

A Watch-type Human Activity Detector for the Aged Care

Chankyu Park*, Jaehong Kim*, Ho-Jin Choi**

*Dept. of robot & recognition research, ETRI(Electronics and Telecommunications Research Institute), Daejeon, Korea

**Dept. of Computer Science, KAIST(Korea Advanced Institute of Science and Technology), Daejeon, Korea

parkck@etri.re.kr, jhkim504@etri.re.kr, hojinc@kaist.ac.kr

Abstract—Human activity recognition is widely researched in the various filed these days. For the aged care, the one of the most important activities of old people is fall, since it causes often serious physical and psychological results. Many researchers have studied human activity recognition techniques in various domains; however none released to a commercial product satisfying the old people requirements, which are comfortable to wear it, weight-lighted and having exact accuracy to detect emergency activity and longer battery durance. Thus, to address them, we propose a systematic evaluating procedure for getting best minimum feature sets and best classification accuracy. We also do experiments for comparing the two features reduction techniques and four classification techniques in order to discriminate four each basic human activities, such as fall for the aged care, walking, hand related shocks and lastly general activity which is just last all arbitrary behaviour except existed three other patterns.

Keywords—Elderly care, wearable device, human activity recognition, fall detection, accelerometers, feature selection, feature reduction

I. INTRODUCTION

For the aged care, it is very important to detect emergency activities from ordinary daily life activities of old people. Falls are one of the most closed to emergency state in old people's daily life at nursing home. Many researchers have studied human activity detection techniques in elderly care domain. With the enhancement of wireless networks and low-power MCU (Microcontroller Unit) technology, wearable fall detectors using inertial based sensors have been released and are capable of sending an alarm to caregivers or medical center automatically. These devices are small, inexpensive, easy to use, and can be worn at various body locations such as waist, chest, or back and have more merit than the vision based devices, in terms of privacy and blind spot problems. The elderly people may well accept a watch-type device as a type of fall and activity detector in many situations, since most people have no aversion about wearing a watch.

On the other hand, most of classification techniques generally require lots of computation, because they require many data sets having high dimension of a feature vector to represent characteristics of human activity based on 3-axis accelerometer sensor. In addition to make sure real-time detection of the old people's basic activities in watch-type low-power devices, it need to be simple and less computation, because a typical low-power MCU has low clock speed while it consumes less battery. Thus, one of the candidates to reduce computation is to reduce dimension of a feature vector used as input of a classifier, but it also must not decrease the accuracy to detect various basic human activities.

In this paper, we explore the feasibility of using various feature reduction and classifiers to detect fall and other activities; walking, hand shock and general activities. Then propose the efficient method to reduce a lot of computation in order to ensure real-time execution in watch-type device having low-power MCU.

II. PROPOSED APPROACH

We propose the various cases to recognize four kinds of human basic daily activities including falls as one of emergency activities and also reduce a lot of computation to ensure real-time detection in wrist-type device having low-power MCU.

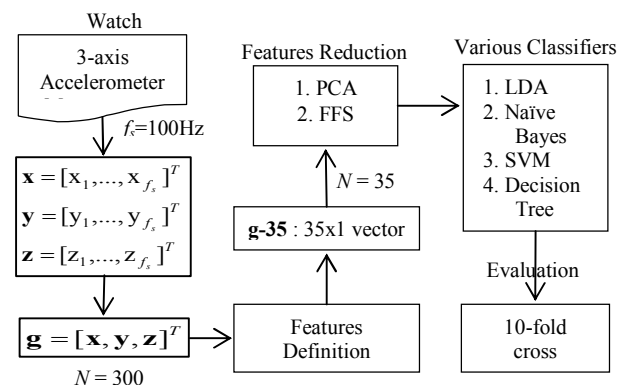


Figure 1. Overall our proposed approach to evaluate feature reduction and various classification techniques

Figure 1 introduces a systematic approach which is composed of three major procedures; (1) signal acquisition from the watch in Figure 2, pre-processing 3-axis accelerometer signal and feature definition for extracting major features to represent the characteristics of recorded activities; (2) feature reduction and classifier modelling for preventing over fitting and complexity; (3) evaluation of performance measure for comparative criterion. From the following sections, we present detail description and processes of each procedure.

A. Data Acquisition and Definition of Activity Patterns

We did experiments to collect data sets of fall and ordinary human basic activities elicited by young subjects wearing the watch for defining features and learning parameters of various machine learning algorithms.



