Abnormal motion detection in an occlusion environment

著者	Mudjirahardjo Panca, Tan Joo Kooi, Kim Hyoungseop, Ishikawa Seiji
journal or publication title	Proceedings of SICE Annual Conference 2013
page range	1398-1402
year	2013-09
URL	http://hdl.handle.net/10228/00006659

SICE Annual Conference 2013
September 14-17, 2013, Nagoya University, Nagoya, Japan

Abnormal Motion Detection in an Occlusive Environment

Panca Mudjirahardjo, Joo Kooi Tan, Hyoungseop Kim and Seiji Ishikawa

Department of Control Engineering, Kyushu Institute of Technology, Kitakyushu, Japan
E-mail : {panca,etheltan,ishikawa}@ss10.cntl.kyutech.ac.jp, kimhs@cntl.kyutech.ac.jp
1-1 Sensuichou,Tobata-ku, Kitakyushu City, Fukuoka, 804-8550, Japan
Tel :093-884-3191

Abstract: We present a motion classification approach to detect movements of interest (abnormal motion) based on optical flow. By tracking all feature points of a moving human in successive frames, we calculate the coordinate space and create feature space. This is done directly from the intensity information without explicitly computing the underlying motions. It requires no foreground segmentation, no prior learning of activities, no motion recognition and no object detection. First, we determine the abnormal scene and speed by using the velocity histogram. Then by using *k*-means clustering over velocity orientation and magnitude, we determine the abnormal direction. The performance of the proposed method is experimentally shown.

Keywords: Harris corner detector, Lucas-Kanade tracker, abnormal motion, velocity histogram, *k*-means clustering.

1. INTRODUCTION

In recent years, abnormal motion detection has attracted great research attention in computer vision. Most current surveillance systems only provide *reactive* security by enabling the analysis of events after the event has already occurred. What is really needed by the security community is, however, *proactive* security to help prevent future attacks.

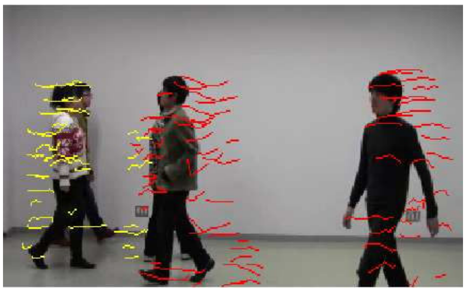
Many approaches of video event analysis are based on the object trajectories extracted from video. Abnormal events can be detected through a prior learning of normal events or without a learning process by analyzing the trajectory result directly.

Jiang et al. [1] used spatial and temporal context and performed frequency-based analysis to detect anomalous video events. The normal observation is modeled by Hidden Markov Model (HMM). This research detected an anomalous car trajectory on the road from top view. Kiryati et al. [2] recognized an abnormal human behavior from high camera view. Before the detection phase, they included the training phase for normal condition. Baranwal et al. [3] detected an abnormal motion indoors and in a static background environment. They trained various motions using radial basis functions networks (RBFN). Park et al. [4] used clustering of motion based on similarity measurement of a feature space. They detected an abnormal motion, especially in a different direction case, from high camera view.

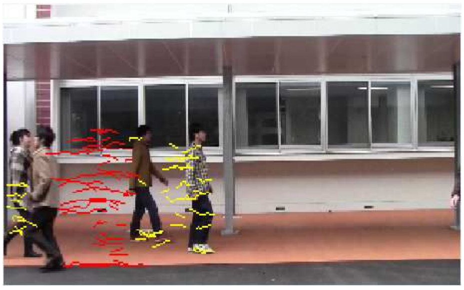
In this paper, we propose an abnormal motion detection in an occlusive environment, which occur often in real life. We capture a scene of a 2 meter height, indoor and outdoor scene, as shown in **Fig. 1**. It requires no foreground segmentation, no prior learning of normal motion, no motion recognition and no object detection. We analyze the trajectory of the points tracked by Lucas-Kanade tracker.

