edge operator; finally, to produce the moving object's edge by double threshold fusion based on motion regions and edge map. Our experimental results show that the proposed method can detect the edge of moving object effectively.

Keywords: *Frog's Visual Characteristics, Spatial-temporal Entropy, Moving object detection, Canny operator*

1. Introduction

Scientists are studding Computer Vision for 40 years now; however, many difficulties are encountered when applying traditional theories of computer vision into the development of practical system. Thus, many concerned researchers begin to doubt about the traditional theories and tend to find new theories affected by the vision behavior of human and animals [1].

The detection of moving objects is the trend in the field of computer vision in the recent years. It is widely used in military and civilian areas, such as: weapons, security equipment, and traffic control devices; therefore, it has an important practical value and broad development prospects. However, there are huge challenges when it comes to design a computer vision system because of the many interference factors in reality especially when people put forward higher requirements on its capability against reactivity (real-time) and robustness [2], respectively; many researchers suggest a series of algorithms. Lipton, *et al.*, [3], used the adjacent two frame differential method to detect moving objects in actual video sequence, this method succeeded to track moving targets to some limits. Next, VSAM [4] tried to improve this methodology; so, he puts forward an algorithm combining adaptive background difference and three frame difference.

Ma and Zhang [5] proposed a detection method for moving objects based on the spatial-temporal entropy. In their method, they built a spatial-temporal entropy image (STEI), which considers those pixels with larger entropy in the STEI as part of moving objects. It is more robust to noises than traditional difference based methods owing to that the STEI method is a statistical measurement of variation. It is easy to implement this method, put it suffers from some weakness to detect the edge of a moving target due to the diversity of spatial structure;

* Corresponding Author

moreover, it requires high computational complexity. Later, Jing, *et al.*, [6] proposed a detection method based on the difference-based spatial-temporal entropy image (DSTEI). This method obtains the motion image after applying spatial-temporal entropy image to frame difference. It wipes out the disturbance of static background edges to some degrees; but, it can't effectively meet the accuracy and the real-time requirements. In sum, we can conclude that in the filed of detecting moving objects using traditional theories, some algorithms are fast but not accurate and others are more efficient but with higher complexity.

To this, various biological visual system which after a long-term evolution provide a good reference to solve traditional algorithms drawbacks [7-8]. Frogs for example have little mental capacity, however they depend heavily on its vision system to continue in life, This can be a good observation to solve the present encountered traditional algorithms drawbacks by simulating the frog's visual behavior [9].

Currently, researches and applications based on the behavior of frog's vision are relatively narrow. Zhi-ling Wang have proposed A Fuzzy Region Understanding Tactic for Object Tracking Based on Frog's Vision Characteristic [10]. In his method, it was noticed that frog is myopia and concern on the contour; thus, he used mean-shift algorithm to track the target profile after by smoothing the image sequence in a fuzzy manner. This method has better robustness for local mutations situation. Next, Kimihiro NISHIO [11] proposed and designed a two-dimensional network for moving objects detection constructed from simple analog circuits based on the frog visual system. Based on simulation and practical experiments the results shows that the network succeeded to detect motion direction and velocity of a moving object. Afterwards, Zhao Liang [9] introduced a detailed system analysis and use the computer 3D to simulate the frog vision behavior, this kind of visual behavior model could be taken as a good reference to study the formation of the vision characteristics and develop the computer vision.

In this paper, we propose a motion detection method by combing the edge feature and motion information based on the features of frog's visual system. The results showed that, this method is highly adaptive to static environment. The reset of this paper is organized as follows: In Section 2 we introduce a new moving detection strategy based on the frog vision features. In Section 3 we explain the algorithm design. In Section 4 we analyze the experimental results, finally we conclude this paper in Section 5.

2. Algorithm strategy based on the features of frog vision

Wiener and his colleagues did a large number of physiological experiments on number of organisms (including frogs) as a basis for the creation of cybernetics in 1948 [12]. 1960s, the representatives of researching physiology and behavior of frog's vision include Ingle D, Ewert JP, Speery RW [13]. Later in 1980s, a new round of the study on the frog's vision began based on the original study with the development of neural networks.

