A Fast and Robust Algorithm for Fighting Behavior Detection Based on Motion Vectors

Jianbin Xie, Tong Liu, Wei Yan, Peiqin Li and Zhaowen Zhuang
Department of Electronic Science and Engineering, National University of Defense Technology
Changsha, CO 410073 - China
[e-mail: jbxie@126.com, liutong1129@126.com]
*Corresponding author: Tong Liu

Received June 16, 2011; revised September 20, 2011; accepted October 17, 2011; published November 29, 2011

Abstract

In this paper, we propose a fast and robust algorithm for fighting behavior detection based on Motion Vectors (MV), in order to solve the problem of low speed and weak robustness in traditional fighting behavior detection. Firstly, we analyze the characteristics of fighting scenes and activities, and then use motion estimation algorithm based on block-matching to calculate MV of motion regions. Secondly, we extract features from magnitudes and directions of MV, and normalize these features by using Joint Gaussian Membership Function, and then fuse these features by using weighted arithmetic average method. Finally, we present the conception of Average Maximum Violence Index (AMVI) to judge the fighting behavior in surveillance scenes. Experiments show that the new algorithm achieves high speed and strong robustness for fighting behavior detection in surveillance scenes.

Keywords: Surveillance scene, fighting behavior, motion vectors, average maximum violence index, image analysis

DOI: 10.3837/tiis.0000.00.000

1. Introduction

Surveillance cameras are inexpensive and widespread these days but the manpower required to monitor and analyze them is expensive and inefficient. Therefore, the research of behavior analysis becomes a hot topic on the field of intelligent surveillance. Much work has been done in this field. The developments and general strategies of behavior analysis are reviewed in [1], including the detail presentment of the following stages: modeling of environments, detection of motion, classification of moving objects, tracking, behavior understanding and description, and fusion of information from multiple cameras. The framework in [2] for behavior analysis consists of four major modules: behavior modeling, feature extraction from video sequences, basic behavior unit discovery and complex behavior recognition. There are many kinds of interesting human behaviors to study. In [3], an intelligent approach is proposed to analyze player's behaviors in racket sports video, which can help referees to distinguish some transgressors. In [4], a system is presented to detect abnormal activities by extracting the features of human postures. In [5], an intelligent video surveillance system is proposed to detect fall incidents of human for indoor safety. In [6], an approach is proposed to analyze driver's behavior for traffic safety. In [7], a method is proposed for the real-time detection of vandalism in video sequences.

Among the studies of behavior analysis, the auto-detection of human's fighting behavior is one of the most active research topics, which has a broad application for society security. As to the field of fighting behavior detection, some work has been done, too. In [8], an algorithm is presented to detect human violent behavior, such as fist fighting, kicking, hitting with objects, etc, which relies on motion trajectory information and direction information of persons' limbs. In [9], the proposed algorithm extracts the distance features from the boundary contour to the centric position of humans, and trains and recognizes them by using PCA-SVM algorithm. However, both of them need to extract the intact contour of human, which is difficult to be achieved when the scene is complicated or the target is fragmentary. In [10] and [11], an algorithm based on analyzing the color feature of the motion region is proposed to detect fighting behavior, and a local complexity concept called Maximum Warping Energy (MWE) is introduced to describe the color changes with the variation of time and space. The algorithm proposed in [12] uses three steps to detect whether the behavior of a single person in video sequence is abnormal or not. 1) to use Mixture Gaussian Model to obtain background model; 2) to use color-shape information and Random Hough Transform to extract the zebra crossing, and segments the background; 3) to use object's rectangle features to judge whether the person's behavior is abnormal or not. However, these algorithms are sensitive to color information of surveillance scenes, so they cannot adapt to a complicated environment such as night or weak illumination. In [13], the proposed algorithm uses optical flow to extract the speed and direction features of the moving objects, and introduces a local complexity concept called Violent Action Detection (VAD) to define violent activities, but optical flow has the disadvantages of complicated calculations and large time-consuming, and the accuracy may be very low if we narrow the scale of surveillance scenes. In [14], the proposed algorithm works well in the simple scenes with a few of people, but cannot adapt to the crowding scenes. In [15], limb movements are characterized by using the statistics of angular and linear displacement, and the entropy of the Fourier spectrum is used to measure the randomness of subject's motions. However, the occlusion of limb movements always exists in surveillance scenes.

