Fall Detecion System for Monitoring an Elderly Person Based on Six-Axis Gyroscopes

Maojie Tang[*] and Dongxiu Ou

Key Laboratory of Road and Traffic Engineering of Ministry of Education, Tongji University, China

[*]Corresponding author

Abstract—**Accidental falls are crucial causes of death due to injury among the elderly. Many researches about fall detection applied complex algorithms and required heavy equipment. However, these approaches can hardly apply to the elderly's daily life. In this paper, we employ a six-axis gyroscope that integrated in a small smart bracelet. Users who wear the smart bracelet can get the information including acceleration for X, Y, and Z movement, and the rate of rotation in space. Then, we introduce three feature vector generation methods based on the information and feed these three vectors into support vector machine (SVM) algorithm for fall detection. From a dataset of 66 people, we show that the geometric parameters method is the best of the three with a high accuracy (100%), low false alarm rate (0%) and low missing alarm rate (0%) in a simulated home environment.**

Keywords—component; fall detecion system; six-axis gyroscope; feature vector generation method; support vector machine

I. INTRODUCTION

The trend for populations in China is to contain an increasing number of elderly people. As shown in [1], the percentage of population aged over 60 years in China is anticipated to rise from 15.3% in 2015 to 35% in 2050. As a consequence, the number of the elderly will exceed the society's capacity for taking care of them. Among such care, one important issue is to detect whether an elderly person has fallen or not [2]. According to [2], falls are the most serious causes of death due to injury and affect over one in every three elderly people. Furthermore, many complications such as fractures and abrasions always happen associated with falls. Although more than half of falls do not result in sever injuries, nearly 48% uninjured fallers can not get up without immediate assistance and this period of time also affect their safety. Therefore, it is crucial to design a system for fall detection. The intelligent system can detect anomalous behavior and send an alarm to certain caregivers or the elderly person's relatives by a modern communication method when accidental falls happen.

To achieve an exact and timely detecting result, a number of approaches for fall detection have been proposed in recent years [3]. Wang and his colleagues [4] presented a fall detector based on a hybrid model (WiFall) using wireless networks. Tapia et al [5] developed a real-time algorithm to detect physical activities and corresponding intensities based on C4.5 and Naive Bayes classifier. Liao, Huang and Hsu [6] introduced a fall detection system based on Bayesian Belief Network and built the probability model in different environments. Bourke and Lyons [7] developed an intelligent threshold algorithm to distinguish normal activities and falls. Mastorakis and Makris [8] designed a video approach for fall detection with no pre-knowledge. Chua and his two colleagues [9] designed a vision system to detect accidental falls and performed well with high detection results.

However, the abovementioned methods do not adequately address the problem at hand. For example, WiFall performs poorly when multiple people are conducting different activities. C4.5 and Naïve Bayes are difficult to deal with time series data. Artificial neural networks are not suitable for for this task, as lacking a sufficient number of fall samples can result in underfitting. In addition, the threshold algorithm relies on experts' experience and it's difficult to decide how to choose for each person. Although a vision system achieves a high accurate without pre-knowledge, 3D bounding box's first derivatives are difficult to collect in reality field. To overcome these limitations, this paper employs support vector machine (SVM) algorithm with three different feature vector generation methods for fall detection. We firstly collect data using six-axis gyroscopes, which surpasses three-axis gyroscopes in data reliability and data dimension. Then, three feature vector generation methods map the measurement space into the feature space by employing principal component analysis (PCA), extracting geometric parameters and reducing dimensions, respectively. Finally, these vectors have been fed into SVM for classification.

The remainder of this paper is organized as follows. Section 2 introduces the six-axis gyroscope and its advantages. The mathematical principles and calculation processes of SVM are explained in Section 3. Section 4 employs three methods for feature vector generation. Section 5 presents a numerical experiment using real human fall data for states detection, follower by a discussion and concluding remarks in Section 6.

