韓國컴퓨터情報學會 論文誌
第14卷 第4號, 2009. 4.
2009-14-4-2-4

# 상황인식 컴퓨팅을 위한 사람 움직임 이벤트 인식

최요환*, 신성윤**, 이창우***

# Recognition of Events by Human Motion for Context-aware Computing

Yaohuan Cui *, Seong-Yoon Shin **, Chang Woo Lee ***


## 요 약

최근 컴퓨터비젼 분야에서 이벤트 검출 및 인식이 활발히 연구되고 있으며, 도전적인 주제들 중 하나이다. 본 논문에서는 사무실 환경에서 발생할 수 있는 이벤트의 검출 및 인식을 위한 방법을 제안한다. 제안된 방법은 MHI(Motion History Image) 시퀀스(sequence)를 이용한 인간의 모션을 분석하며, 사람의 체형과 착용한 옷의 종류와 색상, 그리고 카메라로 부터의 위치관계에 불변한 특성을 가진다. 제안된 방법은 기존의 방법들 중, 칼라 정보를 이용한 방법에 비해 조명의 변화에 민감하지 않은 장점이 있으며, 관심의 대상이 되는 객체의 외형과 같은 사전지식에 의존하는 방법에 비해 스케일에 민감하지 않은 장점이 있다. 에지검출 기술을 HMI 순서 영상 정보와 결합하여 사람 모션의 기하학적 특징을 추출한 후, 이벤트 인식의 기본정보로 활용한다. 제안된 방법은 단순한 이벤트 검출 프레임웍을 사용하기 때문에 검출하고자 하는 이벤트의 설명만을 첨가하는 것으로 확장이 가능하다. 또한, 제안된 방법은 컴퓨터비젼 기술에 기반한 많은 감시시스템 뿐 아니라 상황인식 기반의 이벤트 검출 시스템에 핵심기술이다.

## Abstract

Event detection and recognition is an active and challenging topic recent in Computer Vision. This paper describes a new method for recognizing events caused by human motion from video sequences in an office environment. The proposed approach analyzes human motions using Motion History Image (MHI) sequences, and is invariant to body shapes, types or colors of clothes and positions of target objects. The proposed method has two advantages; one is thant the proposed method is less sensitive to illumination changes comparing with the method using color information of objects of interest, and


---

• 제1저자 : 최요환 교신저자 : 이창우
• 투고일 : 2009. 03. 11, 심사일 : 2009. 03. 30, 게재확정일 : 2009. 04. 20.
*군산대학교 컴퓨터정보공학과 대학원 **군산대학교 컴퓨터정보공학과 교수 ***군산대학교 컴퓨터정보공학과 교수
※ 이 논문은 2008년도 정부(교육과학기술부)의 재원으로 한국과학재단의 지원을 받아 수행된 연구임(No. R01-2008-000-20642-0(2008)).
※ 이 논문은 2009년 한국컴퓨터정보학회 제39차 동계학술대회에서 발표한 논문("휴먼 모션분석을 통한 이벤트 검출 및 인식")을 확장한 것임.

the other is scale invariance comparing with the method using a prior knowledge like appearances or shapes of objects of interest. Combined with edge detection, geometrical characteristics of the human shape in the MHI sequences are considered as the features. An advantage of the proposed method is that the event detection framework is easy to extend by inserting the descriptions of events. In addition, the proposed method is the core technology for event detection systems based on context-aware computing as well as surveillance systems based on computer vision techniques.

▸ Keyword : 이벤트 인식(Event Recognition), HMI(Motion History Image), 기하학적 특징(Geometrical Characteristics), 휴먼모션분석(Human Motion Analysis), 상황인식 컴퓨팅(Context-aware Computing), 감시시스템(Surveillance System)

## Ⅰ. Introduction

Event detection and recognition has become an important and popular topic recently. One of the most obviously applications of this technology is the visual surveillance system which detects, recognizes, and tracks certain objects from image sequences, moreover, understands and describes their behaviors. In recent years, researches in the event detection and recognition field based on computer vision technology have solved much challenging problems.

