Activity Recognition for Risk Management with Installed Sensor in Smart and Cell Phone

Daisuke Honda, Nobuchika Sakata, and Shogo Nishida

Division of Systems Science and Applied Informatics Graduate School of Engineering Science, Osaka University 1-3, Machikaneyama-cho, Toyonaka-city, Osaka, 560-8531 Japan
{suzuki,sakata,nishida}@sys.es.osaka-u.ac.jp

Abstract. Smart and cell phone with self-contained sensor such as accelerometer, gyroscopic and digital magnetic compass sensor have been popular. Combining certain algorithm and those sensors, it can estimate user's activity, situation and even user's absolute position. However, estimation of user's activity, situation and user's absolute position become difficult when once sensors posture and position are changing from original position in user's motion. Also, according to stored, worn and handheld position and posture of those cell and smart phone are often changed. Therefore, we exclude estimation of user's position and we focus to only estimation of user's activity and situation for risk management. Basically, we design special classifier for detecting user's unusual behavior and apply other user's position data from internet to the results detected by the classifier which are combined wavelet transform and SVM. We assume that user's unusual activity and situation can be detected by smart and cell phone with high accuracy.

Keywords: Activity recognition, Wearable computer, SVM, nearly fall incident, cell phone and smart phone.

1 Introduction

In 1929, Heinrich a safety engineer in U.S. found the principle that there are 29 small accidents and 300 incidents behind 1 fatal accident. This principle tells us the way to prevent fatal accidents. To prevent fatal accidents, we should not overlook small accidents and incidents and should take countermeasure based on it. Aviation industry in U.S. founded ASRS in 1976 and started incident report system. Learning from mistake is more profitable than punishment to mistake they thought. Also they have achieved a certain result. Currently, the importance of incident report is acknowledged in all over the world. There are a lot of organizations conducting incident report system, such as IMO, MAIB, MARS, IMISS and TSB.

Community involving high-risk job specialist such as aviation industry, medical field and marine tend to be adopted incident report activity. On the other hands, in community involving low-risk, it is difficult to take such kind of incident report activity. People cannot keep motivation and perceive incentive to report their incident experiences. Finally, such incident report system is adopted only after a fatal accident happens.

J.A. Jacko (Ed.): Human-Computer Interaction, Part III, HCII 2011, LNCS 6763, pp. 230–239, 2011.

For this problem, automatic accumulation system of incident information has been tried in motor vehicle industry. InterNavi[1] developed by HONDA is actual system of road maintenance and improvement using information of sudden braking cases accumulated automatically from general public drivers. Conducting road maintenances based on danger place and point obtained from accumulated data involving sudden brakes, and then, the number of sudden braking was reduced 70%. Currently, such activity is not tried to pedestrian. Therefore, we assume that there is effect to reduce the fatal incidents with applying automatic incident accumulation to pedestrian. In order to apply automatic incident accumulation system to pedestrian, we consider major two problems not to be found in motor vehicle. First, there are a lot of flexibilities in human movement when comparing to motor vehicles. There is no obvious key to detect incident occurrence, such as sudden braking in motor vehicle. There are many researches for studying detection of complete fall accidents [2],[3],[4]. The detection key of complete fall accidents is clearer than that of nearly fall incident. Second, there is a practical problem to deploy sensors on pedestrian's body.

2 Policy of Sensor Deployment

Motor vehicle has been computerized drastically. There are a lot of ICs and sensors everywhere in a modern car. In case that we can deploy ICs and sensors on human body wherever we like, human activities can be recognized automatically to some extent. However, it is difficult to deploy sensors on human body wherever we like. In case of guarantying a lot of merits directly for the person to wear computers and sensors, pedestrians may wear computers and sensors. For example, smart computer realize context awareness [5], learn application for commuter adapting to user context [6], and monitor health care of elderly people [7] have been reported. In such applications, single purpose sensors are deployed at upper arms, thighs, shoes and cane wherever and whatever developer like. However, there is no merit for each pedestrian directly in the application we propose. The merit of our proposed application is not for individual but for community. Without profiting individually merit, people may not accept burden of wearing special sensors. In addition, incidents of pedestrian are not occurs frequently. Thus single purpose device is not worth for our proposed application in terms of time cost, wearable burden and bothers.