II. GYROSCOPE

A gyroscope is a device used for measuring or maintaining orientation and angular velocity. It is a spinning wheel or disc in which the axis of rotation is free to assume any orientation by itself. When rotating, the orientation of this axis is unaffected by tilting or rotation of the mounting, according to the conservation of angular momentum. The device that make use of this technology include aircraft, race cars, video games, digital cameras and the most common use for the gyroscope is in our smartphone. In fall detection system, it can be integrated into a small smart bracelet and can sense our body motion information such as step speed and acceleration.

Nowadays, three-axis gyroscopes (Fig. 1a) have been widely integrated in smart bracelets for motion detection. Users can wear a smart bracelet on their wrists. Gyroscopes then

collect corresponding information for detecting user's posture and movement. However, the data collected by three-axis gyroscopes are rough and insufficient to deal with complex work. Therefore, six-axis gyroscopes (Fig. 1b) have been introduced to replace three-axis gyroscopes. Six-axis gyroscopes can collect acceleration for X, Y, and Z movement, and the rate of rotation in space (roll, pitch and heading). Compared with three-axis gyroscopes, the major advantage of six-axis gyroscope is that combing the six sensors can detect both unusual attitude and a fall. By centralizing the pitch controls and applying throttle, the six-axis gyroscopes come to a stable hover and are fit for our detection task.
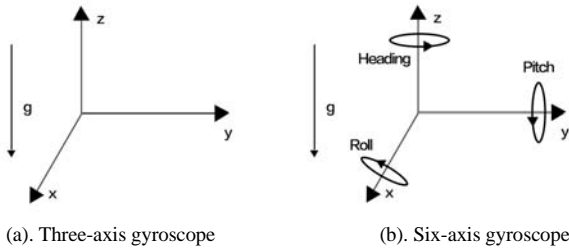

(a). Three-axis gyroscope   (b). Six-axis gyroscope

FIGURE I. THREE-AXIS AND SIX-AXIS GYROSCOPE

III. HUMAN FALL DETECTION WITH SUPPORT VECTOR MACHINE

Support Vector Machine (SVM) is a robust machine learning model using small-sample data for classification. An SVM model is a representation of samples as points in measurement space, mapped so that the samples of the separate categories are divided by a clear gap that is as wide as possible. New samples are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall. The gap referred to as a "hyper-plane" can be calculated by solving the following convex optimization problem,

$$\min \frac{1}{2} \|w\|^2$$

$$s.t. \ y_i(w^T x_i + b) \geq 1, i = 1,2,\dots,n$$

where x_i and y_i represents the sample and corresponding label ,respectively. (w, b) is the hyper-plane.

A more recent study [10] have summarized three mainly sub-modules in the literature for SVM, including data collection, feature vector generation and classification. In this paper, we will focus on feature vector generation module with three methods introduced below. For data collection module, we get the original sample from six-axis gyroscopes. Feature vector generation module maps the measurement space into the feature space and consists of feature extraction, feature selection and feature reduction [11].

- Feature extraction method is to find the suitable feature for classification directly influences the effect of classification results.

- Feature selection method inspects these features chosen by feature extraction method and eliminate features with no significant difference between two categories. There are many potential benefits of feature selection,

including reducing training time, improving classification performance and facilitating data visualization [12]. Further, several available feature selection methods in the literature have been proposed such as forward selection, backward selection, decision tree method, hypothesis testing (t-test) and so on.

- Feature reduction method reduces the measurement space into a low-dimensional space and easily visualize samples.

In addition to performing linear classification, SVM can effectively solve nonlinear classification using kernel trick. The kernel trick can map their inputs into high-dimensional feature space and separate the two category in the high-dimensional space.

IV. GENERATION METHODS FOR FEATURE VECTORS

A. Dimension Reduction Method with PCA

In this method, we form feature vectors by employing PCA. We firstly unify the original data to 13 dimension, which means that each feature vector consists of 13 elements. The reason that we choose 13 as our united dimension is that the process during human fall is very short and the longest curve collected in six-axis gyroscopes only exists 13 points. For the sample of less than 13 points, we supply zeroes into the sample. Finally, all samples are used in dimensionally reduction by PCA and the details of PCA are as follows.