In the literature, motion detection methods were designed to sense the movement of objects such as people or cars, and segment regions corresponding to moving objects from the rest static regions of an image. Nowadays, identifying moving objects from a video sequence is a fundamental and critical task in many computer vision fields. At present, there are usually three conventional approaches in motion detection: background subtraction, temporal differencing and optical flow. Background subtraction is a particularly popular method for motion segmentation especially under those situations with a relatively static background. However, it is extremely sensitive to changes of dynamic scenes due to lighting and extraneous events [1-2]. Optical flow [3] is another mostly used method for motion detection by using characteristics of flow vectors of moving objects over time. It is advanced in detecting independently moving objects even with camera motion. However, it is sensitive to noise and a large amount of computational works causes complication. In addition, there are some other approaches for motion detection [4-5]. Based on these motion detection techniques, more researches have been done in various applications [6-7]. Like human motion analysis [8-10], it usually can be divided into three categories: human activity recognition, human motion tracking, and analysis of body and body part movements. Motion detection techniques are mostly used in many real time video surveillance applications [1, 11].

With the technique advance in the analysis of digital video mentioned above, it is becoming possible to detect and recognize some high-level semantic events in video sequences. The considered events are all based on computer vision and usually named visual events. As events usually are related to activities, they could be regarded as composition of action threads. Several challenges exist in the event detection and recognit ion field, for example, a good definition of what an event consists of and the recognition problems due to conflicting indications given by the individual cues. Various techniques are researched to detect and recognize events in video sequences. Most approaches are classified into the following two categories: state-space based [12] and template matching based approaches [11]. The other efficient classifiers such as K-nearest neighbors (KNN), Neural Networks (NN), Support Vector Machine (SVM) and Bayesian [13] could also be trained by features for recognizing human actions. Davis et al. [2] represented simple periodic events by constructing dynamic models of periodic pattern of people's movements. All the above mentioned techniques deal with the data in the pixel domain; therefore they are very computationally expensive.

In event detection and recognition field, many applications have already played a very important role in modern life, ranging from security, finance, public health, medicine, biology, environmental science, manufacturing, astrophysics, business, economics and so on [14-16]. Many previous researchers have proposed kinds of methods to detect and recognize the indoor events similar with the events considered in this paper. As mentioned by Douglas Ayers and Mubarak Shah [17], the action of a person's entering is able to be recognized by using skin detection and prior knowledge about the layout of the room. Olson and Brill [18] developed an automated security system to detect events such as entering, loitering and depositing. Given the reliable head tracking just described, recognition of several actions like getting up and sitting-down was made relatively simple by virtue of the success of the head tracker as proposed by Hammadi Nait-Charif and Stephen J. McKenna [19].

In this paper a vision based indoor environment event recognition approach is presented for visual surveillance systems. These events mainly comprise sitting-down, standing-up, entering, exiting, and approaching, leaving, respectively. Without expensive equipments such as hardware sensors, the abovementioned events are recognized via a combined approach of edge detection, HMI, and geometrical characteristics of human shapes from MHI sequences. The proposed method has a wide range of potential applications such as monitoring systems, human-computer interaction and video indexing and retrieval.

## II. Proposed approach

In this paper, an event is considered as a series of continuous actions. Different events are always composed by different actions. There must be an event occurring when the door's open regardless of whether this event is entering or exiting. In this paper, entering or exiting event is defined as a combination of the door's states and human movements. In addition the geometric features extracted from the human motion analysis are also used for recognizing sitting-down and standing-up events, effectively. With our own definitions, a brief overview of the proposed method is illustrated in Fig 1. An image obtained from the video sequence is applied to the Gaussian smoothing as a low-level preprocessing to remove the noises. In addition, the detection of the door's state is performed with Hough Transform in a pre-defined region of interest(ROI). Meanwhile motion regions are segmented form the static background for updating. Including motion history information, MHIs are used to extract features for motion analysis. Finally, an event is recognized by taking advantage of both motion and appearance information.

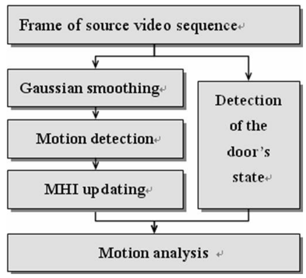



Fig 1. Overview of the framework

### 2.1 Detection of the door's state

In general, entering or exiting event occurs while the door's state changes. Therefore, it is obviously necessary to perform a detection of the door's state for monitoring the changes of the door's states in real time.

