韓國 컴퓨터情報學會 2008 제38차 하계학술발표논문집
第16卷 第1號, 2008. 06.

출입 이벤트 인식

취 야오환*, 이 창 우*

Event recognition of entering and exiting

Yaohuan Cui*, Changwoo Lee*

요 약

Visual surveillance is an active topic recently in Computer Vision. Event detection and recognition is one important and useful application of visual surveillance system. In this paper, we propose a new method to recognize the entering and exiting events based on the human's movement feature and the door's state. Without sensors, the proposed approach is based on novel and simple vision method as a combination of edge detection, motion history image and geometrical characteristic of the human shape. The proposed method includes several applications such as access control in visual surveillance and computer vision fields.

▸ Keyword : Visual Surveillance; Event recognition; entering and exiting; Motion History Image

• 제1저자 : 취야오환
* 한국군산대학교 컴퓨터공학과

1. Introduction

Visual surveillance system usually attempts to detect, recognize and track certain objects from image sequences understand and describe their behaviors. Nowadays, a computer is not only looking at people but also understanding people. Numerous event detection and recognition methods have been proposed for visual surveillance system. For example, the author automatically detects and recognizes unusual events on stairs from video data to rapidly find and analyze the events of interest within large quantities of video data [7], fall detection works well to detect human falls in real time for the enhanced safety in [6]. Michael Spann developed event detection in the traffic system to detect tailgating for intelligent car park [8]. As a human, it is easy to identify a man is entering or leaving the room. But as a computer, it can not tell them apart or even detect one of such complex events as easily as human can do. In the proposed vision based method entering and exiting events are recognized by a combined application of edge detection, motion history image and geometrical characteristic of the human shape.

2. Event recognition method

Event can be considered as a series of continuous actions. Different events always could be composed by different actions. There must be an event occurred when the door's open regardless of entering or exiting. In this paper, entering or exiting event is defined as a combination of door's movement and human activity. With our own definition, a brief overview of the proposed method is illustrated in Figure 1.

No matter entering or exiting, usually the door's state changes same in most instances as closed, open, and then closed. But the significant difference is that human activity differs when the door's states are the same. So certain feature must be extracted as a unique description of the difference between entering and exiting events. In the proposed method the changing laws of the bounding box's size and the relation of bounding boxes' position in single motion history image are selected as the features to tell apart the entering and exiting events, combined with the door's state.

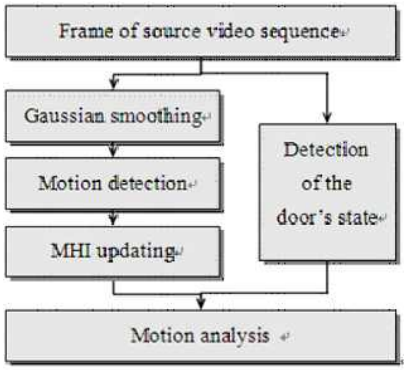


Figure 1: Overview of the framework

2.1 Detection of the door's state

In order to detect the door's state, the moving doorframe is detected with Canny edge detector and Probabilistic Hough Transform. A region of interest (ROI) is set in the higher position of the door manually to decrease the calculation and avoid the obstruction that due to human's motion. Additionally, Gaussian smoothing algorithm is applied to the ROI for denoising.

Table 1: The judgment of the door's state

```
If (the moving doorframe is detected)
 {If (distance between the doorframe and the side of the door < 2
       pixels)
       Return the door is closed
      Else   return the door is opening
 }
 Else return the former state of the door
```

As always in most right side in ROI according to the moving direction in experiments, the moving doorframe is considered as the most right vertical line among the detected vertical lines. Some man-made thresholds are set in Probabilistic Hough Transform (PHT) through experiments. The minimum line length is 28 pixels, while the height of ROI is 30. And the maximum gap between line segments lieing on the same line is 1. It was considered as a vertical line when the abscissa difference of the two points in the detected line is less than 3. A standard of measurement is summarized in Table 1 for door's state detection after the moving doorframe detection.