- Step1: Collect a sample from six-axis sensors, and start with an input vector,

$$X = [x_1, x_2, \dots, x_n] \tag{1}$$

where n represents the cardinality of X.

- Step2: Construct the correlation matrix of X, which has n linear independent eigenvectors. For each element in the matrix, x is determined to satisfy

$$X = \sum_{k=1}^{n} c_k \phi_k \tag{2}$$

where c and ϕ represents the eigenvalues and eigenvectors of the matrix, respectively. Also, we can estimate x by

$$\hat{X} = \sum_{k=1}^{d} c_k u_k \tag{3}$$

where d represents the required dimensions and the u_k satisfies

$$u_i^T u_j = \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \tag{4}$$

- Step3: Employ the mean squared error ζ the Lagrange multiplier method. Then, we can get following formulas

$$\zeta = E\left[(X - \hat{X})^T (X - \hat{X})\right] = \sum_{k=d+1}^{\infty} u_k^T \gamma u_k \tag{5}$$

$$g(u_k) = \sum_{k=d+1}^{\infty} u_k^T \gamma u_k - \sum_{k=d+1}^{\infty} \alpha_k [u_k^T u_k - 1] \quad (6)$$

where γ equals to $E[XX^T]$ and α_k is the Lagrange multiplier.

- Steo4: Solve the simultaneous equations which contains (2) to (6), and we can conclude

$$[\gamma - \alpha_k I] u_k = 0 \quad (7)$$

- Step5: Calculate the mean squared error ζ associated with Lagrange multiplier

$$\zeta = \sum_{d+1}^{\infty} \alpha_k \quad (8)$$

If we take these d eigenvectors corresponding with the d biggest eigenvalues to represent x, the mean squared error is the smallest. This is the principle of PCA.

We experiment many times with varying dimensions of reduced feature vector and find that when reducing two dimensions, the classification result is the best. Figure 2 illustrates the distribution of the samples reduced to two dimensions.
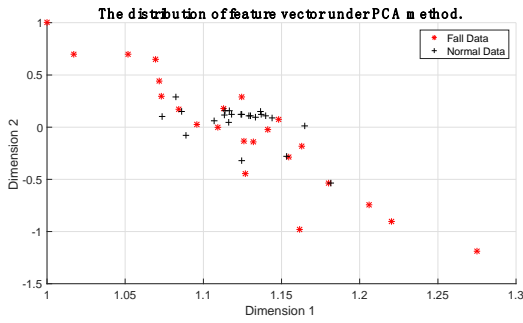

FIGURE II. THE DISTRIBUTION OF FEATURE VECTOR UNDER PCA METHOD

B. Geometric Parameter Feature Extraction

In this method, we extract statistical parameters of each sample and then establish the feature vector by these parameters. For one sample X sorted by its value,

$$X = [x_1, x_2, \dots, x_{13}] \quad (9)$$

we extract six parameters, including the first range, the second range, the third range, sum of the first to the third range and the fluctuation factors. These features are illustrated by Table 1.

TABLE I. FEATURES AND CORRESPONDING ATTRIBUTES

No.	Feature	Remark
1	$P_1 = x_1 - x_{13}$	The first range
2	$P_2 = x_2 - x_{13}$	The second range
3	$P_3 = x_3 - x_{13}$	The third range
4	$P_4 = \sum_{k=1}^{3} P_k$	The sum of 1st to 3rd range
5	$P_5 = x_1$	Maximum
6	$P_6 = \dfrac{P1}{\sqrt{\frac{\sum_{k=1}^{13} x_i^2}{13}}}$	Fluctuation factor

Then, this paper introduces a hypothesis testing (t-test) for feature selection, and the statistic results are shown in Table 2. For the t-test, the null hypothesis that two category have an equal mean. A low p-value ($P \le 0.05$) indicates that there existed distinctive difference between two categories and can be used for SVM. Table 2 shows that the p-value results are all less than 0.05, which indicate that all features can be adopted for classification.