To identify the changes of the door's states, the moving doorframe is detected with canny edge detector and Probabilistic Hough Transform(PHT). Since it is less susceptible to occlusion and noise than the bottom part of the door, 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 to remove noises. After performing the PHT, there are usually more than one candidate lines for the doorframe detection. To decide which candidate line is the moving doorframe, the next step is performed as follow. As always in the most right sides in ROI according to the practical situation in experiments, the moving door-frames are considered as the most right vertical line among the detected vertical lines.

Some manually defined thresholds are set in the PHT through the experiments. The minimum line length is 28 pixels, while the height of ROI is 30. And the maximum gap between line segments lying 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.

After the door-frame detection, the distance between the door-frame and the side of the door is considered as the judgment that whether the door is open or closed. A standard of measurement is summarized for door's state determination after the moving door-frame detection. If the size of the door decreases to less than 2 pixels for two consecutive frames, the door is closed; otherwise it is open. If the change of the door state is not detected, the former state is preserved.

## 2.2 Motion detection

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 [20]. In the proposed method temporal differencing is used to detect motion and as an input of the following MHI updating. The referred knowledge of the MHI is illustrated detail in Intel open source Computer Vision Laboratory [21]. The MHI is a new view-based template approach to represent actions. The work is motivated to solve the problem when to recognize actions in extremely low resolution image with no strong features or information about the three-dimensional structure of the scene. The MHI is a static image template where the pixel intensity is a function of the recency of motion in a sequence [21]. So the more recently moving pixels appear brighter than others. The MHI as a kind of static image template is commonly used for human activity recognition such as gait, gesture and fall incidents [14]. The MHI is constructed by layering motion regions over time with an update rule in Eq. (1).

$$mhi_t(x,y)=\begin{cases} timestamp & \text{if } silh(x,y)\neq 0. \\ 0 & \text{if } silh(x,y)=0 \text{ and } \\ & mhi_{t-1}(x,y)<timestamp-duration. \\ mhi_{t-1}(x,y) & otherwise. \end{cases} \quad \text{Eq. (1)}$$

As in equation, the binary sequence silh(x,y) indicating region of motion is extracted from the original image sequence by image differencing. silh(x,y) is a mask that has non-zero pixels where the motion occurs. It is obtained by a binary threshold process after the image differencing step. 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 the timestamp. For each image differencing result, the binary sequence, silh(x,y), is obtained by a threshold process. If the silh(x,y) exists, then the corresponding mhi(x,y) value equals the current timestamp. But if the silh(x,y) equals 0 and the duration between the last and current timestamp is less than or equal to the maximal duration, mhi(x,y) holds its previous value. Otherwise assign zero value to the mhi(x,y).

The MHI is useful in our proposed method, because it offers not only the moving human's current position but also the history information of the human's movement 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

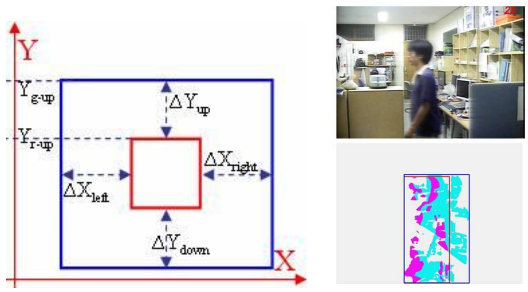
Similar with the method described in [22], after getting the MHI, the whole moving regions are enclosed by a minimum bounding rectangle, also known as bounding box in the two dimensional case as in Fig 2, which we call as global bounding box to distinguish with the other bounding boxes used in the follow section. After background subtraction, the binary image is projected onto the x and y axes. Therefore, the location of left/right borders of the bounding box is calculated according to the x projection. In addition the top and the bottom of the bounding box are calculated similarly. Although the background is noisy with kinds of objects such as bags, cabinets and computer terminals, the motion based background segmentation method has a good effect with the room in disorder. Patrick Peursum [23] outlined foreground by a bounding box after foreground objects are segmented out from the background and tracked based on the centroid of the bounding box. The most recently moving pixels are brighter than others in the MHIs.