There are two problems in the former algorithms for human's fighting behavior detection: 1) the speed of these algorithms is too low to meet the requirement of real-time detection on embedded platform; 2) these algorithms cannot adapt to the complicated environment. In other words, the accuracy rate may reduce obviously in the dark or crowding environment. To solve these problems, this paper proposes a fast and robust algorithm for human's fighting behavior detection. The new algorithm includes four steps: 1) to select the motion regions in videos as the research object, and to extract features from magnitudes and directions of motion vectors; 2) to normalize these features by using Joint Gaussian Membership Function; 3) to fuse these features by using weighted arithmetic average method, and then to present the concept of Average Maximum Violence Index (AMVI) to describe the variation of the velocity and direction of motion object; 4) to detect fighting behavior in terms of the confusion degree of motion region in surveillance scenes.

The paper is organized as follows. Section II discusses and calculates the motion vectors in fighting scene. Section III proposes a fast and robust algorithm for fighting behavior detection. Section IV describes the experiments by using our algorithm to detect fighting behavior, and discusses the performance of our algorithm. Finally, we conclude in Section V.

2. Motion Vectors in Fighting Scene

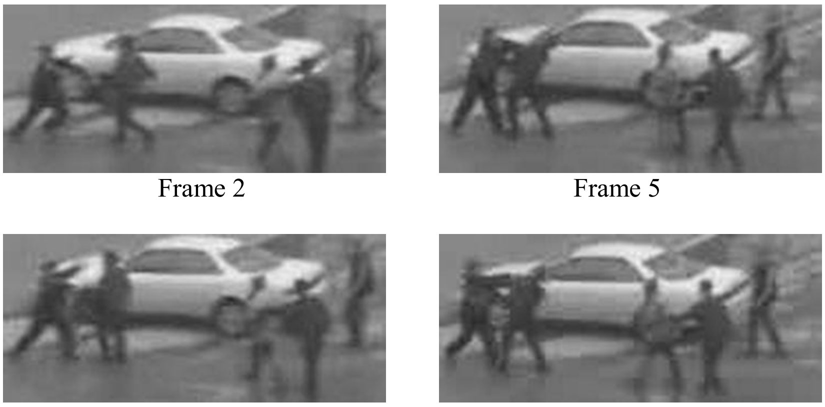
Fighting behavior is a kind of antagonistic behavior among two or more persons, which represents as drastic movements by means of persons' bodies or weapons, such as fighting, robbing, murdering, etc. The scene that contains fighting behavior is called fighting scene.

Compare to normal scene, fighting scene has some peculiar features. When violence occurs, the distance between two persons may be very close, and the contour of persons' body is difficult to extract. Some of persons' body such as arms and legs may move quickly, so the movement energy in scene may increase. Some of persons' body may overlap each other, so the scene may be very disordered.

Such features of the fighting scene can be reflected well by Motion Vectors (MV). Motion vectors carry much motion information such as trajectory, velocity, direction of movement, etc. When fighting behavior occurs, some magnitudes and residual data of motion vectors may increase, and the directions of these vectors may become disordered. As is shown in sequence diagrams of Fig.1, people in the left part of the scene are fighting with each other, while the others in the right part are peacemakers. Obviously, the former is fighting behavior while the latter is normal behavior. We divide the screen into two parts, and extract motion vectors from 6 frames, then draw motion vectors' distribution on (b) and (c) of **Fig. 1**. From the figure, we can see clearly that the motion vectors of fighting behavior are obviously different from the normal ones. The magnitudes and residual data of MV in the left part are larger than that in the right part, and the directions of MV in the left part are more disordered than that in the right part.