2. OVERVIEW OF THE PROPOSED METHOD

Definition of abnormal motion in this paper is a human motion with different speed and orientation rather than speed and orientation of motion of a group people in a scene, whether in the same direction or opposite direction.



(a) scene A



(b) scene B

(c) scene C

Fig. 1. Examples of successful abnormal motion detection: (a) An indoor case, and (b) an outdoor case: The abnormality is in opposite direction against normal motion. (c) An outdoor case, where the abnormality is in the same direction but faster speed against normal motion.

Based on this definition, we describe the proposed system in the following.

There are three main processes in the proposed method. The first process is the tracking of human motion. We extract tracking points on motion objects in every two successive frames by using Harris corner detector followed by the Lucas-Kanade tracker.

The second process is creation of coordinate spaces. After tracking the motion of objects, there are many tracking points belong to a moving human and noise. To reduce the noise caused by the change of light intensity, we eliminate static points. Then, we create coordinate spaces. Coordinate spaces are the tracking point series from the observation frames. Each feature point in the last frame will have one coordinate space, which shows a trajectory of the point in an observation time.

Next, we create feature spaces as in [4]. A feature space has four dimensions. It consists of velocity of x, velocity of y, magnitude of velocity, and orientation of the velocity.

The third process is detection of abnormal objects and speed. This detection process analyzes the velocity histogram. The histogram shows the velocity of each point in an observation time. With this histogram, we can decide whether abnormality has occurred or not in a scene and detect its speed.

After abnormality has been detected, we determine the abnormal direction by using k-means clustering over velocity orientation and magnitude. **Figure 2** depicts the overview of the proposed system.

3. METHOD

3.1. Extracting Feature Points

Harris corner detector, a popular feature point detector, is applied to extract feature points in a given image. The Harris corner detector is based on the local auto-correlation function of a signal, where the local auto-correlation function measures the local changes of the signal with patches shifted by a small amount in different directions.

For a small shift $[u,v]$, we have bilinear approximation as follows:

$$E(u,v) \cong [u,v] M \begin{bmatrix} u \\ v \end{bmatrix} \tag{1}$$

where M is a 2×2 matrix of the following form computed from image derivatives;

$$M = \sum_{x,y} w(x,y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \tag{2}$$

The Harris measure of a corner is defined by

$$R = \det M - k(\operatorname{trace} M)^2 \tag{3}$$

where

$$\begin{aligned} \det M &= \lambda_1 \lambda_2 \\ \operatorname{trace} M &= \lambda_1 + \lambda_2 \end{aligned} \tag{4}$$

Here λ_1 and λ_2 are the eigenvalues of the matrix M.

Find the points with larger corner response function R (R > threshold), and take the points of local maxima of R.