(a) Sample bounding boxes


(b) Global bounding boxes

Fig 2. Example of motion detection

The MHI offers important information as the direction of a moving motion. With the aim of seeing clearly, most recently moving pixels are colored with magenta and the other moving pixels are colored with cyan as shown in Fig 3 (c). The outer blue rectangle represents the global bounding box; while the inner red one denotes the most recently moving region in the MHI. Obviously, the inner bounding box is always inside the global bounding box in the same MHI. The relation between the positions of inner and outer bounding boxes is defined as the geometrical characteristic for the event recognition as illustrated in Fig 3 (a). We call these relations as the geometric patterns. The leaving and approaching geometric patterns are illustrated in Section 3.4.



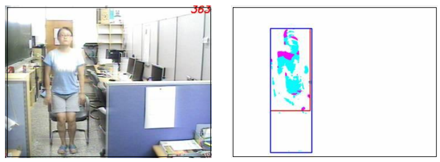

(a) (b)

Fig 3. Example of Geometrical patterns (a) relation between inner and outer bounding boxes, (b) a frame and its colored MHI

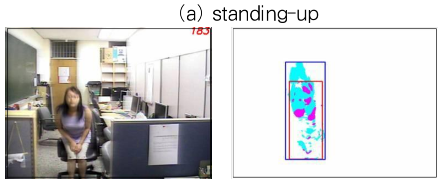
Usually after a motion detection step, it is important to select features that can distinguish these events well, such as appearance information from the moving objects. This includes not only instantaneous information of the spatial features of objects such as width, height, and aspect ratio, but also temporal information about changes in the sizes of objects as well as motion features, such as the direction of movement and speed [24]. In the proposed method, all these information cues which are important features for behavioral analysis are described in two-dimensional coordinate system. In most cases, the x-axis is the frame number and the y-axis is the specific feature value for that frame. In our proposed method, several variables are selected as features from the geometric patterns, and their variation rules are considered to describe human behaviors for the event recognition.

Yg-up is the y value of the up horizontal of global bounding box. Yr-up is the y value of the up horizontal of recently moving region bounding box. The distances between these two bounding boxes are calculated in pixels in four directions (up, down, left, and right) as ΔYup, ΔYdown, ΔYleft, and ΔYright.

For the history information shown in one single MHI and the geometric patterns indicating the changes of human movements, the MHI combined with the geometric patterns is of great use in some recognition work.



(a) standing-up



(b) sitting-down

Fig 4. Sample frames and its MHIs with the geometric patterns

In Fig 4, it is difficult to distinguish whether a person is standing or sitting from a single frame. However the problem is easily solved by analyzing the MHI with the proposed geometric patterns. Obviously a person in Fig 4 (a) is standing when ΔYup becomes zero and ΔYdown makes a great difference. Whereas a person in Fig 4 (b) is sitting when ΔYdown becomes zero and ΔYup makes a great difference.

In addition, as the size of the human body is changed all the time along with its motion, the size of the global bounding box is selected as a feature to represent the events. For tracking the size of the global bounding box, a diagram is constructed as shown in Fig 5. The y value indicates the scaled size of the global bounding box and x value indicates the frame time. And the frame is marked with a green line in the case when the door's state is closed and its state was open in the adjacent previous frame. The marked line denotes the candidate frame that the event occurs. In order to analyze the characteristics of human motion, the other variables in the diagram could also be constructed similarly, except where the y axis denotes Yg-up, Yr-up, or, ΔYup, respectively.

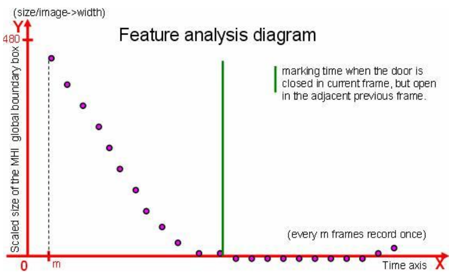


Fig 5. Diagram: "." indicates the scaled size of the current global bounding box

## 2.4 Definition of events

Combined with the detected door's state and the geometrical characteristics of the human shape from the MHI sequences obtained from the previous process, events are ready to be recognized by their definitions. The features obtained from the abovementioned motion analysis are used to recognize the events considered in this paper. Table 1 shows all the events to be recognized in the proposed method.

