Security and Observation System for Solitary Person Using Accelerometer and Surveillance Camera

Satoshi Hasebe†, Shigeyoshi Nakajima††, ThiThi Zin†† and Takashi Toriu††,

Osaka City Univ., 3-3-138 Sugimoto-cho, Sumiyoshi-ku, Osaka City, 5588585 JAPAN

Summary
Recently people living alone increase according to a change of the state of the society in the world. Notably elderly people living alone increase more and more. The increasing number of hunger deaths and solitary deaths becomes a big social problem. Municipal governments and care service companies provide services monitoring an alone elderly person using a television-phone or a remote camera and deliver a service person in emergency On the other hand, there were many cases that a turnover accident of an alone elderly person causes a solitary death. Then in this study we propose a system to discriminate human postures and behaviors using camera images and accelerometer data to detect quickly a turnover or else. We can discriminate with a camera and an accelerometer a posture which is not discriminable with only a camera. With an acceleration data we can easily discriminate a face up lie from a face down lie. We show the results of our experiments to discriminate postures which are not discriminable only with images also not discriminable only with accelerations and show the efficiency of the proposed method.
Key words:
Acceleration, image, discrimination, posture, behavior

1. Introduction

Turnover accidents of people in houses increase recently especially accidents of elderly people [1]. In houses elderly people may turnover after hitching an electric cord, a rim of a carpet, of a rag. Such cases increase more and more. After the accidents, some broke their bones, couldn't call phone for help, became debilitated and died. An automatic system to watch an alone person in a house, detect an accident and quickly alert to someone is needed now. We can discriminate a standing posture from a lying posture in broad terms using a surveillance camera but cannot discriminate details of postures. And if we know the behavior of the person before the accident or the details of turnover motion we use such information in the first-aid treatment efficiently. So we propose a system to analyze human postures and behaviors according to images from cameras and data from accelerometers in this paper.

There were many recent works to analyze human postures and behaviors with wearable accelerometers. C. Sun et al. [2] analyzed a motion of a human arm with an accelerometer. Y. Kawahara et al. [3] estimated human behaviors of "walking", "running" and "sitting". N. Ravi et al. [4] discriminated 10 behaviors such as "standing," "walking," "running," "climbing stairs" and etc. Others used more accelerometers. A. Gallagher et al. [5] used 3 accelerometers and discriminated 3 behaviors "walking," "running" and "brushing of teeth." S. S. Intille et al. [6] used 5 accelerometers and discriminated some human behaviors. N. Kern et al. [7] discriminated 10 behaviors such as "standing," "sitting," "climbing stairs," "descending stairs," "shaking hands," "writing a word on a blackboard" and "typing a keyboard" etc. with 12 accelerometers in a shoulder, elbow, knee, foot etc.

On the other hand there were studies about behavior patterns of elderly alone persons [8][9].

Our proposal is a sensor fusion of accelerometers and a surveillance camera for application of watching behavior of elderly people living alone.

The rest of this paper is organized as shown below. The section 2 shows the analysis method and the algorithm of the proposed method. The section 3 shows the experiments for evaluation. The section 4 shows the experimental results and the consideration. The section 5 shows the conclusion.

2. Analysis Method and Algorithm

In this section we will explain the proposed the analysis method and algorithm of the proposed method to discriminate postures and behaviors according to acceleration data from accelerometers and images from a camera. We talk about the detail of the method to extract a region of a human silhouette, detecting a posture, detecting posture and motion according to acceleration data from accelerometers.

Manuscript received March 5, 2011
Manuscript revised March 20, 2011

2.1 Process of Images from Camera

Images from a camera is preprocessed first. The color images are changed to monochrome gray scale images. A human silhouette region is extracted from the differential image between the human image and the background image. A Gaussian filter makes the image smooth. We used a threshold method to make a binary image. Opening and closing delete convex-concave details of the contour of the region. After that the height and width of the body and etc. are measured and the posture is discriminated.

2.1.1 Image Binarization

We explain how to extract a region of a human from an image taken by a camera. The purpose of this preprocess is to change a color image taken by a camera to a binary image to identify a region of a human.

We calculate a mean value of RGB component of an image and make gray scale image i.e. monochrome image.

Before taking an image with a human we took an image without a human and use such an image as a background image.

We make a differential image between an image with a human and the background image.