2.2 Motion detection and MHI updating

Nearly every visual surveillance system starts with motion detection. Motion detection aims at segmenting regions corresponding to moving objects from the rest of an image [3]. At present, most motion segmentations use the methods in background subtraction, temporal differencing and optical flow. In the proposed method temporal differencing was used to detect motion and prepared for MHI updating. The MHI is a static image template where pixel intensity is a function of the recency of motion in a sequence [1]. It is constructed by layering motion regions over time with an update rule in Equation. (1).

$$mhi(x,y)=\begin{cases} timestamp & silh(x,y)\neq 0 \\ 0 & silh(x,y)=0 \\ & and\ mhi(x,y)<timestamp-duration \\ mhi(x,y) & otherwise \end{cases} \quad \cdots\cdots (1)$$

As in equation, the binary sequence silh (x, y) indicating region of motion is extracted from the original image sequence by image differencing. Timestamp is the current time in milliseconds or other units, and the duration means the maximal duration of motion track in the same units as timestamp. Some MHI examples are shown in Figure 2.

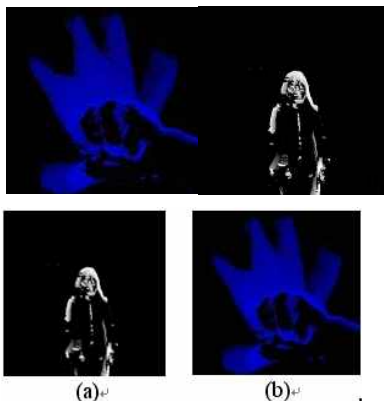


Figure 2: (a) MHI for walking straight forward to camera (b) MHI for making a fist

In the proposed method, the useful MHI could offer not only the moving human's current position but also historical movement information during certain duration in a single image. And this motion history information is effectively used in the following motion analysis step.

2.3 Motion analysis

Similar with the method described in [4], after getting the MHI, the whole moving region for the updating duration could be enclosed by a minimum bounding box as shown in Figure 3, which we call as global bounding box.

Figure 3: Global bounding box located in MHI

The human profile in image would shrink as going steady far away from the camera, vice versa. But leaving or approaching should be determined by not only the size's changing tendency of being greater or smaller but also how it changes, such as enlarging or shrinking.

For tracking the size of the global bounding box, a diagram is constructed as shown in Figure 4. The

y value indicates the scaled size of the global bounding box and x value indicates the frame time. And the frame time that when the door's state is closed and open in the adjacent previous frame is marked with a green line.

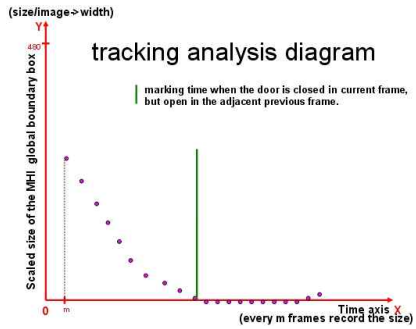
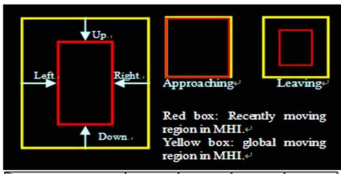


Figure 4: Diagram: "." indicates the scaled size of the current global bounding box

In order to detect the leaving and approaching actions more robustly, our method encloses the most recently moving region in MHI with another rectangle, and extracts the relation of bounding boxes' position as geometrical characteristic for event recognition as illustrated in Figure 5. We call these relations leaving pattern and approaching pattern. The red box denotes the recently moving region while the yellow one shows the whole moving region in MHI. The distances from rectangle of recently moving region to global bounding box in four directions (up, down, left, and right) are calculated in pixels to determine its pattern.

As the human is walking straightly, so the leaving and approaching patterns could be presented in Figure 5. In approaching pattern the absolute distance values in four directions are all smaller than certain threshold, while in leaving pattern they are greater and the sign in up, down, left, right direction is plus, minus, plus, minus respectively.

To determine it is entering or exiting, we define these two events as follows: If the leaving pattern is detected when the door's closed and after the marking time, the average global bounding box size is almost zero, then we define it as an exiting event; else if the size is almost zero when the door's closed, but after marking time, the approaching pattern is detected, then it's an entering event.



Distance (Pixels) / Action	Up	Down	Left	Right
Leaving	> Thr	< -Thr	> Thr	< -Thr
Approaching	<|Thr|	<|Thr|	<|Thr|	<|Thr|

Figure 5: Geometrical characteristic

3. Experiment results

We tested our reference implementation on video sequences captured ourselves using an EVI-D70 SONY camera set on the table about 1.5m above the floor. The test platform is windows XP professional running on a desktop PC with Pentium IV 3 GHz CPU and 1 GB of memory. One test video that contains both entering event and exiting event is described in Table 2 below.