The human body in image would be shrunk when human is leaving far away from the camera steady, while it would enlarge when human is approaching to the camera. But leaving or approaching could not be determined only by the size's tendency of being greater or smaller of the global bounding box but also how it becomes greater or smaller, such as enlarging or shrinking. Geometric pattern is used to represent the enlarging and shrinking tendency. To recognize approaching and leaving events, leaving and approaching patterns are defined below to robustly distinguish between the two actions.

Table 1. List of events and recognition conditions

| Event | Recognition conditions |
|---|---|
| Approaching | The size of the global bounding box increases continuously, approaching pattern is detected. |
| Leaving | The size of the global bounding box decreases continuously and leaving pattern is detected. |
| Entering | No movement, door opens, person approaches. |
| Exiting | Person leaves, door opens, person disappears. |
| Sitting-down | Yg-up varies from level condition and continuous falling condition, in the same time, Yr-up rises and ΔYup falls continuously. |
| Standing-up | Yg-up varies from continuous rising condition to level condition, in the same time, Yr-up rises continuously and ΔYup stays zero level condition. |

For a human is walking straightly, the leaving and approaching patterns in Fig 6 are illustrated in Table 2. In the approaching pattern the absolute distance values between the two defined bounding boxes are all smaller than a certain threshold in four directions, while in the leaving pattern they are greater. The selected threshold is obtained by experiments.

For two events, entering and exiting, the changes of the door's state usually are in the same order in most instances as follows: closed, open, and then closed. However the significant difference is that human activity differs when the door's states are the same. Thus, a 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 the single MHI are selected as the features to distinguish between entering and exiting events. For recognizing entering/exiting events, the diagram of the scaled global bounding box's size is constructed; and the detection of leaving and approaching patterns defined above to robustly distinguish between the two actions are used.

Table 2. Definition of leaving/approaching pattern

| | Leaving Pattern | Approaching Pattern |
|---|---|---|
| $\Delta$Yup | > threshold | < threshold |
| $\Delta$Ydown | > threshold | < threshold |
| $\Delta$Yleft | > threshold | < threshold |
| $\Delta$Yright | > threshold | < threshold |

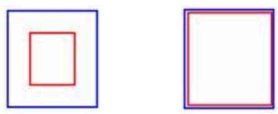


Fig 6.  Leaving and approaching pattern

To determine whether it is entering or exiting, we define these 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 approaching 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 is determined as an entering event.

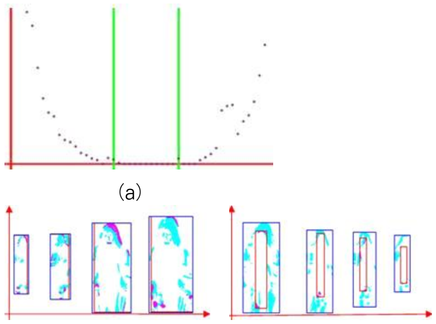
For standing-up and sitting-down, three diagrams of Yg-up, Yr-up and $\Delta$Yup are constructed as the specific features of each event. The time is considered as the candidate time at which sitting-down or standing-up event occurs when the value of Yg-up varies between the level condition and the continuous rising or falling condition. Furthermore, unique features of the changing rule in Yr-up and $\Delta$Yup value are combined to determine the event type. Therefore, if Yg-up varies from the level condition to continuous falling condition and at the same time, Yr-up rises and $\Delta$Yup falls continuously, sitting-down event occurs; if Yg-up varies from the continuous rising condition to the level condition and at the same time, Yr-up rises continuously and $\Delta$Yup stays zero level condition, standing-up event occurs.

## III. Experiment results

The propose method was tested on video sequences with 640 x 480 pixel resolution. We captured video sequences of ourselves using an EVI-D70 SONY camera set on the table about 1.5m above the floor. The pan/tilt/zoom camera has a wise angle, and its features are good advantages for the potentially feature work like tracking. The test platform is Windows XP Professional running on a desktop PC with Pentium IV 3 GHz CPU and 1 GB of memory. To evaluate the proposed method, video sequences are collected in a testing office room, including several events.