Frame 1

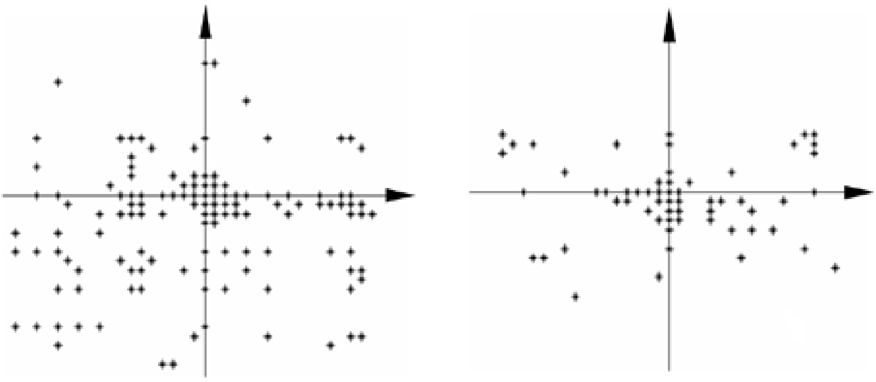
Frame 4



Frame 2 Frame 5

Frame 3 Frame 6

(a) Surveillance scene



(b) Distribution of MV in left part (c) Distribution of MV in right part

Fig. 1. Surveillance scene and distribution of MV

3. A Fast and Robust Algorithm for Fighting Behavior Detection

3.1 Motion region segmentation

The motion targets in surveillance scenes may be persons, animals, moving cars, etc. When fighting behavior occurs, there must be mutual movement or overlap among two or more persons. Therefore, the distance between two persons may be very close, and it is difficult to extract intact contour of persons' body. So we regard the adjacent motion targets as an integral part of the whole, and use Motion Region (MR) to mark it. In this way, the problem of mutual overlap among motion targets can be solved effectively, and it can enhance the robustness of the algorithm to adapt to complicated surveillance scenes. Meanwhile, the speed of the algorithm can be improved by analyzing the motion region of the surveillance scenes.

(1) Motion target detection

Motion target detection algorithm mainly includes background subtraction method, adjacent frame difference method, consecutive frame difference method, optical flow method, etc. Among them, the adjacent frame difference method is efficient, fast, easy to implement, and has been widely used in motion target detection. In this paper, we use adaptive three-frame difference method to detect motion targets. The implementation process is as follows:

***Step*1** We calculate the frame difference image *Erro* and the absolute gray difference images $D_{(k,\,k\text{-}1)}$, $D_{(k,\,k+1)}$ by using consecutive images $I_{k\text{-}1}$, I_k and I_{k+1}.

$$Erro = \left| I_{k+1} - \alpha I_k - (1-\alpha) I_{k-1} \right| \tag{1}$$

$$D_{(k,k-1)} = \left| I_k - I_{k-1} \right| \tag{2}$$

$$D_{(k+1,k)} = \left| I_{k+1} - I_k \right| \tag{3}$$

Where α is a weight value, and the initial value is set to 0.5. k is the serial number of video frame;

***Step*2** We calculate the mean value of image *Erro*. And then calculate the adaptive threshold T_1 by multiplying the mean value by a weight value.

$$m = \frac{1}{W \times H} \sum \sum Erro \tag{4}$$

$$T_1 = \beta \times m \tag{5}$$

Where $W \times H$ is the size of the image; β is a weight value, which is set to 10 according to experiments.

***Step*3** We update the weight value α, and then extract motion region M_k, which is a binary image. In image M_k, the pixels' values of motion targets are set to 1, while the ones of background are set to 0.

$$\alpha = e^{-2/m} \tag{6}$$

$$M_k = \begin{cases} 1 & , D_{(k,k-1)} \geq T_1, D_{(k+1,k)} \geq T_1 \\ 0 & , else \end{cases} \tag{7}$$

(2) Motion region marking

There are many discrete target points in image M_k, which need to be combined to motion regions. This process is as follows:

***Step*1** We use median filter algorithm to eliminate the isolated pixels in image M_k;

***Step*2** We use image morphological operations such as dilation and erosion to remove the hole in image M_k, and combine the adjacent target points together;

***Step*3** We use 8-connected neighborhood method to search target blocks, and calculate the number of target points in a target block. When the number is larger than the threshold T_2, we regard the block as motion region and mark it. Where, T_2 is set to ($W \times H \times 0.01$) according to experiments.

3.2 Motion Vector Calculation

The calculation of motion vector is the basic of fighting behavior detection. There are many fast motion estimation methods based on block matching, such as Three-Step Search (TSS) method, Two-Dimensional Logarithmic Search (2D LOGS) method, Block-Based Gradient Descent Search (BBGDS) method, Four-Step Search (FSS) method and Hexagonal Search (HEXBS) method.

These methods use different search patterns and strategies in order to achieve lower computational complexity and higher accuracy. This paper uses Improved Three-Step Search (ITSS) method proposed in [16], which has considered that the distribution of motion vectors in videos is center biased. This method uses the new design of diamond-shaped search pattern to reduce the possibility of being trapped in the local minimum, in this way it achieves lower computational complexity and higher accuracy. The implementation process is as follows:

***Step*1** In **Fig. 2**, we search for 17 points to obtain the location of Minimum Block Difference (MBD). The 17 points are marked with 1 in **Fig. 2**. If the MBD point is located in the center of the search window, the method ends; if the MBD point is in the large diamond template, then goes to step 2, otherwise goes to step 3.