Considering constraint above-mentioned, the most practical way is to use mobile phone. Almost of all pedestrian carry mobile phone. Also smart phones equipped with inertial sensor are quite popular such as iPhone. We assume that more and more people may have it in the future. With using installed sensors in mobile phone, we can induce many pedestrian to accept automatic incident report system with no concern of user burden and cost of special device.

When we use the sensor in mobile phone, we must consider the stored position of mobile phone on human body. iShare Corporation reported one of the result of awareness survey [8] about mobile phone. According to this survey, stored position of mobile phone varies from person to person, and also varies from day to day in same person depending on attire or activity. In case building applications of human activity recognition with installed sensors in mobile phone, recognition algorithms should be changed according to the stored positions of mobile phone. Basically, device contexts should be recognized first. After that user context can be recognized. We assume that Two-phased approach is needed.

Fujinami et al. conducted a research about device context of mobile phone with tri-axial acceleration sensor [9]. According to their research, stored positions of mobile phone can be recognized with appropriate accuracy using sensor signals of stationary movement period like walking. Therefore, we suppose that the stored position of mobile phone is obtained from their method. In this paper, we only focus a method to detect nearly fall incidents occurrence during walking.

Applications for individual user need perfectible recognition of user activities even in any stored position and posture of mobile phone. However, our proposed application is for community. We assume that the automatic incident reports system gather a lot of data from many people carrying with cell and smart phone. Thus, our proposed application does not chase individual activities perfectly. The application can ignore difficult device context with a lot of flexibility such as handheld situation or holding in bag. Finally, we plan to activate incident detection application only when mobile phone is in bottom of user's pocket. Bottom of user's pocket can be expected not to have considerable individual difference in shape. At the same time, bottom of user's pocket is close to center of gravity of human body. It can be expected that discriminative signal is observed when the person lose balance by trip and slip. In addition, it is possible to apply to one-third of pedestrians even if we ignore the position other than bottoms of user's pocket from ishare report result [8]. In the future, there is possibility to extend the scope of application to other stored position of mobile phone. However, in this paper, we believe that it is enough for initial approach.

3 Detection Method

In this research, we use Wii and Wii motion plus instead of smart phone equipped with sensors to measure human activity. It contains a tri-axial acceleration sensor and a tri-axial gyro sensor. We configure sampling frequency at 16 Hz. It is determined consulting with the research [10], [11]. Also we suppose that device context is in bottoms of user’s pocket (Fig. 1). There are three chief functions.

Fig. 1. Appearance of sensor position. We plan to activate incident detection application only when mobile phone is in bottom of user's pocket. Bottom of user's pocket can be expected not to have considerable individual difference in shape.

1. Transformation into coordinate independent of sensor attitude.
2. Division of time-series data and feature extraction.
3. Learning and recognition by SVM.

3.1 Transformation into Coordinate Independent of Sensor Attitude

We suppose that mobile phone is in bottoms of user's pocket. Even so, there is flexibility of sensor attitude. We can simplify and exclude the flexibility by the following steps. In order to exclude the flexibility of sensor attitude, we transform the original sensor data into the data according to coordinate on pedestrian. (Fig. 2) The three axis are consist of

1. Direction of gravitational force,
2. Direction of pedestrian's movement
3. Side direction of pedestrian.

Transformation of coordinate is conducted during a stationary period of walking. Mobile phone in the bottoms of user's pocket is in certain restraint. It is unthinkable that mobile phone spin round and round in the bottoms of user's pocket. Small rotary motion may be occurred along to walking movement, but it can be expected like the swing of pendulum. Also device may keep stable attitude in average viewpoint.