Figure 2. A watch-type human activity detector including a 3-axis accelerometer

Figure 2 shows watch-type activity detector that includes a 3-axis accelerometer and micro-SD memory to store its data. In this study, 15 activity patterns were chosen to estimate parameters of classifiers to be more robust against false alarms, and simulated trials which are trying to play similar fall actions and other ordinary activities were performed by 21 volunteers (ages range: 20 ages-5 people, 30 ages-9 people, 40 ages-6 people, 50 ages-1 people). The chosen lists of various activity patterns we adopted are categorized by characteristics of activity as shown in Table 1.

TABLE 1. THE FOUR KINDS OF HUMAN ACTIVITIES FOR CLASSIFICATION

The four human activity patterns for the aged people			
1. Fall	2. Walking	3. Shock	4. General Activity
Fall-Forward	Normal walking	Clapping hand	Standing
Fall-Back	Up stairs	Twisting hand	Sitting
Fall-Left	Down stairs	Trembling hand	Lying
Fall-Right		Hitting hand	Steady State

We grouped 15 patterns into four classes such as fall, walking, shock and general activity and then labelled each record into training data sets as four classes for multi-class classification. Consequently, we acquired 2,800 training sets according to four activities.

B. Feature Definition from Accelerometer's Signal

In this study, we use features derived from the watch's accelerometer with $\pm 6g$ range. Many researchers reported that falls mostly happened within 0.8 seconds [6,9]. To satisfy this condition for detection of emergency states, we used the sliding window corresponds to a second for checking a fall event at sampling rate of 100 Hz without overlapping.

For instance, in the case of walking pattern, waveforms of signals are not changed abruptly and amplitude of signals is generally restricted to the level within $-2g \sim 2g$. To differ fall patterns from similar shocks occurred by hand, we gathered various hand patterns. These are used to train classifiers to be more reliable for decreasing false positive errors.

For examining the influence of different size of dimensions and characteristic of features, we take account of all possible candidates of features as many as possible and then we can try feature analysis like feature selection and transformation techniques. We define a feature vector g_{35} [35x1] described in Figure 1 and Table 2 and we consider the magnitude of the mean, variance, energy, maximum, minimum, the pairwise correlation of the acceleration in x, y and z direction as well as the magnitude of the norm mean, norm variance, norm maximum, norm minimum, so that a g_{35} feature vector was created.

TABLE 2. EACH ELEMENT DESCRIPTION OF g_{35} VECTOR

Feature Vector $g_{35} = [g_1, \dots, g_N]^T$ $N = 35$
$[g_1, g_2, g_3]^T = [\hat{\mu}_x, \hat{\mu}_y, \hat{\mu}_z]^T$ with $\hat{\mu}$ is sample mean
$[g_4, g_5, g_6]^T = [\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z]^T$ with $\hat{\sigma}$ is sample variance
$[g_7, g_8]^T = [\hat{\mu}_m, \hat{\sigma}_m]^T$ for $m = [m_1, \dots, m_{f_s}]^T$ $m_k = \sqrt{x_k^2 + y_k^2 + z_k^2}$, $1 \leq k \leq f_s$
$[g_9, g_{10}, g_{11}]^T = [\max_x, \max_y, \max_z]^T$ for $\max_v = \max_k \{v\}$, $1 \leq k \leq f_s$
$[g_{12}, \dots, g_{15}]^T = [\min_m, \min_x, \min_y, \min_z]^T$ for $\min_v = \min_k \{v\}$, $1 \leq k \leq f_s$
$[g_{16}, \dots, g_{27}]^T = [\xi_{1,2}, \xi_{1,3}, \xi_{2,3}, \xi_{4,5}, \xi_{4,6}, \xi_{5,6}, \xi_{9,10}, \xi_{9,11}, \xi_{10,11}, \xi_{13,14}, \xi_{13,15}, \xi_{14,15}]^T$ with $\xi_{i,j} = \frac{g_i \cdot g_j}{g_i^2 + g_j^2}$
$[g_{28}, g_{29}, g_{30}]^T = [\text{median}_x, \text{median}_y, \text{median}_z]^T$ for $\text{median}_v = \text{median}(v)$
$[g_{31}, g_{32}]^T = [\sqrt{\sum \text{diag}(\text{cov}(x, y))}, l_m]^T$ with $l_m = n(\{m_k \mid m_k > 3.0\}) / f_s$, $1 \leq k \leq f_s$
$[g_{33}, g_{34}, g_{35}]^T = [\hat{\mu}_m^{t-1}, \hat{\sigma}_m^{t-1}, \max_m^{t-1}]^T$ with $t-1$ = index of the pervious window frame