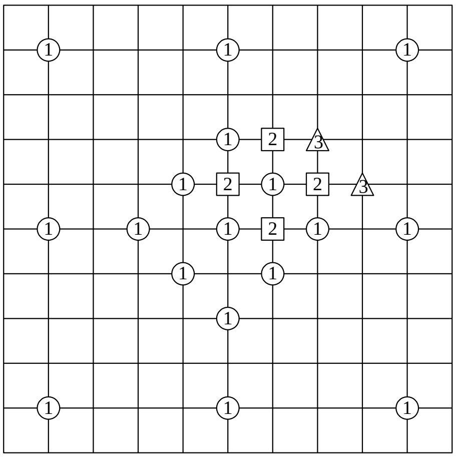


Fig. 2. ITSS search procedure

***Step*2** We use the MBD point in *step* 1 as the center, and reuse the small diamond search patterns until the MBD point is in the center of the search window.

***Step*3** We reduce the step length by half and identify the new MBD point until the step size is equal to 1.

There are three matching criteria in the Block Matching Algorithms (BMA): Normalized Cross-Correlation Function (NCCF), Mean Square Error (MSE) and Mean Absolute Error (MAE). Experiments show that NCCF is highest computational complexity and MAE is least sensitive to the difference of two blocks. Therefore, we select MSE as the matching criterion:

$$MSE = \frac{1}{W \times H} \sum \sum (I_{k+1} - I_k)^2 \tag{8}$$

3.3 Features extraction

When the surveillance scenes contain fighting behavior, the magnitudes and residual data of MV may become great, or the movement directions of MV may become disordered. However, the experiments show that the residual data of MV cannot characterize the differences between fighting behavior and normal behavior well, because some normal behavior may also bring on the increase of the residual data of MV. Therefore, this paper mainly extracts features from magnitudes and directions of MV. The following five kinds of features can characterize the differences between fighting behavior and normal behavior.

(1) Coincidence indicator f_U

Generally speaking, relative motion among persons occurs more frequently in fighting scenes. In this situation, the centric position of a motion region moves slowly, which lead to the small mean value of motion vectors. We define coincidence indicator f_U to characterize this phenomenon.

Simply, we suppose that a motion region can obtain M nonzero motion vectors by using block-matching criteria. The motion vector of the ith block is (Vx_i, Vy_i). We calculate the mean value of the motion vectors in direction x and y respectively:

$$\overline{Vx} = \frac{1}{M}\sum_{i=1}^{M} Vx_i \tag{9}$$

$$\overline{Vy} = \frac{1}{M}\sum_{i=1}^{M} Vy_i \tag{10}$$

Then, we calculate coincidence indicator f_U by using the following formula:

$$f_U = e^{-\lambda\sqrt{\overline{Vx}^2 + \overline{Vy}^2}} \tag{11}$$

Where, λ is a fixed coefficient, which is set to 0.014 according to experiments.

(2) Mean value M_R and variance value σ_R of the motion vectors' magnitudes

When fighting behavior occurs, some parts of the motion region, such as arms, legs and weapons, may move quickly, but the other parts may move slowly. Therefore, the variance of motion vectors' magnitudes becomes large. We use the mean value M_R and variance value σ_R of the motion vectors' magnitudes to characterize this phenomenon. The calculation process is as follows:

Firstly, we calculate motion vectors' magnitude R_i of the $i$thblock by using the following formula:

$$R_i = \sqrt{Vx^2 + Vy^2} \tag{12}$$

Then, we calculate M_R and σ_R:

$$M_R = \frac{1}{M}\sum_{i=1}^{M} R_i \tag{13}$$

$$\sigma_R = \sqrt{\frac{1}{M}\sum_{i=1}^{M}(R_i - M_R)^2} \tag{14}$$

(3) Entropy value E_o and minimum mean value M_o of the motion vectors' directions

When fighting behavior occurs, conflicts among persons may lead to disordered directions of motion vectors. We use the entropy value E_o of normalized direction and the minimum mean value M_o of relative direction to characterize this phenomenon. The implementation process is as follows:

***Step*1** We normalize the direction of motion vectors. The direction space is divided into N regions, which are marked by 0 ~ N-1 respectively. The larger N is, the more prominent direction's diversity feature is; the smaller N is, the more prominent direction's consistency feature is. In this paper, N is set to 16. By calculating the probability of movement vector in every direction, we obtain the histogram $H(\theta)$ of normalized direction, which is shown in **Fig. 3**.

***Step*2** We calculate the entropy value E_o of normalized direction as follows:

$$E_o = \sum_{i=0}^{N-1} p_i \log p_i \tag{15}$$

Where p_i is the possibility, it indicates the possibility that the motion vector locates at the *ith* direction.