Frog is known for the unique visual characteristics, it relies mainly on the vision to prey and find their natural enemies. Frog's vision is sensitive to moving objects, and can't focus on the static details of the surrounding, so it may even starve to death in a pile of static food [4]. Researchers found that frogs rely mainly on the movement and the size of external objects to prey or to escape and so on, and we can easily push the frog to respond by swinging small moving objects. Moreover, the frog's eyes are unable to act as human's eyes such that it is not easy to transfer the attention when the frog's eyes are focus on an object; however, the. compensatory motion of its eyes will be achieved when the frog's body position changes, or its entire visual world flips. Besides, the frog can't see distant objects clearly, because the surface will be formed disturbed in frog's eye when the light comes from long distance [13].

Neuroscience researchers made a detail research and description about frogs amphibian's physical structure, such as the brain structure, nerve distribution, visual characteristics and behavior. With a lot of speculation and various kinds of visual stimulation experiment in analysis of retinal ganglion cells, Lettvin, Maturan, Pitts and McCulloch divided its ganglion cells into five types [13]:

- Edge or contrast detector: Response to the local comparison between the bright and dark.
- Motion convex edge detector: Response to the dark convex side which have entered the receptive field.
- Moving side or change of contrast detector: Response to the movement of border between light and dark area.
- Darken detector: It don't need boundary stimulation, this cell responses to darking of the whole receptive field.
- Dark detector: Tension discharge and its discharge frequency are inversely proportional to the diffuse light felling in its larger receptive field.

In brief, the frog can pay attention to its interesting object quickly and effectively, this ability has a a direct relationship with its visual features and constitution. In this paper, according to the frog vision system, we propose a method to extract the fringe of the moving objects by fusing the motion and the edge information. This method based on utilizing a set of techniques includes moving region detection, Canny edge extraction and double threshold fusion the detail algorithm design is as follow:

3. Algorithm design

3.1. Moving region detection

The Frame Difference is a common method of motion detection. The basic idea of this method is to extract an irregular image by subtracting adjacent frames in the video sequence, then to smooth this image to remove the noise. Finally, to evaluate the value of each point in the extracted image, if the value is greater than a predefined set of threshold, it is considered as a part of motive area, if not, it is considered as a part of the static background.

In information theory, entropy represents the amount of information the data contains, the greater the entropy value, the more information contained. Shannon [14] proposed an information measure based on probability and statistics model in his literature "A Mathematical Theory of Communication". As he indicates that, since entropy is a measure of information indicating the diversity of events, so it can be used to characterize the diversity of pixels state, thus reflect the strength of pixels movement.

In this paper, we calculate the spatial-temporal entropy of gray image as follows: Figure 3 shows the spatial-temporal sliding window, there are three steps for calculating the spatial-temporal entropy in gray image [7]:

Step 1: A spatial-temporal sliding window is firstly formed for each pixel to obtain the corresponding spatial-temporal histogram. The spatial-temporal sliding window S($w \times w \times L$) for (x,y) is defined as:

$$S = \{(i,\, j)_k \mid |i - x| < \lfloor w/2 \rfloor, |j - y| < \lfloor w/2 \rfloor, 0 \le t - k < L\}$$

Step 2: The w×w×L pixels in the sliding window are then classified into Q bins according to their intensity, and are accumulated to obtain the spatial-temporal histogram. Let

$H_{x,y,q}$ denote the number of pixels belonging to bin q in the sliding window of pixel (x,y).

Step 3: The corresponding probability density function (pdf) for component q can be calculated by:

$$E_{x,y} = -\sum_{q=1}^{Q} P_{x,y,q} \log(P_{x,y,q})$$

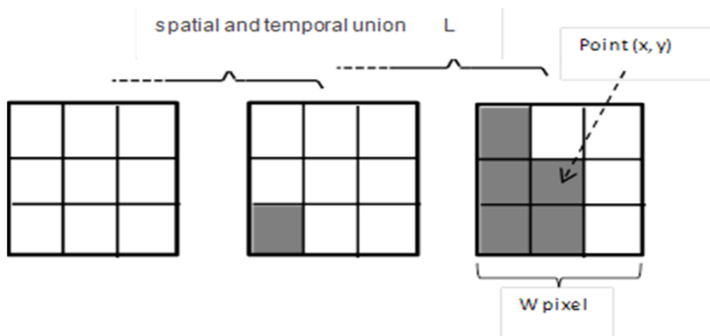


Figure 3. The spatial-temporal sliding window

3.2. Canny edge detection operator

Edge is the basic characteristic of images; it contains valuable target boundary information, which can be used for image analysis, target recognition, *etc.* Methods of edge detection usually based on different techniques such as, differential equations, wavelet, fractal theory and mathematical morphology.