C. Feature Reduction and Activity Classification

Reducing the dimensionality of features is important to parameter estimation of classifiers in machine learning. In some cases, when the dimension N of a feature vector g can be very high, many elements of g may be redundant or even

irrelevant in terms of the classification. We have to be cautious to resolve the dimension of a feature vector according to two major reasons: One is that the computational complexity comes to be too large. The other is that increasing the dimension ultimately results in over-fitted on the training data. To consider those problems, one can apply two main approaches to reduce features: feature selection and feature transformation. Feature selection algorithms select a subset of features from the original feature set; feature transformation methods transform data from the original high-dimensional feature space to a new space with reduced dimension. Feature transformation methods use dimension reduction algorithms to be achieved by computing an optimal subset of predictive features measured in the original data. We use PCA (principal component analysis) as feature transformation and SFS(Section of Forward Feature) as feature selection methods used in this study.

PCA transforms set of measurements of possibly correlated features into a set of features of uncorrelated features called principal components. PCA can find optimal linear transformation which can best represent the data in a minimum mean square error sense. On the other hand, SFS considers meaning of features which is significant and the goal of feature reduction which is to identify a dominant sub-feature set. Selection criteria usually involve the minimization of a specific performance measure for models fit to different subsets. The algorithms search for a subset of predictors that optimally model measured responses, subject to constraints such as added or removed features and the size of the subset.

Finally, after completing feature reduction, we have to ensure that classification models are feasible to be implemented into embedded H/W having little computing power. From experiments, to avoid very complex and heavy algorithms, we examine very compact and simple classification techniques, such as LDA (linear discrimination analysis), naïve Bayes, SVM (support vector machine), and a typical decision tree classifier.

III. EXPERIMENTS

This section introduces some of the methods commonly used to evaluate the performance of a classifier and then various combination procedures of feature analysis and classification model conducted in this experiment.

First, a good classification model must not only fit the training data well, it must also accurately classify records it has never seen before. This is important since a classification model that fits the training data too well can have a worse generalization error than a model with a higher training error, since the model must be over fitting to the training data. In this paper, we used 10-fold cross-validation method for estimating the generalization error of a chosen model in each procedure.

Second, we create eight combination procedure using two feature analysis techniques and four classification techniques as described previous section, apply these combinations to training data set, and then finally evaluate the result of them.

To begin with SFS, evaluation criterion should be decided that which features are effective to the target classifiers. We use wrapper methods to use the performance of the chosen target classifier's learning algorithm to evaluate each candidate feature subset[9]. Therefore, we apply LDA, naïve Bayes, SVM, and decision tree classifiers as target classifier's learning algorithm. Likewise SFS, we apply PCA method to g_35 training set as a method of feature reduction.

A. Experimental Results

The recall rates of fall pattern is presented to Figure 3, where the x axis of the graph indicates the number of reduced features; the y axis represents respectively recall rates over eight combination from PCA and SFS as feature reduction techniques to four different classification methods using 10-fold cross validation. In terms of recall rates, Figure 3 shows that all SFS based classifications are better than the all PCA based classifications over the number of g_35 feature dimension, because the recall rates are usually over 92%.