Step3 We calculate the minimum mean value M_o of relative direction. The relative direction θ_{ij} between the ith and jth direction in histogram $H(\theta)$ can be calculated as follows:

$$\theta_{ij} = \begin{cases} |\theta_i - \theta_j| & ,|\theta_i - \theta_j| \le 8 \\ 16 - |\theta_i - \theta_j| & ,else \end{cases} \tag{16}$$

The mean value of the ith relative direction is:

$$\overline{\theta_i} = \frac{1}{N}\sum_{j=0}^{N-1}(\theta_{ij} \times p_i) \tag{17}$$

The minimum mean value M_o is:

$$M_o = \min_{0 \le i \le 15}\{\overline{\theta_i}\} \tag{18}$$

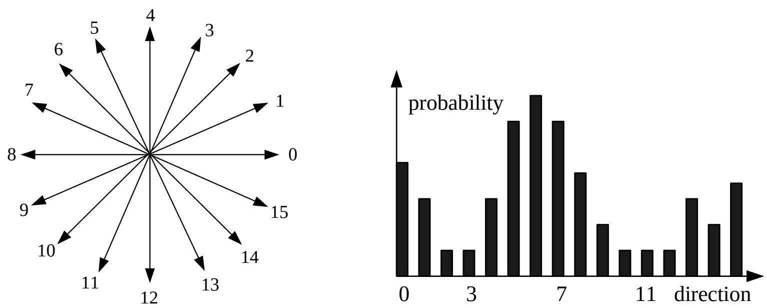


Fig. 3. The normalized motion direction and its histogram

3.4 Fighting behavior detection

When fighting behavior occurs in the motion region, the features f_U, σ_R, E_o and M_o may be large, and M_R may be in a stable range. Meanwhile, other activities do not meet these statistical characteristics. For example, the f_U and M_R values of slow activities such as walking may be similar to the ones of fighting behavior, while the σ_R , E_o and M_o values may be small especially. As for the higher speed activities like running, the f_U value may become very small, E_o and M_o values may become small correspondingly, but the M_R value may become large obviously. According to these statistical characteristics, we use the Joint Gaussian Membership Function to normalize the features:

$$Nf_i = \begin{cases} e^{-\frac{(f_i - c_1)^2}{2\sigma_1^2}} & , f_i \le c_1 \\ e^{-\frac{(f_i - c_2)^2}{2\sigma_2^2}} & , f_i \ge c_2 \\ 1 & ,else \end{cases} \tag{19}$$

Where, f_i represents σ_R , E_o, M_o or M_R. Nf_i is the normalize feature of f_i. c_1, σ_1 are the mean value and variance value of the left membership function curves respectively. c_2, σ_2 are the mean value and variance value of the right membership function curves respectively. c_1, σ_1, c_2 and σ_2 can be determined according to experiments, which will be discussed in Section IV.

Experimental results show that the features have good statistical properties after normalization. When fighting behavior occurs in the motion region, these features can all attain large value; for the other behaviors, some of the features may always attain small value. We use the weighted average method to fuse these features, and then we obtain the feature called Region Violence Index (RVI):

$$RVI_i = \sum_{j=1}^{5} w_j \times Nf_i \tag{20}$$

Where Nf_i represents f_U, M_R, σ_R, E_o or M_o. w_j is a weight value. In this paper, they are set to 0.24, 0.18, 0.18, 0.20 and 0.20 respectively according to experiments.

RVI is a comprehensive index that can reflect trajectory, speed, direction changes and the degree of chaos of movement in the motion region, and it has the strong ability to characterize the fighting behaviors in the scene. There usually are several motion regions in each frame; we select the largest RVI value as the index of the current frame, which is defined as Maximum Violence Index (MVI):

$$MVI = \max\{RVI_i\} \tag{21}$$

Because there are some complicated human behaviors and other unpredictable factors in surveillance scenes, false alarm may be caused in a frame by using MVI to detect fighting behavior. In order to reduce the false alarm rate, we propose the concept of Average Maximum Violence Index (AMVI), which uses the mean value of P frames' MVI to construct the detection criteria for fighting behavior.

$$AMVI = \frac{1}{P}\sum_{j=1}^{P} MVI_j \tag{22}$$

The detection criteria can be described as follows:

$$flag = \begin{cases} 1 & , AMVI \geq T_3 \\ 0 & , else \end{cases} \tag{23}$$

Where, T_3 is a threshold value. If flag equals to 1, it means that fighting behavior occurs in the scene; otherwise, the scene is normal. In this paper, P is set to 15 and T_3 is set to 8 according to experiments.