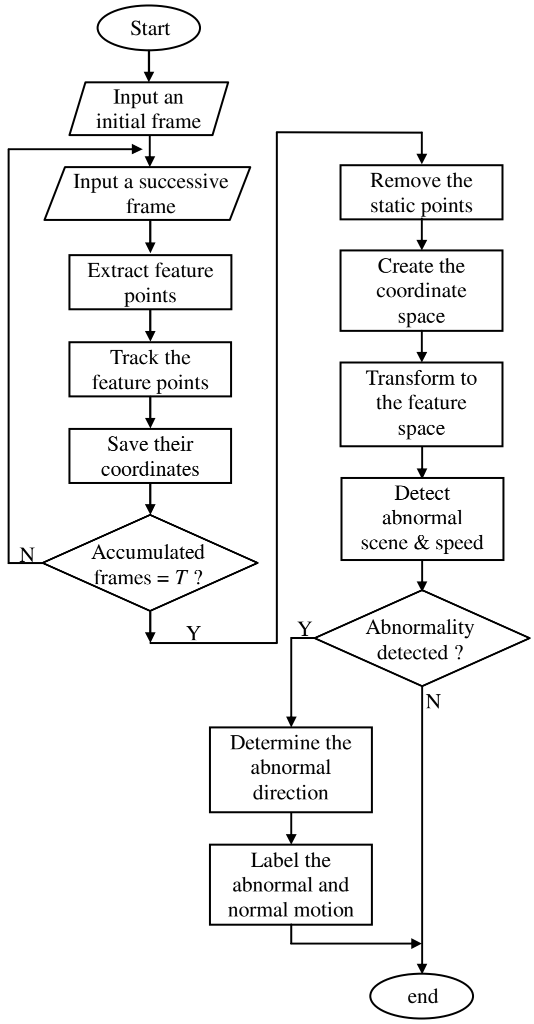


Fig. 2. Overview of the proposed abnormal motion detection method

3.2. Tracking Feature Points

The Lucas-Kanade tracker, one of the most well-known feature points tracking algorithms is employed in the proposed method.

The L-K algorithm relies only on local information that is derived from some small windows surrounding each of the points of interest. Based on the condition, we can get the final expression in the form of

$$d = \left(A^T A\right)^{-1} A^T b \tag{5}$$

where

$$A = \begin{bmatrix} I_x(x_1,t) & \dots & I_x(x_n,t) \\ I_y(x_1,t) & \dots & I_y(x_n,t) \end{bmatrix}^T \tag{6}$$

The disadvantage of using small local windows in the

Lucas-Kanade algorithm is that large motions can move points outside of the local window and thus become impossible for the algorithm to find them. This problem led to the development of the pyramidal L-K algorithm, which starts tracking from the highest level of an image pyramid and working down to lower levels. Tracking over image pyramids allows large motion to be caught by local windows. **Figure 3** shows the image of motion vectors.

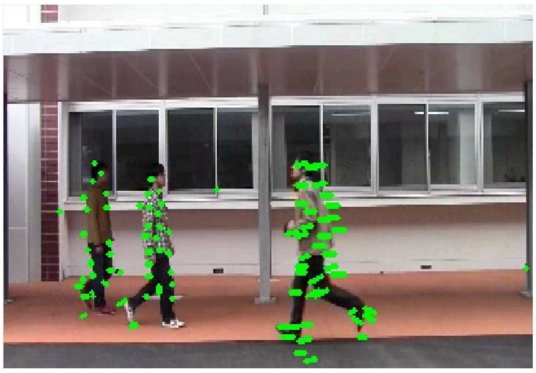


Fig. 3. Motion vectors

3.3. Creating the Feature Space

Since it is difficult to detect an abnormal motion only by the motion vectors between successive frames, we need to accumulate some frames from the first frame.

The feature points are tracked over the entire frames and their location information is stored into a coordinate space [4]. Suppose that a feature point n (n=0,1,2,…,N-1) is tracked through T image frames and its position on the frame t (t=0,1,2,…,T-1) is denoted by $(x_1^{(n)}, y_1^{(n)})$. We then define a sequence of T coordinates of the feature point by the following form [4];

$$
\begin{aligned}
X_0 &= \left[x_0^{(0)}, y_0^{(0)}, x_1^{(0)}, y_1^{(0)}, \ldots\ldots, x_{T-1}^{(0)}, y_{T-1}^{(0)}\right] \\
X_1 &= \left[x_0^{(1)}, y_0^{(1)}, x_1^{(1)}, y_1^{(1)}, \ldots\ldots, x_{T-1}^{(1)}, y_{T-1}^{(1)}\right] \\
&\vdots \\
X_{N-1} &= \left[x_0^{(N-1)}, y_0^{(N-1)}, x_1^{(N-1)}, y_1^{(N-1)}, \ldots\ldots, x_{T-1}^{(N-1)}, y_{T-1}^{(N-1)}\right]
\end{aligned} \tag{7}
$$

Figure 4 shows the tracking result throughout the T frames.