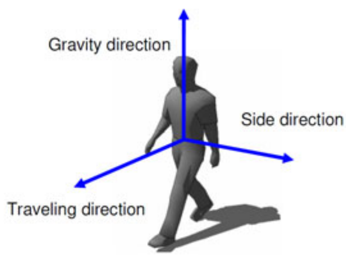

Fig. 2. Coordination on pedestrian. We transform the original sensor data into the data according to coordinate on pedestrian.

For this reason, we simply evaluate gravity acceleration vector by putting raw data of tri-axial acceleration sensor through low pass filter. We use third degree Butter worth low pass filter. Cutoff-frequency of the filter is configured at 0.5Hz. Signals under 0.5Hz is unlikely to be mixed in, no matter how slow the walking pace is. We can obtain nearly flat signal by this way. Also it is the approximation of gravity acceleration vector. (Fig. 3)

Next, we exclude element parallel to the gravity acceleration vector from the signals of tri axial acceleration sensor. We obtain the acceleration element existing in horizontal plane. In horizontal plane, swing along travelling direction is larger than that alongside direction. We can obtain the vector of travelling direction axis using principle component analysis (PCA). The vector of side direction axis can be obtained from pre mentioned two axis. By this way, we can obtain three axis installed on pedestrian's body. Along to the three axes, we can obtain acceleration and gyro signals independent of sensor attitude. (Fig. 4) The process above mentioned uses duration of stationary walking activity for 2 second. This process is based on the assumption that mobile phone attitude is remained steadily in average viewpoint. It cannot be used under the condition having more flexibility. However, it is enough for the condition of bottoms of user's pocket.

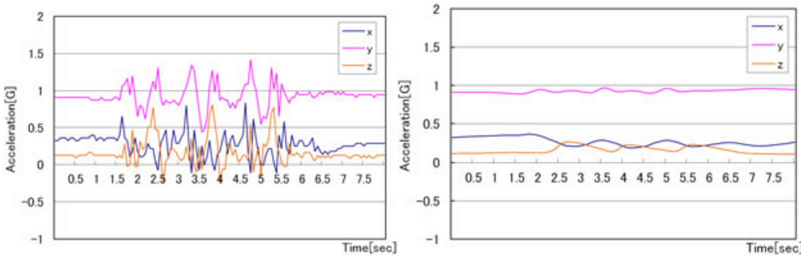

Fig. 3. We use third degree Butter worth low pass filter. Cutoff-frequency of the filter is configured at 0.5Hz. Left side figure is raw data from sensor. Right side figure is after applying the low pass filter.

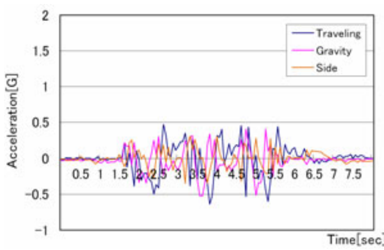

Fig. 4. The vector of side direction axis can be obtained from pre mentioned two axis. By this way, we can obtain three axis installed on pedestrian's body. This figure shows sensor data removed gravity.

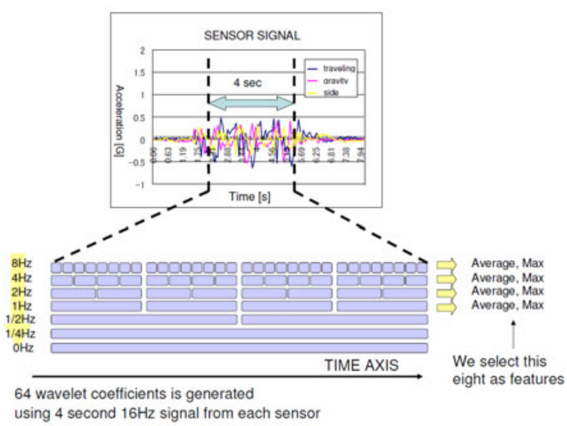

Fig. 5. We calculate maximums and averages of 1Hz, 2Hz, 4Hz and 8Hz coefficients individually. By this way, we obtain 8 features for each sensor. In this phase, time flexibility disappears.