4. Simulation and Analysis

All the experimental videos are surveillance videos, which contain 330 video clips of fighting scene and 170 video clips of normal scene. Each clip lasts for 10 to 120 seconds, which contains 1 to 4 segments of fighting behaviors and each fighting behavior lasts for 3 to 10 seconds. All the videos are of 352×288 (pixels/frame), 24 (bit/pixel) and 25 (frame/second). **Fig. 4** shows some video frames in these experimental video clips.

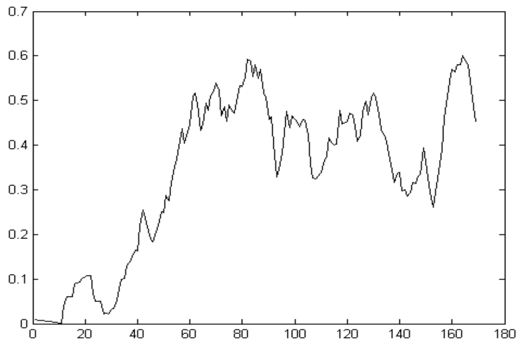
(a) Two persons fight outdoor

(b) Four persons fight outdoor

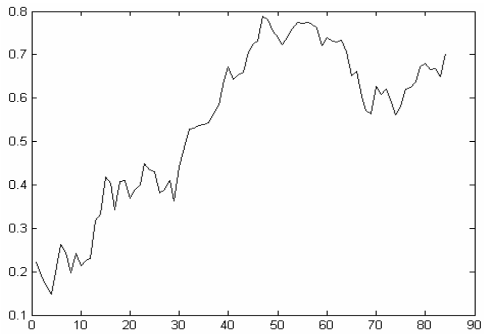
(c) Five persons fight indoor (d) Two persons fight at night

Fig. 4. Video frames in fighting scenes

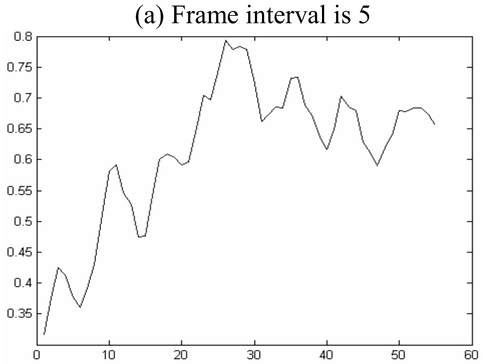
Because fighting behaviors usually last for a period of time, it is not necessary to detect fighting behaviors frame by frame. We can detect fighting behaviors at regular frame intervals. **Fig. 5** shows the AMVI curves extracted from the same video clip when the frame interval is 5, 10, 15 and 20 respectively. From the figure we can see, although the difference of frame interval has an influence on AMVI value, the statistical properties of AMVI value are similar. In experiments, we can achieve the similar performance by using different Gaussian membership parameters according to different frame interval. Due to the page limit, only one group of Gaussian membership parameters is shown in **Table 1**.



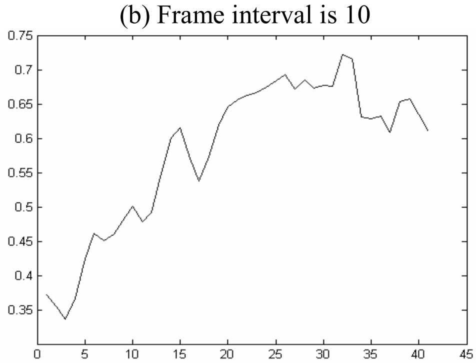
(a) Frame interval is 5



(b) Frame interval is 10



(c) Frame interval is 15



(d) Frame interval is 20

Fig. 5. The AMVI curves with different frame interval

In order to test the performance of our algorithm comprehensively, we classify the experimental videos according to time, site and persons' number, and then calculate the False Alarm Rate (FAR) and Missed Alarm Rate (MAR) under different frame interval. The result of experiments is shown in **Table 2**. Averagely, the FAR value is about 7.7% and the MAR value is about 3.4%. Our algorithm has strong robustness for complicated surveillance scenes.