The video sequences are captured in different illumination conditions like day and night. And the layout of the room is different. The persons in the video sequences are dressed in different clothes. In all the cases the experiment results are not effects because the proposed method is based on motion information which is invariant to color and illumination conditions, etc. Although sometimes human's head occludes the ROI, the experiment result of the door's state detection has a good accurate. The MHI is extracted from the video sequences with a fixed duration, $\tau$ of 3 seconds in our experiment. By getting the motion information from the MHIs, the whole moving regions are enclosed with a global bounding box.

The size variation of the global bounding box in our test videos is shown in Fig 7. The dot indicates the scaled size of the global bounding box at a corresponding time frame, where the duration between dots is 15 frames. Also the marking times are shown as green vertical lines in the figure below. Before the first marking time, the y value decreases and the leaving pattern is detected. After the marking time however, the y value is almost zero for a while, and thus we determined that the exiting event occurs. Meanwhile, the entering event happens after the second marking time y values are increasing and the approaching pattern is found. The distance threshold for our pattern matching was 3 pixels through the experiment. Several defined patterns are detected as shown in Fig 7.

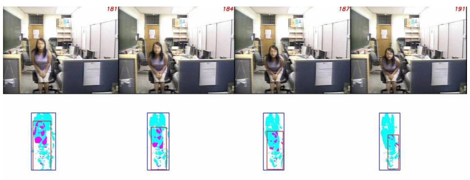


(a)

(b) (c)

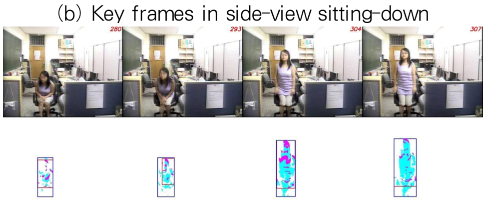
Fig 7. Test results: (a) diagram of the scaled global bounding box's size (b) approaching pattern (c) leaving pattern

Two views – front view and side view – were considered for sitting-down and standing-up event recognition. As shown in Fig 8, accompanied by variations of human motion, the geometric patterns fluctuated regularly from either of the two views.



(a) Key frames in front-view sitting-down

(b) Key frames in side-view sitting-down



(c) Key frames in front-view standing-up

(d) Key frames in side-view standing-up

Fig 8. Key frames of sitting-down and standing-up event in two front- and side-view

Corresponding to the geometric patterns of the same video sequence, the three variables (Yg-up, Yr-up, ΔYup) are plotted as curves in Fig 9. The four vertical dashed lines indicate the sitting-down and standing-up events where the value of Yg-up changes between level and continuously rising or falling. The horizontal axis represents the time reference. When Yg-up varies steady, Yr-up falls rises continuously, and when ΔYup rises simultaneously for a space, sitting-down event occurs; when both Yg-up and Yr-up rise continuously and the ΔYup becomes zero simultaneously for a space, then standing-up event occurs. Although there were some noises in the graph, the geometrical changes of the features represent the sitting-down and standing-up events uniquely.

The method achieves a good performance with single moving object. So when there are more than two moving objects, the implementation needs to be extended further. Thus with the segmentation techniques of the background subtraction, binary connected component analysis and clustering the binary connected components, the proposed method also works on the video sequences with two moving people. In Fig 11, it shows the test result of video sequence with two moving people.

The other relative video sequences are performed by different persons, including 35 entering/exiting and 32 sitting-down/standing-up events respectively. And they are dressed in different clothes.

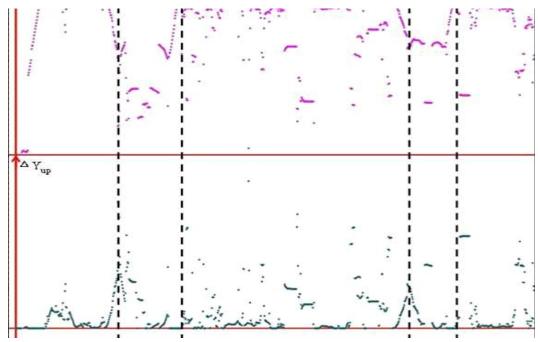



Fig 9. Variation diagram of Yg-up, Yr-up and ΔYup from top down

However the results show that the proposed method is invariant to body shape, clothing and position. The experiment accuracy is shown in Table 3. The false recognition rates are caused by the obstacles occlusion such as chairs, which scenarios evidently limits the presented framework. Because of the obstacles occlusion such as chairs and tables in the office room, the changes of the human bounding box's size will be enlarged and shrink irregularly. And due to the instability of human motion speed, the noise appears lot in diagram processing step.