3.2 Division of Time-Series Data and Feature Extraction

Pattern recognitions of time-series data require the process of dividing the data and feature extraction from the divided data. This time, we divide sensor data at the interval of four second. There are six sensors. Sampling frequency is 16 Hz. Thus, one compartment consist of $4 \times 6 \times 16 = 384$ data. In case of that we use this 386 data as a feature vector, there remains flexibility of time axis, and the flexibility make pattern recognition process complex. In order to exclude flexibility, feature extraction is generally conducted before pattern recognition process. In speech recognition, transformation into frequency domain is commonly used such as Fourier transformation. However, signal of nearly fall incident is not a periodic signal. It is sudden signal like single peaked pattern. We assume Fourier transformation that intend for periodic signal is not suitable, and conduct feature extraction by Wavelet transformation which is strong for the analysis of sudden signals. In this time, we divide signals simply into Haar wavelet. We apply wavelet transformation to each sensor data series. Applying wavelet transformation to 64 data (16Hz $\times$ 4sec), we obtain the same number of wavelet coefficients which consist of one coefficient corresponding to 0Hz, one coefficient corresponding to 1/4Hz, two coefficients corresponding to 1/2Hz, four coefficients corresponding to 1Hz, eight coefficients corresponding to 2Hz, sixteen coefficients corresponding to 4Hz, thirty two coefficients corresponding to 8Hz. Using wavelet transformation, there remain time flexibilities in this phase. We calculate maximums and averages of 1Hz, 2Hz, 4Hz and 8Hz coefficients individually. (Fig. 5) By this way, we obtain 8 features for each sensor. In this phase, time flexibility disappears. There are six sensors. So now we obtain 48features. We add variances of each sensor data and finally obtain 54 features in total.

At the last step, we normalize 54 features individually. For this normalization, we calculate averages and variances of each feature from data of various activities of various people.

3.3 Learning and Recognition by SVM

SVM is often-used algorithm of pattern recognition because of its high performance in experimentation, existence of theoretical basis, existence of simple implementation, and capability of nonlinear recognition using Kernel method [12].

We use the SVM algorithm written in Reference [13]. We use Gaussian Kernel. The parameter variance of this Kernel is configured at 2000. Another parameter of soft margin is configured at 1000.

4 Experiment

Our research objective is detecting nearly fall incidents of pedestrian in outdoor situation of regular daily activities. In other word, we suppose to distinguish nearly fall incidents from daily activities. The method described above is constructed for this purpose. Currently, we must evaluate classification ability of the method. In order to evaluate the method, we need to collect samples of nearly fall incident and regular daily activity. Collecting real incident data sufficiently in real life is very difficult, because incident do not occur frequently. Therefore we collect incident samples by imitation of subject. Also we define four activities as regular daily activities at outdoor. The four activities are ”WALK”,”RUN”,”GO UPSTAIRS”, and ”GO DOWNSTAIRS”. We let subject perform the four activities in small excursion. Then we define incidents as nearly fall incidents this time. The incidents can be caused by each side of foot, by right foot or left foot. By contrast mobile phone is held in only one side of the bottoms of user’s pocket. Thus, we define following two incidents; ”SENSORSIDE TRIP” and ”OTHER SIDE TRIP”.

Resultantly, we collect samples under six type conditions, ”WALK”,”RUN”,”GO UPSTAIRS”, and ”GO DOWNSTAIRS” as daily activity ”SENSOR SIDE TRIP” and ”OTHER SIDETRIP” as nearly fall incident.

We let twelve subjects (man, age 22-25) serve each activity ten times. As a result, we collect total 720 samples.

Fig. 6 shows the tendency of 54 features distribution in “SENSOR SIDE TRIP”. Unfortunately we cannot illustrate the 54-dimenal space. We compromise and illustrate the tendency of distribution by box-and-whisker plot.