Table 1. The Gaussian membership parameters of M_R, σ_R, E_o and M_o when the frame interval is 10

	σ_1	c_1	σ_2	c_2
M_R	1.0	4.2	1.4	4.2
σ_R	0.6	2.8	1.0	3.0
E_o	0.4	3.5	0.4	3.5
M_o	1.0	4.0	0.9	4.0

Table 2. Statistics on FAR and MAR for fighting behaviors detection

videos' property		videos' number	frame interval is 5		frame interval is 10		frame interval is 15		frame interval is 20	
			FAR	MAR	FAR	MAR	FAR	MAR	FAR	MAR
time	day	350	7.4%	3.1%	6.3%	1.7%	7.1%	1.7%	8.6%	2.0%
	night	150	8.7%	8.7%	6.7%	6.7%	9.3%	5.3%	9.3%	6.0%
site	indoor	200	8.0%	4.0%	6.0%	2.5%	7.0%	2.0%	10.0%	3.0%
	outdoor	300	7.7%	5.3%	6.7%	3.7%	8.3%	2.7%	8.0%	3.3%
persons' number	≤3	220	6.8%	5.9%	5.0%	3.6%	5.9%	3.6%	7.7%	4.5%
	4~7	200	8.5%	4.0%	7.5%	3.0%	9.0%	1.5%	9.5%	2.5%
	>7	80	8.8%	3.8%	7.5%	2.5%	10.0%	1.3%	10.0%	1.3%

The computation of our algorithm focuses on the segmentation of motion regions and the calculation of motion vectors. During the process of motion region segmentation, there need about $5 \times W \times H$ addition operations. During the process of motion vectors calculation, we only calculate the motion vectors in the motion region. In general, there is about 20% to 40% macro blocks needed to be calculated in the whole frame, and the average search points of each macro block are 17.92. In our experiments, the detection process of one frame costs less than 85 milliseconds on TMS320DM642 platform. The algorithm can detect fighting behaviors lasting for not less than 2 seconds.

Furthermore, we compare our algorithm with existing ones. Because the color of many real surveillance videos is not clear (at night environment), and the boundary contour of persons cannot be obtained easily (when they are far from cameras or overlap with others), only algorithms in paper [8][13] can be used in our experiments, and the performance comparison between our algorithm and them is shown in **Table 3**. From the table we can see, the FAR value of algorithm in paper [8] is too large when the number of persons is above 7, because in this situation the persons' limbs are not clear or overlap with others. Algorithm in paper [13] has large FAR value too, because the accuracy of optical flow reduces with complicated scenes. More over, the computational time of algorithm in paper [13] is too large, which is more than 2000 milliseconds to detect one frame on TMS320DM642 platform. However, our algorithm extracts features from motion regions, not from each person, so we can achieve high speed and strong robustness for fighting behavior detection in real complicated surveillance scenes.

Table 3. Performance comparison between our algorithm and the others

videos' property		videos' number	Algorithm in paper [8]		Algorithm in paper [13]		Our algorithm	
			FAR	MAR	FAR	MAR	FAR	MAR
persons' number	≤3	220	5.0%	4.8%	6.2%	4.2%	5.0%	3.6%
	4~7	200	18.2%	3.3%	9.9%	3.6%	7.5%	3.0%
	>7	80	31.8%	2.0%	12.0%	2.4%	7.5%	2.5%

5. Conclusion

This paper proposes a fast and robust algorithm for fighting behavior detection based on motion vectors. The new algorithm has two advantages: 1) it is fast enough to meet the requirement of real-time detection on embedded platform; 2) it is robust enough to adapt to complicated surveillance scenes with different illumination environment and crowded site. However, this algorithm can only be valid when the surveillance cameras are static, and the research towards to moving cameras will be the future work.

References

[1] T. Ko, "A Survey on Behavior Analysis in Video Surveillance for Homeland Security Applications," in *Proc. of IEEE Conf. on Applied Imagery Pattern Recognition Workshop*, pp. 1-8, Oct. 15-17, 2008. Article (CrossRef Link)

[2] X. Xue, T.C. Henderson,"Video-based Animal Behavior Analysis from Multiple Cameras," in *Proc. of IEEE Conf. on Multisensor Fusion and Integration for Intelligent Systems*, pp. 335-340, Sep. 2006. Article (CrossRef Link)

[3] G. Zhu, Q. Huang, C. Xu, L. Xing, W. Gao, H, Yao, "Human Behavior Analysis for Highlight Ranking in Broadcast Racket Sports Video," in *Proc. of IEEE Trans. on Multimedia*, pp. 1167-1182, Oct. 2007. Article (CrossRef Link)

[4] C. Liu, P. Chung, Y. Chung, "Human Home Behavior Interpretation from Video Streams," in *Proc. of IEEE Conf. on Networking, Sensing and Control*, pp.192-197, Mar. 21-23, 2004. Article (CrossRef Link)

[5] J. Tao, M. Turjo, M. Wong, M. Wang, Y. Tan, "Fall Incidents Detection for Intelligent Video Surveillance," in *Proc. of IEEE Conf. on Information, Communications and Signal Processing*, pp. 1590-1594, Sep. 11, 2006. Article (CrossRef Link)