We used Gaussian filtering to the differential image to remove noise and make a smooth image. Gaussian filter is one of weighted mean filters which has heavy weight values near its center and light weight values far from its center. We used a 5 x 5 size Gaussian filter

Threshold method gives us a binary image from a gray scale image. If absolute values of the pixel of Gaussian image from the differential image are larger than a threshold value we assume the pixel is included in a human region. The threshold value is 20% of the maximum (i.e. 51 ≒ 0.2 x 256 in 256 gray level image) on an empirical basis.

2.1.2 Remove Noise

We used opening method and closing method to remove convexo-concave details of the contour of the human region. In opening method, if there is one pixel or there are more positive (i.e. included in a human region) pixels in the neighbor of a current pixel we make the current pixel positive. In closing method, if there is one pixel or there are more negative (i.e. out of a human region) pixels in the neighbor of a current pixel we make the current pixel negative.

Fig.1(a) shows the raw data taken from camera. Fig.1(b) shows a binary image which is a result of the preprocess.

2.2 Parameters from Images

We discriminate a posture of a human such as “standing”, “sitting ”, or “lying” with a height and a width of a human region from an image after preprocess. We raster-scaned an image like Fig.1(b) from left-top to right-bottom and decide the top border and the left border of a human region. Then we raster-scaned the same image inversely to decide the bottom border, the right border, the width and the height. We call one’s body length as a height.

(a)Before (b)After

Fig.1 Preprocess for Images

2.3 Process for Acceleration Data

We separate acceleration data from accelerometers with a certain interval. We calculate the mean value, the angle of the body (using the gravity direction), the deviation, the maximum value and power spectrum of Fourier transform. Then we decide the posture and motion of the human body in the time period.

2.3.1 Mean Value and Angle of Body

We decide the posture and motion of the human body in the time period from an acceleration data along three aces *x*,*y* and *z*. We decide the direction of the body according a formula shown below.

$$Angle = \arcsin(Average) \qquad (1)$$

Where *Average* is a mean acceleration value along *z*-axis (*G*). If a person stands up strait the value becomes about 1.0.

2.3.2 Deviation

We calculate the deviation of the acceleration data to estimate variability of the accelerations. If the deviation is larger than a threshold value we decide that the person is in action. But smaller we decide in still.

2.3.2 Synchronization

We wanted to synchronize the image data and the acceleration data. We requested a volunteer actor to jump at the start and the end of action. We detect the jumps with local maximum peaks in acceleration data also detect local maximum peaks of differential of positions of centroids of human regions in the binary images.

2.3.4 Power Spectrum

We transformed the time sequence of the acceleration along *z*-axis during a certain time interval and calculated the heights of the peaks to discriminate “walking” or “running”. Heights of peaks in walking data are low but those in running data are high. We assumed that method is suit for discrimination of “walking” and “running”.

2.3.4.1 Fourier Transform

Fourier transform is an expression of amplitudes of a frequency components of waves. In this case we assume that a time sequence of the acceleration is constructed of several waves.
Discrete data are transformed by discrete Fourier transform. The real part of the transformed result is represented as a_n. And the imaginary part is represented as b_n.

$$a_n = \sum_{n=0}^{N-1} x(n)\cos(\frac{2\pi}{N}kn) \qquad (2)$$

$$b_n = \sum_{n=0}^{N-1} -x(n)\sin(\frac{2\pi}{N}kn) \qquad (3)$$

The power spectrum pw is the root square of the sum of square of the real part a_n and the square of imaginary part b_n.

$$pw = a_n{}^2 + b_n{}^2 \qquad (4)$$

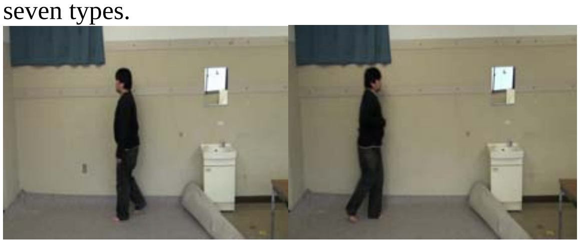
2.4 Analysis using Image and Acceleration

We synchronized images and accelerations and analyze postures and behaviors. We detected a start jump timing in image sequences. Then also we detected a start jump timing as a peak in acceleration time sequences. Also we detected end jumps to synchronize.
The steps of the proposed method are shown below.