Canny [9] operator is the most commonly used edge detection operators, and also recognized as the one with excellent performance, it is often cited by other operators as the standard operator for comparative analysis of the advantage. Although the Canny operator is a first-order differential operator, it was extended by adding a non-maximum suppression and dual-threshold on the basis of original first-order differential operator. Using non-maximum suppression don't only suppress unnecessary edge effectively, but also can improve the positioning accuracy of the edge; taking advantage of dual-threshold can effectively reduce the rate of undetected the edge. The Canny operator procedure include the following steps:

Step 1: Image denoising.

Step 2: calculation on Gradient amplitude and direction angle. Computing the gradient M_x and M_y in the x and y direction of the image respectively after denoising.

Gradient amplitude:

$$|\nabla f| = \sqrt{M_x{}^2 + M_y{}^2}$$

Direction angle:

$$\theta = \arctan(\frac{M_y}{M_x})$$

Step 3: Non-maximum suppression. For each point, comparing the two gradient amplitude along the gradient direction. If one is the largest, retain it and remove the other. Through all gradient amplitude, the candidate image M is obtained.

Step 4: Double threshold. A high threshold T_h and a low threshold T_l are first set for the candidate image M. If the gradient amplitude of any pixel in M is greater than T_h, it will be considered as the edge, and non-edge if less than T_l. For the pixel whose gradient amplitude is between T_h and T_l, it will be marked as the edge if the gradient amplitude of its 4- adjacent or 8-adjacent pixel is greater than T_h ,otherwise, it will be not [10].

3.3. Double threshold fusion

In this paper, According to the moving region detection techniques mention before, we define two thresholds Mov_h and Mov_l; where, Mov_l based on Frame Difference and Mov_h based on Spatial-temporal entropy. For the edge map, if the edge points reach threshold Mov_h, it is considered as a moving edge, if the edge points which belong to 4-adjacent or 8-adjacent points of moving edge arrive threshold Mov_l, it is also considered as a moving edge. Finally, we recognize the whole edge of moving object with the following fusion method:

1. Through the Frame Differential, we group motion points coordinates in set A, and then we create another set B by calculating Spatial-temporal entropy of pixels in set A, such that, $B \subseteq A$;

2. Through Canny edge operator extraction, we group the edge points coordinate in set C;

3. $D = B \cap C$, $E = A \cap C$;

4. We define another three sets, set D contains seeds point, and set E represent the growth areas and conditions, using 4- neighborhood or 8- neighborhood to region grow, finally, we obtain set F which includes the coordinates of moving edge points.

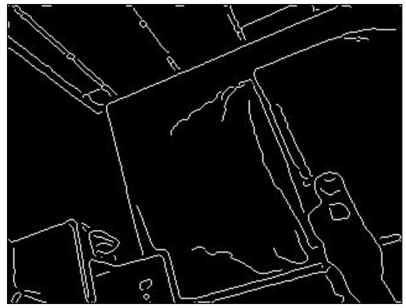
4. Experimental results

The proposed method was tested using various real video sequences. Figure 4 show the experimental results for the indoor video such that: (a) shows the input image of 14^{th} frame of original sequence, (b) shows the result of edge detection, (c) shows the result of Frame-Difference, (d) shows the spatial-temporal entropy image calculated based on Frame-Difference image, (e) shows the fusion of figures (b) and (c), (f) shows the fusion of figures (b) and (d), (g) shows the final results based on double thresholds. Moreover, we have conducted the same experiment for the outdoor images sequence as shown in Figure 5 for the auto surveillance video.

From the experimental results, we can see that through the double thresholds, some background information is suppressed, but the counter of the target is effectively reserved (see figure (g)). Figure (e) shows the result of using threshold Mov_l by fuseing (b) and (c), and the image still has some background interference; while, figure (f) shows the result using threshold Mov_h by fusing (b) and (d) such that the information of target's edge is incomplete. By using double thresholds, we have enhanced the application range and robustness of the system.

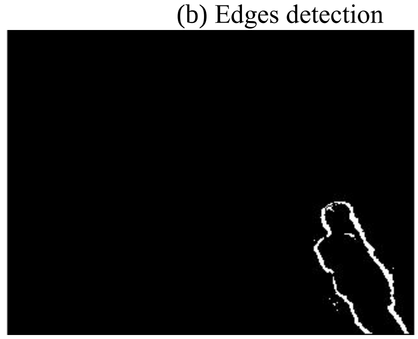
Figure 6 shows the motion detection results of three methods: (a) shows the original image from the video sequences, (b) shows the result of STEI method, (c) shows the result of DSTEI method; finally, the result of our method is shown in (d). In this experiment, the parameters are set as follows: the width of the sliding window w is set to 3, the frame length of the sliding window L is set to 3 too, and parameter Q is set to 33.