[6] F. Naghiu, D. Pescaru, "Influence of Driver Behavior Patterns in Correcting Video Sensing Errors in Traffic Surveillance Applications," in *Proc. of IEEE Conf. on Intelligent Computer Communication and Processing*, pp. 173-176, Aug. 27-29, 2009. Article (CrossRef Link)

[7] M. Ghazal, C. Vazquez, A. Amer, "Real-Time Automatic Detection of Vandalism Behavior in Video Sequences," in *Proc. of IEEE Conf. on Systems, Man and Cybernetics*, pp. 1056-1060, Oct. 7-10, 2007. Article (CrossRef Link)

[8] A. Datta, M. Shah, N.D.V. Lobo, "Person-on-Person Violence Detection in Video Data," in *Proc. of IEEE Conf. on Pattern Recognition,* pp. 433-488, Aug. 11, 2002. Article (CrossRef Link)

[9] X. Wu, Y. Ou, H. Qian, "A Detection System for Human Abnormal Behavior," in *Proc. of IEEE Conf. on Robots and Systems*, pp. 1204-1208, Aug. 2-6, 2005. Article (CrossRef Link)

[10] A. Mecocci, F. Micheli, "Real-Time Automatic Detection of Violent-Acts by Low-Level Colour Visual Cues," in *Proc. of IEEE Conf. on Image Processing*, pp. 345-348, Sep. 16-Oct. 19, 2007. Article (CrossRef Link)

[11] A. Mecocci, F. Micheli, "Real-Time Recognition of Violent Acts in Monocular Colour Video Sequences," in *Proc. of IEEE Workshop on Signal Processing Applications for Public Security and Forensics*, pp. 1-4, Apr. 11-13, 2007. Article (CrossRef Link)

[12] W. Wang, R. Wang, Y. Chen, "The Abnormal Behavior Analysis of Single Person on the Road based on Region and Behavior Features," *MIPPR, SPIE*, vol. 6786, pp. 67862N-1, Nov. 15, 2007. Article (CrossRef Link)

[13] K. Hayashi, M. Seki, T. Hirai, T. Koichi, "Real-Time Violent Action Detector for Elevator," *Proceedings-SPIE The International Society for Optical Engineering*, vol. 6051, pp. 60510R, Dec. 5, 2005. Article (CrossRef Link)

[14] K. Supriya, P.S. Sastry, "Abnormal Activities Detection in Video Sequences using Learnt Probability Densities," in *Proc. of Conf. on Convergent Technologies for Asia Pacific Region*, pp. 369-372, Oct. 15-17, 2003. Article (CrossRef Link)

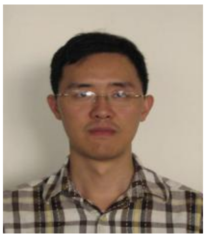
[15] C.-P. Lee, W.-L. Woon, K,-M. Lim, "Statistical and Entropy based Human Motion Analysis,"

KSII Transactions on Internet and Information Systems, vol. 6, pp. 1194-1208, 2010. Article (CrossRef Link)

[16] A. Barjatya, "Block Matching Algorithms for Motion Estimation," *DIP 6620 Spring 2004 Final Project Paper*, Utah State University, 2004.

Jianbin Xie received his B.S., M.S. and Ph.D. degrees from the National University of Defense Technology in 1993, 1995 and 1999. He is currently working as a research fellow at National University of Defense Technology. His main research interests include biometric identification, pattern recognition and machine learning.

Tong Liu received his B.S. and M.S. degrees from the National University of Defense Technology in 2004 and 2006. He is currently working toward a Ph.D. degree at National University of Defense Technology. His current research interests include biometric identification, pattern recognition and machine learning.



Wei Yan received his B.S. and M.S. degrees from the National University of Defense Technology in 2000 and 2006. He is currently working as an instructor at National University of Defense Technology. His current research interests include biometric identification, pattern recognition and machine learning.

Peiqin Li received his B.S. and M.S. degrees from the National University of Defense Technology in 2003 and 2007. He is currently working as an instructor at National University of Defense Technology. His current research interests include biometric identification, pattern recognition and machine learning.

Zhaowen Zhuang received his B.S. and M.S. degrees from the National University of Defense Technology in 1981 and 1984, and received his Ph.D. degrees from Beijing Institute of Technology in 1989. He is currently working as a professor in the National University of Defense Technology. His current research interests include target recognition and satellite navigation.