TABLE II. THE P-VALUES OF ALL FEATURES

Axis	Feature	P-Value	Axis	Feature	P-Value	Axis	Feature	P-Value
1	P_1	1.40E-06	2	P_1	1.00E-08	3	P_1	6.20E-06
	P_2	8.00E-06		P_2	3.90E-07		P_2	2.40E-07
	P_3	6.00E-06		P_3	1.70E-06		P_3	5.40E-07
	P_4	6.20E-05		P_4	1.20E-08		P_4	9.90E-11
	P_5	3.40E-13		P_5	1.30E-06		P_5	4.80E-05
	P_6	6.80E-06		P_6	1.30E-17		P_6	8.20E-10
4	P_1	3.60E-05	5	P_1	3.60E-05	6	P_1	4.70E-07
	P_2	4.70E-05		P_2	4.70E-05		P_2	2.40E-05
	P_3	4.50E-05		P_3	4.50E-05		P_3	4.80E-05
	P_4	2.20E-05		P_4	2.20E-05		P_4	5.30E-07
	P_5	4.70E-05		P_5	4.70E-05		P_5	1.70E-06
	P_6	0.0178		P_6	0.0119		P_6	0.001

After selecting features for each axis, we can get a vector with 36 elements. In addition, we reduce its dimension to 2 by PCA. The 2-dimension features are referred as to Geometric Parameters with PCA (GP-PCA) method and shown in Figure 3.
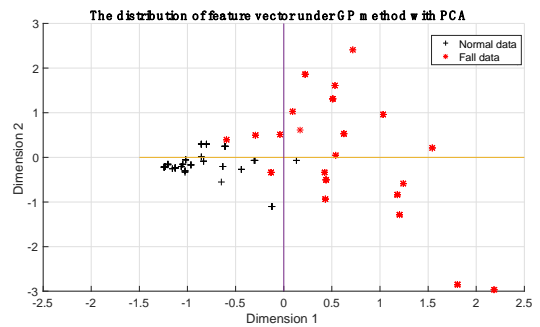

FIGURE III. THE DISTRIBUTION OF FEATURE VECTOR UNDER GP WITH PCA

V. EXPERIMENT RESULTS

We experimented on the volunteers' data collected from smart bracelets with JY-901 six-axis gyroscopes. The sample set we experimented on covers 43 normal data and 23 fall data. Then, we added Gaussian noise with average 0 and standard deviation

0.01 to the real field data so as to expand the experimental sample set. The dataset was split randomly into two subsets (training set and testing set) which account for 70% and 30% of the dataset, respectively. Moreover, we choose Gaussian kernel $\kappa(x_1, x_2)$ as the kernel trick,

$$\kappa(x_1, x_2) = Ce^{(-\gamma|x_1 - x_2|)^2} \qquad (10)$$

and determine penalty factor C and kernel parameter γ by 10-fold cross validation method [13]. In 10-fold cross validation, the dataset has been split into 10 equal sized subsamples. Of the 10 subsamples, one is retained as the validation set for testing and the rest 9 subsamples is merged as training set. The cross-validation process is then repeated 10 times (the folds) and 10 results from the folds can be averaged to produce the best parameter. Finally, the best (C, γ) can be chosen when the average accurate reaches the highest.

Table 3 illustrates the detection results of three feature vector generation methods. In Table 3, we introduce two parameters, false alarm rate (FAR) and missing alarm rate (MAR), to evaluate the generalization performance of different feature generation methods. FAR means the probability of identifying the normal data as fall data and the MAR means the probability of identifying the fall data as normal data.

TABLE III. THE FALL DETECTION RESULTS

	PCA	GP-PCA	GP
Accurate rate	88.77%	93.58%	100%
FAR	6.95%	1.60%	0
MAR	4.28%	4.81%	0

From the results, we can compare the advantages of these three methods.