（a）Original sequence



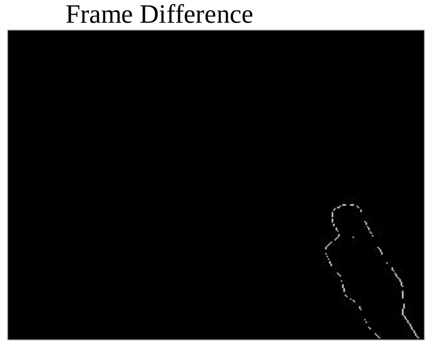
(b) Edges detection

（c）Frame Difference

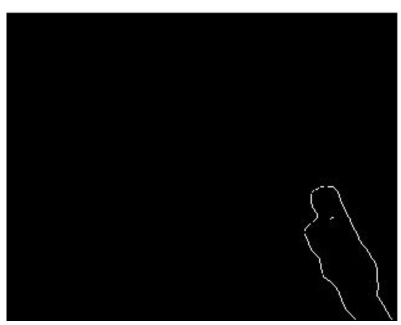


(d) Spatial-temporal entropy method base on Frame Difference

(e) Threshold based on (b) and(c)



(f) threshold based on (b) and (d)

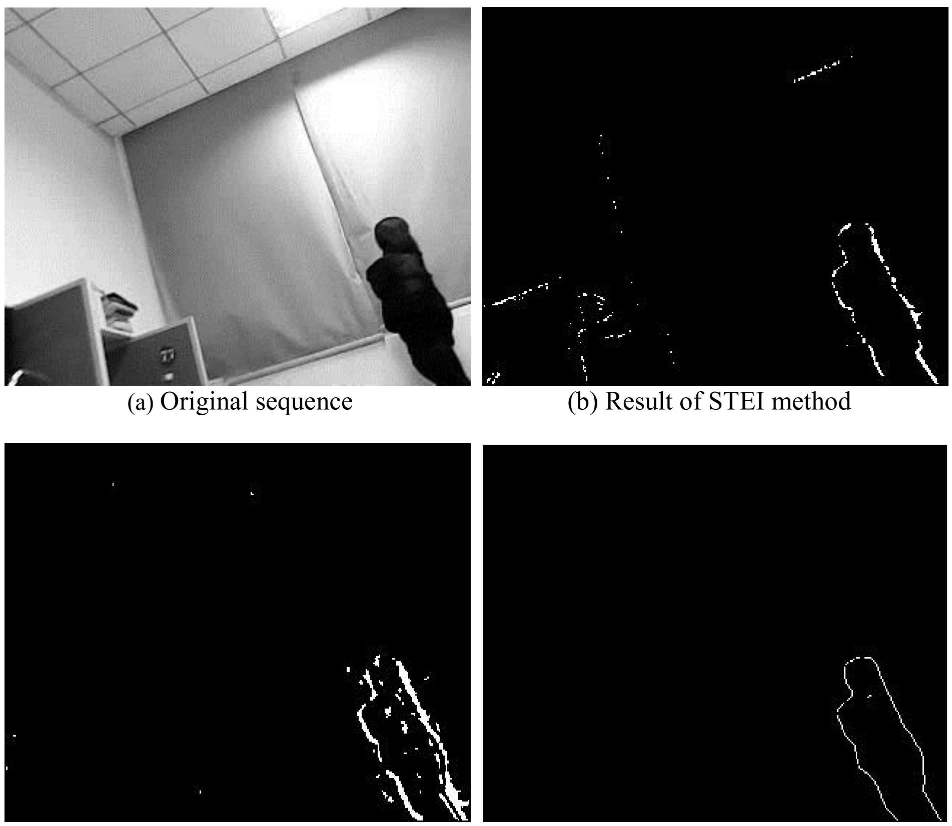


(g) The final results base on double thresholds

Figure 4. Experimental results in indoor video

(a) Original sequence (b) Moving detection result

Figure 5. Experimental results in outdoor video



(a) Original sequence (b) Result of STEI method

(c) Result of DSTEI method (d) Result of our method

Figure 6. Experimental results of three algorithms

5. Conclusions

Inspired by the visual characteristics of frog eyes, in this paper we proposed a moving object detection algorithm using Frame Difference for motion detection, Canny operator for edge detection, and double threshold fusion to simulate the frog eye's detector. The main features of this algorithm is:(1) We Learn from the frog visual system, and simulate part of the visual neural mechanisms of frog eyes; (2) We only calculate the spatial-temporal entropy based on the Frame Difference image, which reduce the computational complexity, remove lots of static information and extract the movement area more accurately; (3) By using double threshold, we extract the edge of target more effective and enhance the systematic robustness. In the Future work we intend to perform intensive deeper research on frog eyes and provide more sophisticated method that could detect moving objects in more complex scenes.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Grant No. 61173107), the Research Foundation of industry-education-research cooperation among Guangdong Province, Ministry of Education and Ministry of science & Technology, China (Grant No. 2011A091000027) and the Science and Technology Plan of Changsha, Hunan Province, China (Grant No. K1109099-11).

References

[1] T. Wang, "Computer Vision and Reconstruction", Pattern Recognition and Artificial Intelligence, vol. 1998, no. 11, **(1998)**, pp. 300-304.
[2] J. C. Enrique, *et a1.*, "A new video segmentation method of moving objects based on blob-level knowledge", Pattern Recognition letters, vol. 2008, no. 29, **(2008)**, pp. 272-285.