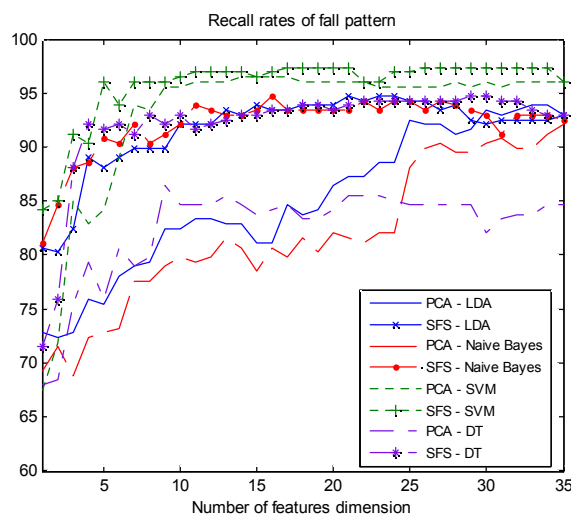


Figure 3. Recall rates of fall pattern over eight combinations

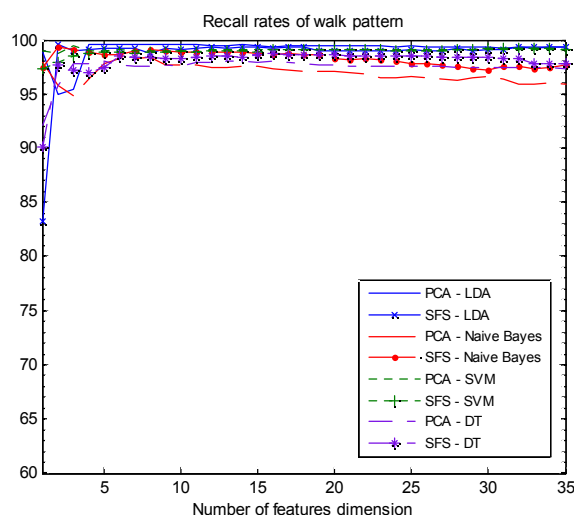


Figure 4. Recall rates of walking pattern over eight combinations

We can realize that increasing number of features does not make a significant contribution to enhance of the recall performance, since the recall rates over 10th feature dimension are constant approximately. In case of walking pattern, Figure 4 shows the recall rates over 5th feature dimension are constant approximately.

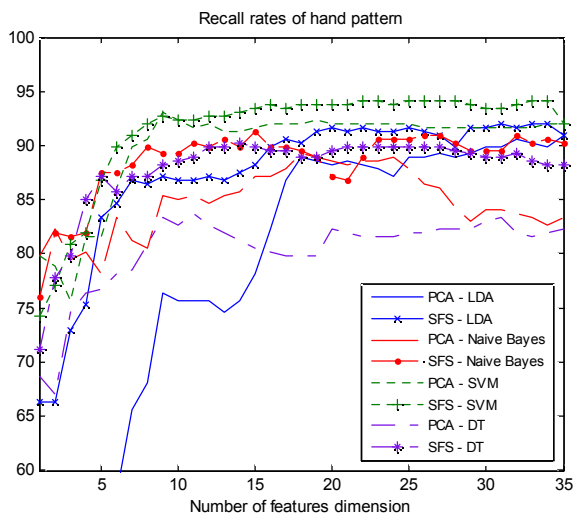


Figure 5. Recall rates of hand pattern over eight combinations

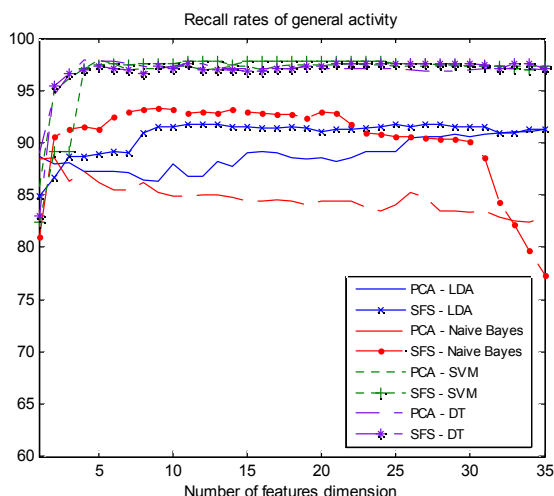


Figure 6. Recall rates of general activity pattern over eight combinations

We can find two facts; one is that SFS based feature selection is more reliable than PCA based on transformation through Figure 3 ~ 6 and the other is that we don't need a lot of features over 5~10 number of features. Figure 7 shows also the overall accuracy to each combination in terms of recognition rates. We can find the pair of SFS-SVM is the best of all in this paper and it is enough to have 10 numbers of features to achieve tolerable performance.

To ensure an embedded online classifier algorithm to be real-time execution in watch-type hardware, we have to find a minimal order of the feature vector satisfying maximal performance of recognition as much as possible, since the

classifier should be simple and tightly depend on the dimension of feature vector as a matrix operation. Consequently, we decided that simple is best and both SFS-SVM and SFS-DT method are the best candidate. The required real-time speed of the both methods in the embedded hardware should be 100Hz. We confirmed that real-time operation of both methods was performed in without delay in the same watch-type activity detector used data collection.