Fig 11. Test result of two moving people

Table 3. Event recognizing accuracy

| Event type | Accuracy |
|---|---|
| Approaching | 85.0% (30/35) |
| Leaving | 80.0% (28/35) |
| Entering | 91.4% (32/35) |
| Exiting | 94.0% (33/35) |
| Sitting-down | 90.6% (29/32) |
| Standing-up | 93.7% (30/32) |

Comparing with other methods, the proposed method has no need of modeling or training process. As mentioned by Douglas Ayers and Mubarak Shah [7], the action of a person's entering into a room was recognized by using skin detection and prior knowledge about the layout of the room. In some situations, the skin detection based entering/exiting recognition method would be detected as a false positive sample, such as a person opening the door and then leaving without entering. Although skin could be detected in the door area, there is no entering/exiting event actually. Therefore, the proposed method is less sensitive to illumination changes comparing with the method using color information of objects of interest

The analysis of the experimental results revealed that the proposed method takes advantage in many factors. In comparison to some sensor based event detection and recognition methods, the proposed method had good performance as a vision based method without any precision equipment and complicated work. The proposed method is easier to expand then the previous works such as [11-12, 25] without template training, matching and modeling work that cost large computation. Comparing with the method using a prior knowledge like appearances or shapes of objects of interest, the proposed method is invariant to scale changes.

## IV. Conclusion

In this paper, motion detection and geometrical characteristics of the human shape with the MHI sequences were proposed as events detection and recognition method for surveillance systems. The features extracted from the human shape in the MHI sequences describe the human behaviors uniquely for event recognition. The proposed framework is also easy to extend for further development, in which it is simple without any model construction, probability calculation and training work. The proposed method is widely applicable to many applications such as the people counting field for security management.

In the future, more investigations are required to describe various human activities. Obviously, an extension of recognizing human actions such as picking up, stealing is another desirable avenue for future work. It is considered to construct an event retrieval system for visual surveillance with good use of the proposed method.

## 참고문헌


[1] Haritaoglu, Harwood, and Davis, "W4: Real-time surveillance of people and their activities," IEEE Trans. Pattern Anal. Machine Intell, Vol.22, pp. 809-830, Aug. 2000.

[2] Lipton, Fujiyoshi, and Patil, "Moving target classification and tracking from real-time video," in Proc. IEEE Workshop Applications of Computer

Vision, pp. 8 - 14, 1998.

〔3〕 Barron, Fleet, and Beauchemin," Performance of optical flow techniques," Int. J. Comput.Vis., Vol. 12, No. 1, pp. 42 - 77, 1994.

〔4〕 Friedman and Russell, "Image segmentation in video sequences: a probabilistic approach," in Proc. 13th Conf. Uncertainty in Artificial Intelligence, pp. 1 - 3, 1997.

〔5〕 Collins, Lipton, Kanade, Fujiyoshi, and Wixson, "*A system for video surveillance and monitoring*," Carnegie Mellon Univ., Pittsburgh, PA, Tech. Rep., 2000.

〔6〕 Cheung, Kamath, "Robust Background Subtraction with Foreground Validation for Urban Traffic Video," EURASIP Journal on Applied Signal Processing, Vol. 14, pp. 1-11, 2005.

〔7〕 Efros, Berg, Mori and Malik, "Recognizing Action at a Distance," Proceedings of ICCV, Vol. 2, pp. 726-733 2003.

〔8〕 Somboon, Hongeng, "Video-based event recognition: activity representation and probabilistic recognition methods," Computer Vision and Image Understanding, Vol. 96, pp. 129-162, 2004.