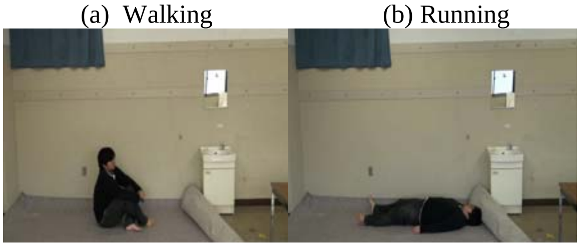
Step1 : Preprocessing Images
Step2 : Synchronize images and acceleration
Step3 : Detect top, bottom, left and right borders of human regions in images.
Step4 : Mean of acceleration
Step5 : Deviation of acceleration
Step6 : Peaks of power spectrum
Step7 : Discrimination of postures and behaviors according features of images and accelerations

3. Experiment

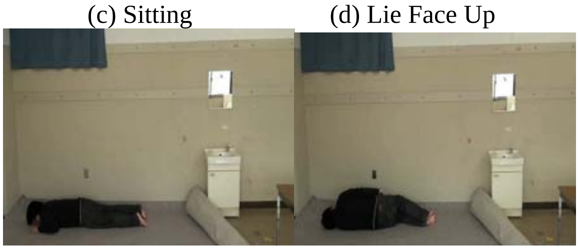
10 test subjects wore accelerometers on their bellies and backs each. A video camera took time sequence of their images. There are 7 types of postures or behaviors, “walking”, “running”, “lie face down” , “lie face up”, “lie left down” and “lie right down”. Fig.2 shows images the seven types.



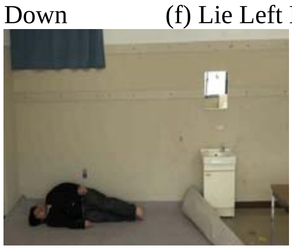
(a) Walking (b) Running



(c) Sitting (d) Lie Face Up



(e) Lie Face Down (f) Lie Left Down



(g) Lie Right Down

Fig.2 Postures and Behaviors

Table 1: Threshold

thresholds about accerelation	**Threshold**
The body angle to the ground	15°
The body angle on belly-back	Under −75°or over 75°
A4 Paper The body angle on sideway in rolling	Under −70°or over 70°
variance	0.1 (ratio vs mean^2)
power spectrum	500
Threshold about image	**Threshold**
The height of human region in standing	Over 65% of one's body length
The height of human region in sitting	45%～55% of one's body length
The height of human region in lying	Under 40% of one's body length
The width of human region in standing	Under 45% of one's body length
The width of human region in sitting	40%～55% of one's body length
The width of human region in lying	Over 60% of one's body length

3.1 Camera and Accelerometer

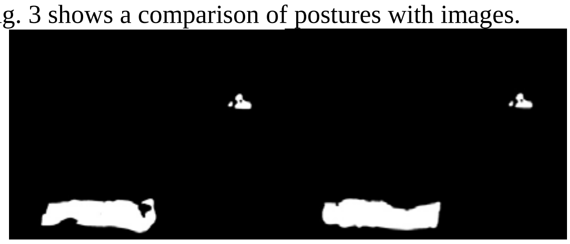
We used HDR-CX550V a hard disk handy video camera made by SONY to take videos of test subjects. The video is captured to images with 3 [fps].
Our accelerometer is WAA-006 a compact wireless hybrid sensor made by ATR-promotion. A WAA-006 has 3-axes acceleration sensor 3-axes gyro sensor and measures acceleration data and angle velocity data and send them to a PC by Bluetooth. The range of acceleration is ± 4 [G]. The sampling interval is 3[ms]. The size is W 39[mm] x H 44[mm] x D12[mm]. The weight is 12[g]. In Bluetooth communication a human body blocks a radio wave. So we used 2 accelerometers on a human belly and a human back.

3.2 Threshold for Discrimination

The image acceleration data set is discriminated with threshold values. We employed trial and error method to decide threshold values. Table 1 shows the threshold values.

4 Experimental Result and Consideration

Fig. 3 shows a comparison of postures with images.