However, the movement cannot be known only by the position information of points in the coordinate space. Therefore the coordinate space is converted to the feature space. There are three kinds of feature spaces such as velocity (8),velocity magnitude (9), and velocity orientation (10). They are defined respectively as follows;

$$
\begin{aligned}
V_0 &= \left[v_x^{(0)}, v_y^{(0)}\right] = \left[x_{T-1}^{(0)} - x_0^{(0)}, y_{T-1}^{(0)} - y_0^{(0)}\right] \\
V_1 &= \left[v_x^{(1)}, v_y^{(1)}\right] = \left[x_{T-1}^{(1)} - x_0^{(1)}, y_{T-1}^{(1)} - y_0^{(1)}\right] \\
&\vdots \\
V_{N-1} &= \left[v_x^{(N-1)}, v_y^{(N-1)}\right] = \left[x_{T-1}^{(N-1)} - x_0^{(N-1)}, y_{T-1}^{(N-1)} - y_0^{(N-1)}\right]
\end{aligned} \tag{8}
$$

Fig. 4. Tracking result throughout T frames

$$
\begin{aligned}
|V_0| &= \sqrt{\left(x_{T-1}^{(0)} - x_0^{(0)}\right)^2 + \left(y_{T-1}^{(0)} - y_0^{(0)}\right)^2} \\
|V_1| &= \sqrt{\left(x_{T-1}^{(1)} - x_0^{(1)}\right)^2 + \left(y_{T-1}^{(1)} - y_0^{(1)}\right)^2} \\
&\vdots \\
|V_{N-1}| &= \sqrt{\left(x_{T-1}^{(N-1)} - x_0^{(N-1)}\right)^2 + \left(y_{T-1}^{(N-1)} - y_0^{(N-1)}\right)^2}
\end{aligned} \tag{9}
$$

$$
\begin{aligned}
\theta_0 &= \arctan\left(\frac{y_{T-1}^{(0)} - y_0^{(0)}}{x_{T-1}^{(0)} - x_0^{(0)}}\right) \\
\theta_1 &= \arctan\left(\frac{y_{T-1}^{(1)} - y_0^{(1)}}{x_{T-1}^{(1)} - x_0^{(1)}}\right) \\
&\vdots \\
\theta_{N-1} &= \arctan\left(\frac{y_{T-1}^{(N-1)} - y_0^{(N-1)}}{x_{T-1}^{(N-1)} - x_0^{(N-1)}}\right)
\end{aligned} \tag{10}
$$

3.4. Detection of abnormal scene and speed

To detect abnormal scene and speed, we propose a method based on a velocity histogram. A resultant image of tracking and their corresponding velocity histogram are shown in **Fig. 5**.

Due to our definition that a normal motion is a motion performed by a group of people, its amplitude of histogram should be greater than the abnormal one. Based on velocity histogram in Fig. 5 (b), the property of the velocity histogram belonging to an abnormal scene is described in **Fig. 6**, where w is the difference between $v_{\max}$ and $v_{\min}$. Then the abnormal scene is determined as follows;

$$
\text{scene} = \begin{cases} \text{normal} & \text{if } w < \text{threshold} \\ \text{abnormal} & \text{if } w \geq \text{threshold} \end{cases} \tag{11}
$$

The abnormal speed is determined as follows;

$$\text{abnormal speed} = \begin{cases} \text{high speed} & \text{if } 0 \le v_{\Sigma\text{peak}} \le v_{th} \\ \text{low speed} & \text{if } v_{th} \le v_{\Sigma\text{peak}} \le v_{\max} \end{cases} \tag{12}$$

where $v_{\Sigma\text{peak}}$ is a velocity at a peak value of number of points.