〔9〕 Thomas Moeslund, Adrian Hilton, Volker Kr üger, "A survey of advances in vision-based human motion capture and analysis," Computer Vision and Image Understanding, Vol. 104, No. 2, pp. 90-126, November 2006.

〔10〕 Jaimes, A. and Sebe, N., "Multimodal human computer interaction: a survey," Computer Vision and Image Understanding, Vol. 108, pp. 116-134, Oct. 2007.

〔11〕 Aron Bobick and James Davis, "The recognition of human movement using temporal templates," In IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 23, pp. 257 - 267, Mar. 2001.

〔12〕 Venkatesh Babu, Anantharaman, Ramakrishnan and Srinivasan, "Compressed domain action classification using HMM," Pattern Recognition Letters Vol. 23, pp. 1203-1213, Aug. 2002.

〔13〕 Venkatesh Babu and Ramakrishnan, "Recognition of human actions using motion history information extracted from the compressed video," Image and Vision Computing, Vol. 22, pp. 597-607, Aug. 2004.

〔14〕 Jean Meunier, "Fall Detection from Human Shape and Motion History using Video Surveillance," 21st International Conference on AINAW, pp. 875-880, 2007.

〔15〕 Georgios Diamantopoulos, Michael Spann, "Event detection for intelligent car park video surveillance," Real-Time Imaging, Vol. 11, pp. 233-243, June 2005.

〔16〕 Osama Masoud, Nikos Papanikolopoulos, "A Method for Human Action Recognition," Image and Vision Computing, Vol. 21, No. 8, pp. 729-743, 1 Aug. 2003.

〔17〕 Douglas Ayers, Mubarak Shah, "Monitoring human behavior from video taken in an office environment," Image Vision Computing Vol. 19, pp. 833-846, Oct. 2001.

〔18〕 Olson and Brill, "Moving object detection and event recognition for smart cameras," Proceedings of IU , pp. 159-175, 1997.

〔19〕 Hammadi Nait-Charif, Stephen McKenna, "Head Tracking and Action Recognition in a Smart Meeting Room," Proceedings of the 9th international conference on Multimodal interfaces, pp. 86-93, 2003.

〔20〕 Weiming Hu, "A Survey on Visual Surveillance of Object Motion and Behaviors," IEEE Systems, Man, and Cybernetics Society, pp. 825-826, Aug. 2004.

〔21〕 Intel open source Computer Vision Laboratory *http://www.cse.ohi-state.edu/~jwdavis/CVL/Research/MHI/mhi.html*

〔22〕 Ji Tao, Mukherjee Turjo, "Fall Incidents Detection for Intelligent Video Surveillance," Information Communications and Signal Processing 2005 5th International Conference, pp. 1590-1594, 2005.

〔23〕 Patrick Peursum, Svetha Venkatesh, Geoff, "Object Labelling from Human Action Recognition," First IEEE International Conference on Pervasive Computing and Communications, pp. 399, 2003.

〔24〕 Yunqian Ma, Ben Miller, Pradeep Buddharaju, Mike Bazakos, "Activity Awareness: from Predefined Events to New Pattern Discovery," Computer Vision Systems, ICVS, pp. 11-11, January 2006.

〔25〕 Hongying, Nick Pears, "Recognizing human actions based on motion information and SVM," 2nd IET International Conference on Intelligent Environments, pp. v1-239-246, July 2006.

## 저 자 소 개


최요환

2006년 중국해양대학컴퓨터과학과(이학사)

2009년 군산대학교 컴퓨터공학과 (공학석사)

〈관심분야〉 텔레매틱스, 패턴인식, 컴퓨터비젼, 상황인식

신 성 윤

2003년 2월 군산대학교 컴퓨터과학과 이학박사

2006년~현재 군산대학교 컴퓨터정보과학과 교수

〈관심분야〉 비디오 인덱싱, 비디오 요약, 멀티미디어

이 창 우

1996년 경일대학교 컴퓨터공학과(공학사)

1998년 경북대학교 컴퓨터공학과 (공학석사)

1998년~2001년 포항1대학 전임강사

2004년 경북대학교 컴퓨터공학과(공학박사)

2004년~현재 : 군산대학교 조교수

관심분야 : 텔레매틱스, 패턴인식, 컴퓨터비젼, 상황인식