[3] A. Lipton, H. Fujiyoshi and R. Patil, "Moving Target Classification and Tracking from Real-Time Video", Proceedings of IEEE Workshop on Applications of Computer Vision, Princeton NJ, **(1998)** August 14.
[4] R. Collins, *et a1.*, "A system for video surveillance and monitoring: VSAM final report", Carnegie Mellon University, Technical Report: CMU-RI-TR-00-12, **(2000)**.
[5] Y. -F. Ma and H. -J. Zhang, "Detecting motion object by spatial-temporal entropy", in Proceedings of the 2001 IEEE International Conference on Multimedia and Expo, **(2001)**, pp. 381-384.
[6] J. Guo, E. S. Chng and D. Rajan, "Foreground motion detection by difference Based spatial-temporal entropy image", In:Proc:IEEE Region 10 Conf TenCon2004, Chiang Mai, Thailand, **(2004)**, pp. 379-382.
[7] J. Fan, Y. Wu and S. Dai, "Discriminative spatial attention for robust tracking", in: ECCV, **(2010)**, pp. 480-493.
[8] L. A. Torres-Mendez and E. J. Olaya, "A Biologically-inspired Robotic Vision System for Tracking Fast Moving Objects", Technologies for Practical Robot Applications, **(2011)**, pp. 162-167.
[9] L. Zhao, "Basic research of frog visual behavior and computer simulation", master dissertation, Wuhan University of Technology, **(2005)**.
[10] Z. -L. Wang, Z. -H. Chen, X. -X. Xu, L. Wu, "A Fuzzy Region Understanding Tactic for Object Tracking Based on Frog's Vision Characteristic", ACTA AUTOMATICA SINICA, vol. 35, no. 8, **(2009)**, pp. 1048-1054.
[11] K. Nishio and H. Younezu, "A Two-Dimensional Network of Analog Circuits for Motion Detection Based on the Frog Visual System", IEICE TRANS, vol. E89(A), **(2006)**, pp. 428-438.
[12] N. Wiener, "Cybernetics", Beijing: Beijing Science Press, **(1963)**.
[13] J. Y. Lettvin, H. R. Maturana, W. S. McCulloch and W. H. Pitts, "What the Frog's Eye Tells the Frog's Brain", Proc. Inst. Radio Engr., vol. 47, **(1959)**, pp. 1940-1951.
[14] Z. Li, J. Zhen and J. Liu, "A New Method of Motion Detection with Biological Intelligence", International Journal of Signal Processing, Image Processing and Pattern Recognition, vol. 5, no. 2, **(2012)** June, pp. 141-151.
[15] S. A. Toussi and H. S. Yazdi, "Feature Selection in Spectral Clustering", IJSIP, vol. 4, no. 3, **(2011)** September, pp. 179-194.
[16] J. K. Basu, D. Bhattacharyya and T. -h. Kim, "Use of Artificial Neural Network in Pattern Recognition", International Journal of Software Engineering and Its Applications(IJSEIA), vol. 4, no. 1, **(2010)** January, pp. 67-74.