(a) Lie Face Up (b) Lie Face Down

(c) Bending

Fig.3 Comparison of Postures with Images

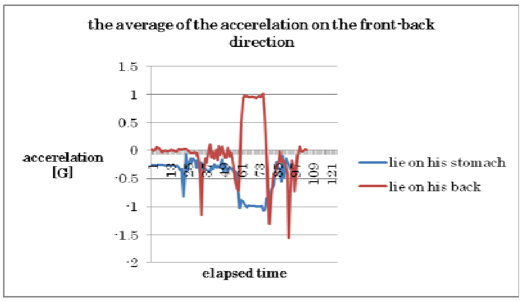
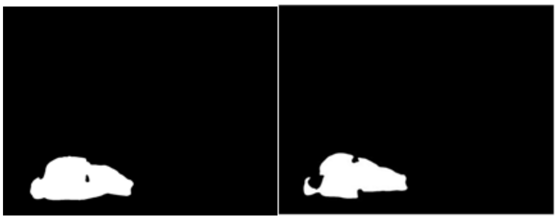


Fig.4 Moving Average of Antero-posterio Acceleration

Fig.3(a) and Fig.3(b) are not discriminated to each other only with images. But they are discriminated to each with acceleration data. Fig4 shows acceleration data. A blue line shows acceleration data along forward-backward axis in a posture of lie face down. And a red line shows acceleration data along forward-backward axis in a posture of lie face up.
Fig.3(a) and Fig.3(c) are not discriminated to each with acceleration data because both acceleration data are

similar to the blue line in Fig.4. But they are discriminated with those images.



(a) Lie Left Down (b) Lie Right Down

Fig.5 Binary Images of Lying along Lateral Axis

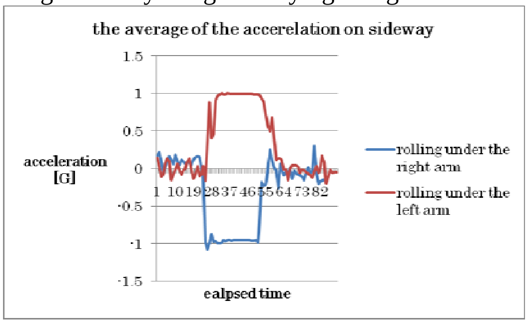


Fig.6 Acceleration along Lateral Axis

Fig.5 shows images of lying along the lateral axis, i.e. lying left down or lying right down. But it is hard to discriminate them only with images. Fig.6 shows acceleration data along the lateral axis, i.e. on sideway. The acceleration in lie left down is very different from lie right down. The two cases are discriminated with acceleration data.

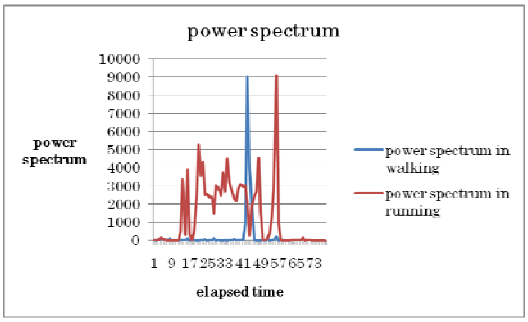


Fig.7 Comparison of Power Spectra of Walking and Running

Fig.7 compares a power spectrum of walking and that of running along one's body axis. If the frequency of the maximum peak is high and it continues we decide that the person is running. But if the frequency is low we decide walking.

Table.2 Accuracy of Result of Discrimination

posture・movement	Result
Waking	$\frac{99}{103}$ (96.1%)
Running	$\frac{101}{105}$ (96.2%)
Lie face up	$\frac{61}{67}$ (91.0%)
Lie face down	$\frac{64}{70}$ (91.4%)
Lie left down	$\frac{86}{93}$ (92.5%)
Lie right down	$\frac{72}{83}$ (86.7%)
Sitting	$\frac{66}{71}$ (93.0%)

Table 2 shows accuracy of result of discrimination of postures and behaviors. For example there were 103 walking data. The proposed method discriminated 99 data as walking data among them. Lacks of data through Bluetooth and threshold method cause errors. But total accuracy is over 85%.

5. Conclusion

We proposed a security and observation System for a solitary person using accelerometers and a surveillance camera. We showed that the proposed method to discriminate human postures and behaviors with camera images and acceleration data is very efficient. The method achieved discriminations which were impossible with only images or impossible with only acceleration data. In most every case two acceleration sensors could bring us acceleration data in spite a human body shielding a radio wave of Bluetooth. In future work we will reduce the ratio of lost Bluetooth data, will apply the method for more test data and will improve the discrimination method to increase the total accuracy.