As a pilot test, we divide 720 data into two group consist of 360 data. Both groups contain each activity of each subject in same number. First, we use one group as learning and conduct recognition test to other group. Then, we exchange the group role and conduct the same test and the result is shown in Fig. 7. There are 120 samples in each activity. Also figures in the table are number of successful samples of recognition. Frequent false positive is fatal obstacle to the incident. So smallness of false positives most important criterion we consider. It is 94.2%. To achieve it near 100%, additional improvement is needed. Evaluating it in often used criterion, the result is 88.0%Precision, and 85.8% Recall.

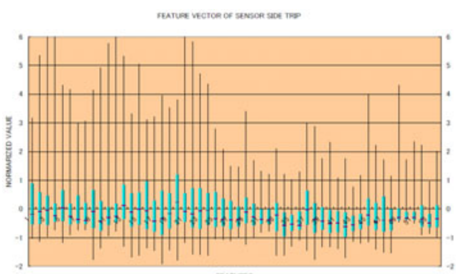

Fig. 6. Figure shows the tendency of 54 features distribution in “SENSOR SIDE TRIP”

USING 1-54 (ALL FEATURES)

	success	%		success	%
WALK	116	96.7	*DAILY*	452	94.2
RUN	116	96.7			
GO DOWNSTAIRS	107	89.2			
GO UPSTAIRS	113	94.2			
SENSOR SIDE TRIP	98	81.7	*INCIDENT*	206	85.8
OTHER SIDE TRIP	108	90.0			

Fig. 7. We use one group as learning and conduct recognition test to other group. Then, we exchange the group role and conduct the same test and the result.

5 Discussion

Using 54 features, we attain 85.8% recall with false positive 5.8%. And precision is 88.0%. Actually this precision cannot be applied to real life data. Generally speaking, nearly fall incident is slightly rare. Even if false positive rate is quite low level, false detection can be accumulated in long daily activities. Consequently, precision is decreased in real life situation. Therefore, more effort and improvement is needed.

On the other hand, automatic incident report system does not need complete accuracy. It is enough to discover the candidate sites having a potential danger. With applying the application to sufficient number of pedestrian, the number of detected incidents at dangerous site may be expected to grow larger comparing with the number of false detections at not dangerous site.

Table 1. The result of recognition with smaller size feature vector

USING FEATURES	PRECISION %	RECALL %
ALL FEATURES	88.0	85.8
ALL WAVELET FEATURES	87.2	85.0
VARIANT OF RAW SIGNAL FEATURES	36.1	88.3
8HZ WAVELET FEATURES	84.0	67.9
4HZ WAVELET FEATURES	86.3	68.3
2HZ WAVELET FEATURES	78.2	70.4
1HZ WAVELET FEATURES	76.6	54.6
MAX OF WAVELET FEATURES	63.9	53.8
AVERAGE OF WAVELET FEATURES	78.2	65.8
ACCELERATION FEATURES	81.6	66.7
GYRO FEATURES	85.3	75.0

Next, we conduct learning and recognition by SVM with smaller size feature vector to see contributions to the recognition of each feature. Table 1 show the result of recognition with smaller size feature vector. As shown in Table 1, we can realize the fact that VARIANT OF RAW SIGNALFEATURES does not contribute so much. The low precision in the line of VARIANTOF RAW SIGNAL FEATURES and high precision and recall in the line of ALLWAVELET FEATURES show the fact. Also you can see the fact that contribution of WAVELET FEATURES makes little difference by changes in frequency. Next you can see the fact that AVERAGE OF WAVELET FEATURES makes larger contribution compared to MAX OF WAVELET FEATURES. Previously, we use only MAXOF WAVELET FEATURES, regarding it as best to detect sharp peak due to tripping and losing balance. Precision is 63.9% and recall is 53.8%. Combining AVERAGE OFWAVELET FEATURES, precision improved 23.3% and recall improved 31.2%. More preferable selection may be existing in 64 wavelet coefficients corresponding 4 second of one sensor. GYRO FEATURES make larger contribution compared to ACCELERATIONS. Using acceleration sensor only, we obtain 81.6% precision and 66.7% recall. Combination use of gyro sensor improves the precision by 6.4% and the recall by 19.1%.For this reason, gyro sensor is essential to high accuracy. Recognition by gyro sensor only cannot work in our framework. Because of that sensor context recognition by Fujinamiet.al uses acceleration sensor, and detection of gravity direction also needs acceleration sensor.