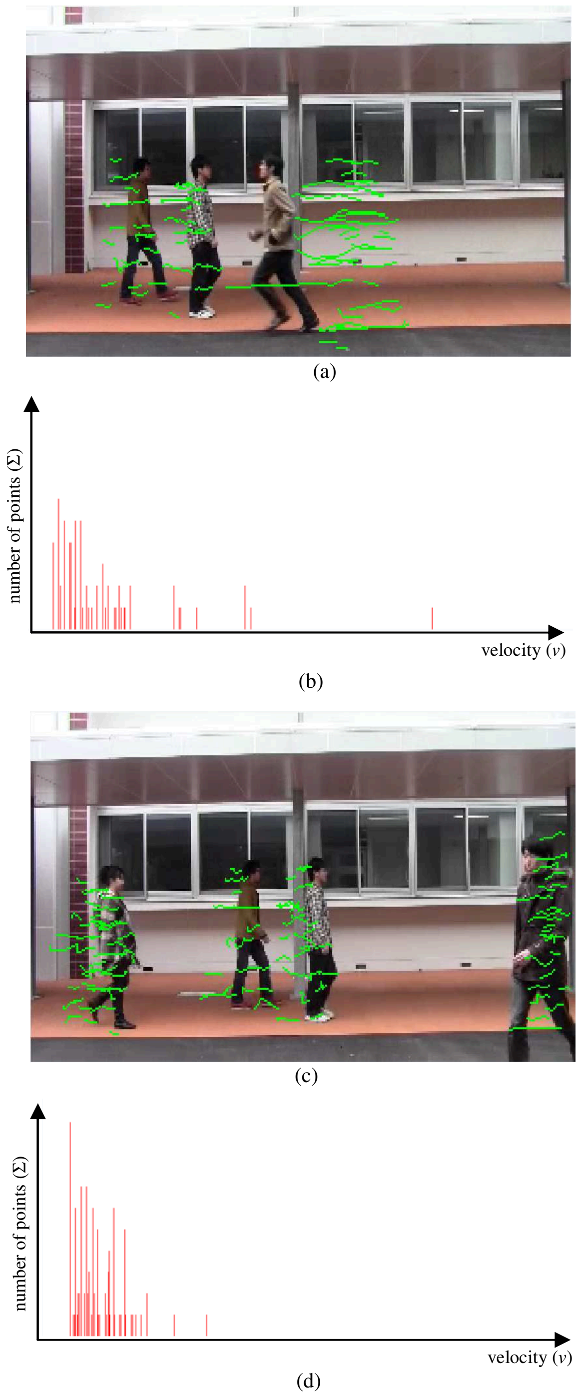


Fig. 5. Result of tracking and their corresponding velocity histogram: (a) An abnormal scene, (b) velocity histogram of (a), (c) a normal scene, (d) velocity histogram of (c)

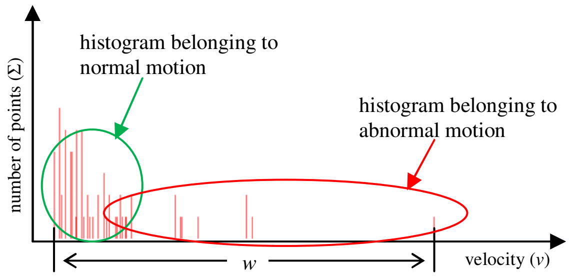


Fig. 6. Property of a velocity histogram to determine a normal or an abnormal scene

3.5. Determining abnormal direction

We determine the abnormal direction by using *k*-means clustering over velocity orientation, θ, and velocity magnitude, $|v|$. The velocity orientation and the corresponding abscissa are shown in **Fig. 7**. To cluster the velocity orientation on the abscissa, we need to shift the original abscissa shown in Figure 7 (b) into a new abscissa shown in Fig. 7 (c). Then we can perform *k*-means clustering.

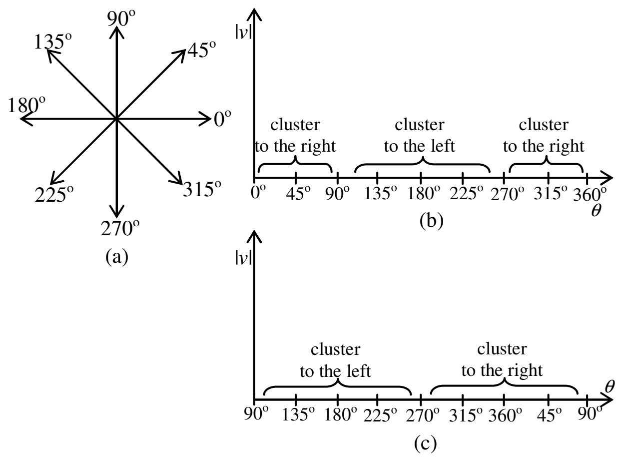


Fig. 7. Finding an abnormal direction of the velocity: (a) Velocity orientation, (b) corresponding abscissa of (a), (c) a shifted abscissa of (b)