- In general, GP method is the best of the three because the features extracted by GP are based on expert's experience. Therefore, SVM algorithm using GP can generate the best hyper-plane and separate each category with high accuracy.

- Compared to GP method, GP-PCA offers a generally acceptable level of classification quality, but it still makes incorrect classifications and generates a relatively high FAR and MAR, which prevents the application of the GP-PCA method in practical applications. The reason that GP-PCA performs worse than GP is that GP makes full of all features, whereas GP-PCA disregards certain information when applying PCA.

- Compared to GP method, the PCA method exhibits classification results and cannot be used for fall detection because of its high FAR and MAR. GP performs better than PCA for two reasons. First, PCA lose some information when reducing dimensions. Second, original data may exist noises that can influence the detection results without data preprocessing.

VI. CONCLUSION

In this paper, we use the smart bracelets worn on the body from the real field data and employ SVM to design a timely fall detection classifier. As for the gyroscopes integrated in bracelets, we adopt the six-axis gyroscopes because of obtaining high-dimensional data. The experimental results show that the GP method can detect falls with 100% accuracy, which is superior to the other two methods introduced above. Furthermore, the low FAR and MAR demonstrate that GP method performs well with strong robustness. Further work will supply samples in different environments and analyze the elderly people's daily behaviors, including sitting, walking, standing and using stairs. Meanwhile, the smart bracelets integrated with six-axis gyroscopes will focus on other applications such as the analysis of youth sports and babysitting.

REFERENCES

[1] Hao X, Gu J, Ying X, et al. Social support and care needs of the disabled elderly population: An empirical study based on survey data from Beijing, China[J]. Bioscience trends, 2017, 11(5): 507-515.

[2] J. Halter, J. Ouslander, M. Tinetti, S. Studenski, K. High, S. Asthana, and W. Hazzard, "Hazzard's geriatric medicine and gerontology," Sixth Edition, McGraw-Hill, 2009

[3] Luštrek M, Kaluža B. Fall detection and activity recognition with machine learning[J]. Informatica, 2009, 33(2).

[4] Wang Y, Wu K, Ni L M. Wifall: Device-free fall detection by wireless networks[J]. IEEE Transactions on Mobile Computing, 2017, 16(2): 581-594.

[5] Tapia, E. M., Intille, S. S., Haskell, W., Larson, K., Wright, J., King, A., and Friedman, R. Real-Time Recognition of Physical Activities and Their Intensities Using Wireless Accelerometers and a Heart Rate Monitor. In Proceedengs of the 11th IEEE International Symposium on Wearable Computers, pp. 37–40, 2007.

[6] Liao Y T, Huang C L, Hsu S C. Slip and fall event detection using Bayesian Belief Network[J]. Pattern recognition, 2012, 45(1): 24-32.

[7] Bourke, A. K., and Lyons, G. M. A threshold-based fall-detection algorithm using a bi-axial gyroscope sensor. Medical Engineering & Physics, vol. 30, issue 1, pp. 84–90, 2006.

[8] Vishwakarma, V., Mandal, C., and Sura, S. Automatic Detection of Human Fall in Video. Pattern Recognition and Machine Intelligence: Automatic Detection of Human Fall in Video, pp. 616–623, 2007.

[9] Chua J L, Chang Y C, Lim W K. A simple vision-based fall detection technique for indoor video surveillance[J]. Signal, Image and Video Processing, 2015, 9(3): 623-633.

[10] Oskoei M A, Hu H. Support vector machine-based classification scheme for myoelectric control applied to upper limb[J]. IEEE transactions on biomedical engineering, 2008, 55(8): 1956-1965.

[11] Pal M, Foody G M. Feature selection for classification of hyperspectral data by SVM[J]. IEEE Transactions on Geoscience and Remote Sensing, 2010, 48(5): 2297-2307.

[12] Weston J, Mukherjee S, Chapelle O, et al. Feature selection for SVMs[C]//Advances in neural information processing systems. 2001: 668-674.

[13] Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection[C]//Ijcai. 1995, 14(2): 1137-1145