Table 2: Test video sequence

Description
1. One person in the office walks towards the closed door.
2. The person is approaching the door.
3. The door opens; he is going out and closing the door.
4. After a while, one person opens the door, enters and closes the door.
5. The person is approaching to the camera.

The MHI could be extracted from the video sequences with the fixed duration τ as 3s in our experiment as shown in Figure 6 (b). By getting motion information from MHI, the whole moving region is enclosed with a global bounding box. The leaving and approaching pattern are detected as presented in Figure 6 (c).

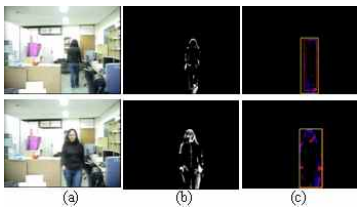


Figure 6: (a) Video frame (b) corresponding MHI (c) corresponding MHI geometrical characteristic

The changing behavior of global bounding box's size in our test video is shown in Figure 7 as illustrated in Section 2.3. The yellow point indicates the scaled size of the global bounding box at corresponding frame time. It's recorded every 15 frames. Also the marking times are shown in green vertical line in the figure below. The red lines indicate the coordinates system.

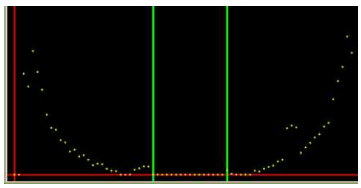


Figure 7: Diagram of the test video

Before the first marking time, y value decreases and leaving pattern could be detected, but after marking time it's almost zero for a while, thus we determined that the exiting event occurs. Meanwhile, entering event happens since after the second marking time y values are increasing and approaching pattern could be found. The distance threshold for our pattern matching is 3 pixels through experiment. Several defined patterns are detected as shown in Figure 8 below.

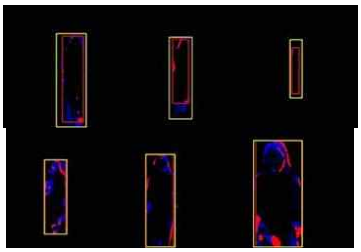


Figure 8: Up row: leaving pattern; down row: approaching pattern

The proposed method also works on the other 34 video sequences including entering or exiting event respectively. The experiment accuracy is shown in Table 3. The false recognition rates are caused by the obstacles occlusion such as chairs and tables in the office room, which scenarios evidently limites the presented framework. However, the theory could also be applied to be extended for further development.

Table 3: Recognizing accuracy of entering and exiting events

	event type	
	entering	exiting
accuracy	91.4% (32/35)	94% (33/35)

4. Conclusion

In this paper, motion detection and geometrical characteristic of the human shape with MHI were proposed as an entering and exiting events detection and recognition method for surveillance system. The proposed method employs Canny edge detection and PHT in order to detect the door's state as a sign of event occurrence. And features extracted from the human shape descript the human behaviors uniquely for event recognition. In comparison of some sensor based event detection and recognition methods, the proposed method had good performance as a vision based method without any precision equipment. In the future, more investigation are required to descript the human behavior better and extend the current single moving object into multi-moving objects environment for more intelligent application system.

References

[1] Intel open source Computer Vision Laboratory http://www.cse.ohi-state.edu/~jwdavis/CVL/Research/MHI/mhi.html

[2] Osama Masoud, Nikos Papanikolopoulos; "A Method for Human Action Recognition" Image and Vision Computing,Volume21, Number8,1August2003.

[3] Weiming Hu; "A Survey on Visual Surveillance of Object Motion and Behaviors" IEEE Systems, Man, and Cybernetics Society, AUGUST2004

[4] Ji Tao, Mukherjee Turjo; "Fall Incidents Detection for Intelligent Video Surveillance" Information Communications and Signal Processing 2005 Fifth International Conference;

[5] A. Bobick and J. Davis; "The recognition of human movement using temporal templates" In IEEE Transactions on Pattern Analysis and Machine Intelligence, volume 23, pages 257-267, March 2001.

[6] Jean Meunier; "Fall Detection from Human Shape and Motion History using Video Surveillance" 21st International Conference on AINAW'2007.

[7] Jasper Snoek; "Automated Detection of Unusual Events on Stairs", the 3rd Canadian Conference on 07-09 June2006, CRV.

[8] Georgios Diamantopoulos, Michael Spann; "Event detection for intelligent car park video surveillance" Real-Time ImagingVolume 11, Issue 3,June2005.