In conclusion, all of 54 features are contributing to a greater or lesser extent. However VARIANT OF RAW SIGNAL FEATURES make less than 1% improve both in precision and recall. Other features are all essential to high precision and recall. For more improvement, incorporation of another sensor data or incorporation of some kind information or twist in feature extraction are hopeful.

6 Conclusion

In this research, we tried to implement nearly fall incident detection system, assuming that mobile phone is in bottom's pocket of pedestrian. We tested the system by imitated trip incident sample and daily activities. As a result, we obtained 88.0% precision and 85.8% recall (and False positive rate is 5.8%). We confirm the delectability of nearly fall incident by imitated samples at first step this time.

Acknowledgments. This research was partially supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Young Scientists (B) "Remote collaboration with Real-world based Interface", No.20700111.

References

1. http://www.honda.co.jp/internavi/ (in Japanese)
2. Lan, M., Nahapetian, A., Vahdatpour, A., Au, L., Kaiser, W., Sarrafzadh, M.: SmartFall:An Automatic Fall Detection System Based on Subsequence Matching for the SmartCane. In: BodyNets 2009 (2009)
3. Yoshimura, T., Nakajima, K., Tamura, T.: An ambulatory fall monitoring system for the elderly. In: Third International Conference of Gerontechnology, Munich, Germany, pp. 10–13 (October 1999)
4. Shinagawa, Y., Tanigawa, T., Ohta, S.: Detection of Human Gait and Falling Using an Accelerometer. Kawasaki Journal of Medical Welfare 9(2), 243–250 (1999) (in Japanese)
5. Takada, D., Ogawa, T., Kiyokawa, K., Takemura, H.: A Context-Aware Wearable AR System with Dynamic Information Detail Control based on Body Motion. Journal of Human Interface Society 12(1), 47–56 (2010) (in Japanese)
6. Nakamura, T., Ogawa, T., Kiyokawa, K., Takemura, H.: Implementation and Evaluation of a User Context Recognition Mechanism for Continuous Wearable Learning Systems. Technical Report of IEICE 106(396), 17–22 (2006)
7. Lan, M., Nahapetian, A., Vahdatpour, A., Au, L., Kaiser, W., Sarrafzadh, M.: SmartFall:An Automatic Fall Detection System Based on Subsequence Matching for the SmartCane. In: BodyNets 2009 (2009)
8. iShare Inc., Survey on mobile phone usage (November 2005), http://bizmakoto.jp/bizmobile/articles/0511/04/news105.html (in Japanese)
9. Fujinami, K., Jin, C., Kouchi, S.: Tracking On-body Location of a Mobile Phone, Late Breaking Results - Cutting Edge Technologies on Wearable Computing. In: The 14th Annual IEEE International Symposium on Wearable Computers, ISWC 2010 (October2010)
10. Bouten, C.V.C., Koekkoek, K.T.M., Verduin, M., Kodde, R., Janssen, J.D.: A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity

11. Tabuchi, K., Naya, F., Ohmura, R., Noma, H., Kiyoshi, K., Kishino, F.: Evaluation of data collection parameters on the daily activity classification with accelerometers. IEICE Technical Report PRMU, 43–48 (May 19,2006) (in Japanese)
12. Tsuda, K.: Overview of Support Vector Machine. Technical Survey IEICE 83(6), 460–466 (2000) (in Japanese)
13. Introduction to Support Vector Machine (in Japanese) ISBN-13: 978-4320121348