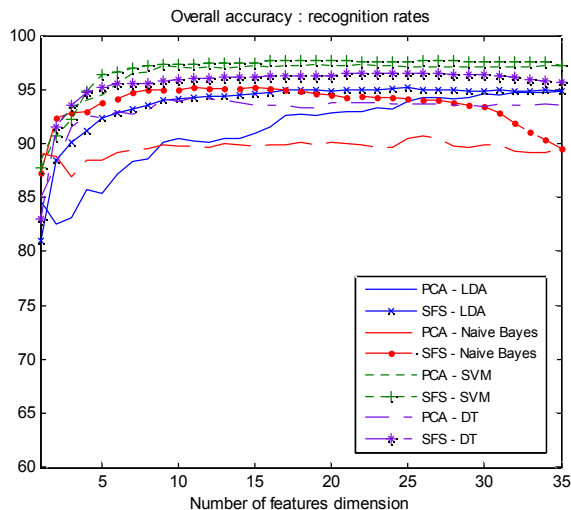


Figure 7. Overall recognition rates of each combination

IV. CONCLUSIONS

In this paper, we demonstrated the feasibility by using feature reduction techniques on large data sets and high dimension of a feature vector for decreasing a lot of computation and to ensure real-time operation under a limited low-power MCU environment. We observed that a reduced dimension of a feature vector could be applied to acquire the recognition rate over 90%. Moreover, we have checked that LDA-linear method and decision tree method show best performance results having just four features and simple decision rules. Besides improving the approach taken in this paper to a larger set of activities, we plan to classify multiple classes such as walking, running, lying, etc., as well as fall in future work. On the other issues, when we are labelling training data for each class, labelling itself is very exhausted task and requires a lot of cost, time and expertise of old people's activities. Another topic is a problem not to collect actual old people's full and general activities. Although we have actual fall data from old people, most of them may be gathered using unsupervised approach. How can we add fall data which are gathered by unsupervised approach to the supervised training sets? We have a plan to research combining unsupervised data set and semi-supervised learning techniques.

ACKNOWLEDGMENT

1. This work was supported by the R&D program of the Korea Ministry of Knowledge and Economy (MKE) and

the Korea Evaluation Institute of Industrial Technology (KEIT). [KI001836, Development of mediated Interface Technology for HRI.] (Acknowledged by Park and Kim)

2. This work was supported by the National Research Foundation (NRF) grant (No. 2011-0018264) of Ministry of Education, Science and Technology (MEST) of Korea. (Acknowledged by Choi)

REFERENCES

- [1] Luukinen H., Koski K., Honkanen R., Kivelä S. "Incidence of injury-causing falls among older adults by place of residence: a population-based study," *Journal of the American Geriatrics Society* vol.43, pp.871-876, 1995
- [2] Blake, A. J., Morgan, K., "Falls by elderly people at home:Prevalence and associated factors," *Age and Ageing*, vol.17, pp.365-372, 1998.
- [3] Pérolle, G., Sánchez, D., Abarrategui, M. I., Eizmendi, G.,Buiza, C., Etxeberria, I., Yanguas, J. J., "Fall Detection: Project of an Improved Solution,"1st International Workshop on Tele-Care and Collaborative Virtual Communities in Elderly Care, 2004.
- [4] Doughty, K., Lewis, R., McIntosh, A., "The Design of a Practical and Reliable Fall Detector for Community and Institutional Telecare," In *Journal of Telemedicine and Telecare*, vol.6, pp.150-154, 2000.
- [5] Kangas, M.; Konttila, A.; Winblad, I.; Jamsa, T.; , "Determination of simple thresholds for accelerometry-based parameters for fall detection," *Engineering in Medicine and Biology Society*, 2007. 29th Annual International Conference of the IEEE , pp.1367-1370, 2007.
- [6] U. Lindemann, A. Hock, M. Stuber, "Evaluation of a fall detector based on accelerometers: a pilot study, " *Medical & Biological Engineering & Computing*, vol. 43, pp.1146 – 1154, 2005.
- [7] Degen, T., Jaeckel, H., Rufer, M., Wyss, S., "SPEEDY: A Fall Detector in a Wrist Watch," In *IEEE2003, 7th International Symposium on Wearable Computers*, 2003.
- [8] Mattia B., Leopoldo R.," Wrist-Worn Fall Detection Device - Development and Preliminary Evaluation," *BIODEVICES* 2009,pp.368-371,2009. .
- [9] Breiman, Leo; Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*. Monterey, CA: Wadsworth & Brooks/Cole Advanced Books & Software. ISBN 978-0412048418.
- [10] F.van der heijden, R.P.W. Duin, D.de Ridder, *Classification, Parameter Estimation and State Estimation*, Wiley, 2004. ISBN: 978-0-470-09013-8.
- [11] J. H. Yo, M. S. Nixon, "Automated Markerless Analysis of Human Gait Motion for Recognition and Classification", *ETRI Journal*, vol. 33, no.2, 2011