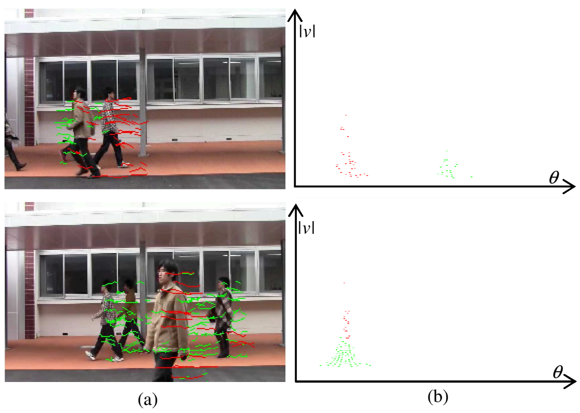


Fig. 8. The result of *k*-means clustering: (a) Tracking images, (b) their clustering by *k*-means algorithm

4. EXPERIMENTAL RESULTS

For experiment, we use indoor and outdoor scene, as shown in Fig. 1, where many people do normal motion (walking) and a person do abnormal motion (running and toward a camera). The video frame rate and the size of an image are 30 fps and 320x180 pixels, respectively.

The observation frames consist of 10 successive frames. The experimental environment is as follows: Operating system is Windows 7 ultimate; processor is Intel® core™ i7 CPU 870 @2.93GHz and the used software is Microsoft Visual Studio 2010. The average execution time per observation frames is 247 ms.

The experimental results are shown in **Fig. 9**. Red tracking lines are abnormal motion, whereas green tracking lines are normal motion.

5. CONCLUSION

In this paper, we proposed an abnormal motion detection method in an occlusive environment, which occur often in real life. We presented a motion classification approach to detect movements of interest (abnormal motion) based on optical flow. This is done directly from the intensity information, without explicitly computing the underlying motions. It requires no foreground segmentation, no prior learning of activities, no motion recognition and no object detection.

As future work, we are going to conduct experiments on the recognition of abnormal motion under stronger occlusion.

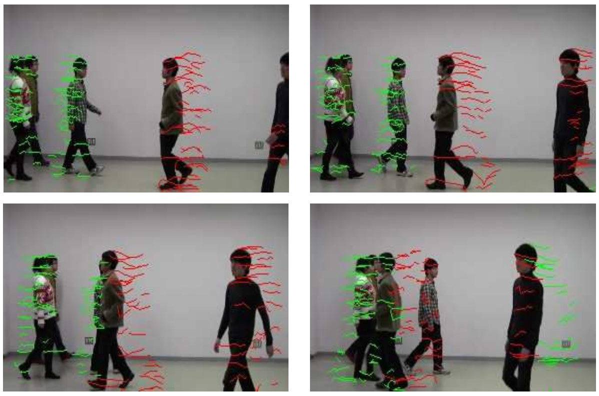
REFERENCES

[1] F. Jiang, J. Yuan, S.A. Tsaftaris, A.K. Katsaggelo, "Anomalous video event detection using spatiotemporal context". *Computer Vision and Image Understanding* 115, pp. 323-333, 2011.

[2] N. Kiryati, T.R. Raviv, Y. Ivanchenko, S. Rochel, "Real-time abnormal motion detection in surveillance video". *Proc. ICPR*, 2008.

[3] M. Baranwal, M. T. Khan, C. W. De Silva, "Abnormal motion detection in real time using video surveillance and body sensors". *International Journal of Information Acquisition.* Vol. 8, No. 2, pp. 103-116, 2011.

[4] M. Park, J. K. Tan, Y. Nakashima, H. Kim, S. Ishikawa, "Detecting human flows on a road different from main flows". *Proc. Int. Sympo. on AROB*, 2011.



(a)

(b)

(c)

Fig. 9. Performance of abnormal motion detection from various scenes. The red line in the frames shows the abnormal notion of a person having a different motion from others. Time elapses in the order of upper left, upper right, lower left and lower right image.