Acknowledgments

This work is partially supported by SCOPE: Strategic Information and Communications R&D Promotion Program (Project ID : 10103768).

Thanks for the students of the School of Engineering and Grad. School of Engineering of Osaka City Univ. to act as our test subjects.

References

[1] Ministry of Health, Labour and Welfare JAPAN, "Vital Statistics 2005," Ministry of Health, Labour and Welfare JAPAN, Tokyo JAPAN,2005.

[2] C. Sun, D. Stirling, F. Naghtdy, "Human behaviour recognition with segmented inertial data," Proceedings of the 2006 Australasian Conference on Robotics & Automation, Auckland New Zealand, December 6-8.2006.

[3] Y. Kawahara, T. Hayashi, H. Tamura, H. Morikawa, and T. Aoyama, "A context-aware content delivery service using off-the-shelf sensors," In Proceedings of The Second International Conference on Mobile Systems, Applications, and Services e (Mobisys 2004), (Poster Presentation), Boston, U.S.A., June 2004.

[4] N. Ravi, N. Dandekar, P. Mysore, M.L. Littma,"Activity recognition from accelerometer data," Conference on Innovative Applications of Artificial Intelligence, pp.1541-1546,2005.

[5] A. Gallagher,Y. Matsuoka,W.T. Ang,"An efficient real-time human posture tracking algorithm using low-cost inertial and magnetic sensors," Proceedings of 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems, Sendai, Japan, September 28-October 2,2004.

[6]

[7] S. S. Intille, L. Bao, E. M. Tapia,and J. Rondoni, "Acquiring in situ training data for context-aware ubiquitous computing applications," in Proceedings of CHI 2004. Connect: Conference on Human Factors in Computing Systems. New York, NY: ACM Press, pp. 1-9, New York, USA, 2004.

[8] N. Kern, S. Antifakos, B. Schiele, and A. Schwaninger, "A model for human interruptability : experimental evaluation and automatic estimation from wearable Sensors," Proceeding of 8th IEEE International Symposium on Wearable Computers(ISWC'04), pp.158-165, Arlington, USA, October 2004.

[9] S. Aoki, M. Onishi, A. Kojima, Y. Sugahara, K. Fukunaga, "Recognition of a Solitude Senior's Behavioral Pattern Using Infrared Detector ,"Technical Report of IEICE (Institute of Electronics, Information and Communication Engineers) WIT (Well-being Information Technology), Vol.101, No.703, pp. 43-48, March 2002.

[10] Y.Ogoshi, S.Ogoshi, S.Hirose, H.Kimura,"A Study on the Recognition of Human Activities by Datamining from Infrared Sensor Information," IEICE Trans. on Information and Systems , J85-D-II(5), pp.959-964, May 2002.

Satoshi Hasebe received the B.E. degree in Information Engineering from Osaka City University in 2011. His research interests include analysis of human postures and behaviors.

Shigeyoshi Nakajima received the B. E. and M. E. degree in Electric Engineering from Kyoto University, Kyoto, Japan in 1982 and 1984 respectively. He received the Ph.D. degree in Information Engineering from Osaka City University, Osaka, Japan in 1997. He is now an Associate Professor in Osaka City University, Japan. His research interests include signal processing, image processing, medical engineering and optimization algorithm and so on. He is a member of IEEE, IEICE and IPSJ.

Thi Thi Zin received the B.Sc (Hons.) degree in Mathematics in 1995 from Yangon University, Yangon, Myanmar and the M.I.Sc degree in Computational Mathematics in 1999 from University of Computer Studies, Yangon, Myanmar. She received the M.E. and the Ph.D. degrees in Information Engineering from Osaka City University, Osaka, Japan, in 2004 and 2007, respectively. She is now a Project Research Associate, in Osaka City University, Japan. Her research interests include ITS, Color Image Processing, Mathematical Morphology and so on. She is a member of IEEE.

Takashi Toriu received the B.Sc. in 1975, M.Sc. and Ph.D. degree in physics from Kyoto University, Kyoto, Japan, in 1977 and 1980, respectively. He was a researcher in Fujitsu Laboratories Ltd. from 1982 to 2002, and now he is a Professor of Osaka City University. His research interests are in the areas of image processing, computer vision, and especially in modeling of human visual attention. He is a member of IEEE, IEICE, IPSJ, ITE and IEEJ.