An Approach Based on Statistical Features to Fall Detection

Hui Wang[1], Xiaohe Chen[2], Xingyu Wu[1], XinjianChen[1] and Lirong Wang[1*]

[1]*school of electronic and information engineering, soochow university, china*
[2]*suzhou institute of biomedical engineering and technology chinese academy of sciences*
wanghuiliuwen@foxmail.com

Abstract

Falls in elderly is a very serious health problem. For these years, the wearable devices based on tri-axial accelerator has been proven to be an effective way to fall detection. Most current methods for fall detection are based on threshold and machine learning. A approach based on statistical features was proposed to distinguish falls and normal activities of daily living(ADL) in this paper. What is worth mentioning is that Kernel Principal component analysis(KPCA) is firstly used to extract the statistical features from the original 3D data of acceleration, we don't need to design features specially. The support vector machine (SVM) algorithm and K-Nearest Neighbor(KNN) algorithm are combined for prediction. Finally the validation of the prediction is done to improve the accuracy. Algorithm is mainly conducted on the public databases(UCI). And our method obtained the result is proved to be better compared with the other literature based on this public databases.

Keywords: fall detection; KPCA; statistical features; SVM; KNN

1. Introduction

As the global population aging rapidly, the number of elderly who live alone increasing rapidly. Falls for young people generally no harm, but for the elderly falls is really a big threat. And the incidence of falls in elderly people is higher compared with young people. If falls of the elderly cannot be found in time, the elderly won't get the timely medical treatment, which is very serious. Falls may cause in a series of damage which directly affects the quality of life of the elderly and even threats life. So for the elderly the real-time automatic fall detection has important significance.

Because of the huge demand of fall detection in health care industry, many fall detection technologies has been developed. The technologies can be generally divided into three categories, wearable device method, camera based method and ambience device method [1-2]. Among these technologies, considering the mobility, portability and the price, the wearable device system is simpler, cheaper and more portable compared with the other two systems [3]. The small device will not bring any serious inconvenience to the elderly. In the wearable device system, the tri-axial acceleration sensor was widely used. The public database our method used is also acquired from the acceleration sensor.

To this end, many researchers have applied many approaches to fall detection based on the wearable device. A.K. Bourke [4] *et. al.,* proposed a threshold-based fall-detection method using a bi-axial gyroscope. The paper points out when a fall activity occurs, the resultant angular velocity and the angular acceleration are all greater compared with the ADL, and the resultant change in trunk-angle is also greater. So three thresholds were set to fall detection.

Luca Palmerini *et. al.,*[5] also utilize the threshold-based method with the acceleration

---

*Corresponding Author

sensor. As the acceleration sum vectors of falls always present a peak and the value is also greater [6-7], a threshold on the value of the peaks is used to construct a window. Then a prototype of typical fall pattern by averaging the available real-world fall signals is created as the mother wavelet. Finally Then the degree of the similarity of the signal window and the mother wavelet is computed, here a threshold of the degree is set to predict the fall.

Except the threshold-based method Most methods used in literature are various features extracted from the data window with machine learning [8-9]. In the paper of Ryan M.Gibson *et. al.,* [10], a method by combining classifiers is used to improve the performance of single classifier .Ryan M.Gibson *et. al.,*[10] obtained the samples by sliding window, then the discrete wavelet transform for a Daubechies2 level-3 function is applied to extract the wavelet approximation coefficients as features for classifier.

Ahmet Turan Ozdemir *et. al. ,*[11] developed an automated fall detection system with wearable device. In this system, a 4s time window around the point of peak resultant acceleration from the west sensor is as a sample. Then the features are extracted from the time series in each window. The features include the maximum, minimum, variance and mean values etc, some values of the autocorrelation sequence are used as the features, then the features in frequency domain are also used. Six machine learning method were used to fall detection. The KNN classifier and the least squares method are proven to be have better accuracy.

In the paper of Shuai Huang *et. al.,* [2], the cost sensitivity analysis in fall detection based on tri-accelerator are presented. This method utilizes the discrete features [2] extracted from the sliding window, and the probabilistic svm is used for prediction, then analyze the cost sensitivity based on the probabilistic outputs. Finally the validation is used to improve the accuracy.

In this paper, we consider the problems of both the preliminary prediction by threshold -base method and the features for prediction with machine learning. We centered the greater peak for time window based on threshold, instead of using the method of sliding window directly, because segmenting the time series data directly may lead to the window cannot completely included the the full data of a fall sample thus increasing the difficulty to prediction.

Considering the features used in machine learning for fall detection or activity recongnition are always extracted by these two types :extracting discrete features artificially from time window [5-12], extracting the wavelet based features by wavelet transform [7-13]. In this paper, we first proposed to extract the statistical features from the original data directly instead of design features artificially. The KPCA is used on the original data to extract the statistical features directly. The type of features extraction by KPCA is simple and can get the representative features. Because there is no need to filter and reduce dimension, and it can still keep all the features of the original data as soon as possible. Here we choose to combine the probabilistic outputs SVM and KNN to get better performance, and we also use validation to improve the accuracy. Figure 1, shows the framework of our method.
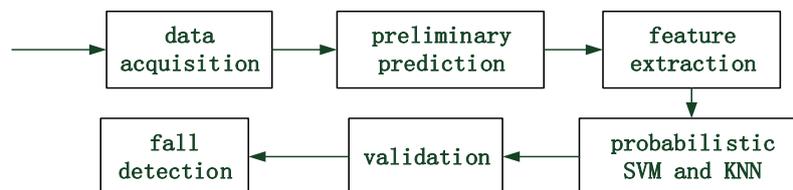

Figure 1. The Framework of our Method

The rest of the paper is organized as follows: the public database and all the algorithms

used in the paper are described in section 2, the performance analysis of our method and the comparison with other methods are described in section 3, section 4 gives the conclusion of this paper and the direction for future research.

2. Method

The method described in this paper is based on the statistical features to fall detection. Figure 2, Shows the procedure of our method.
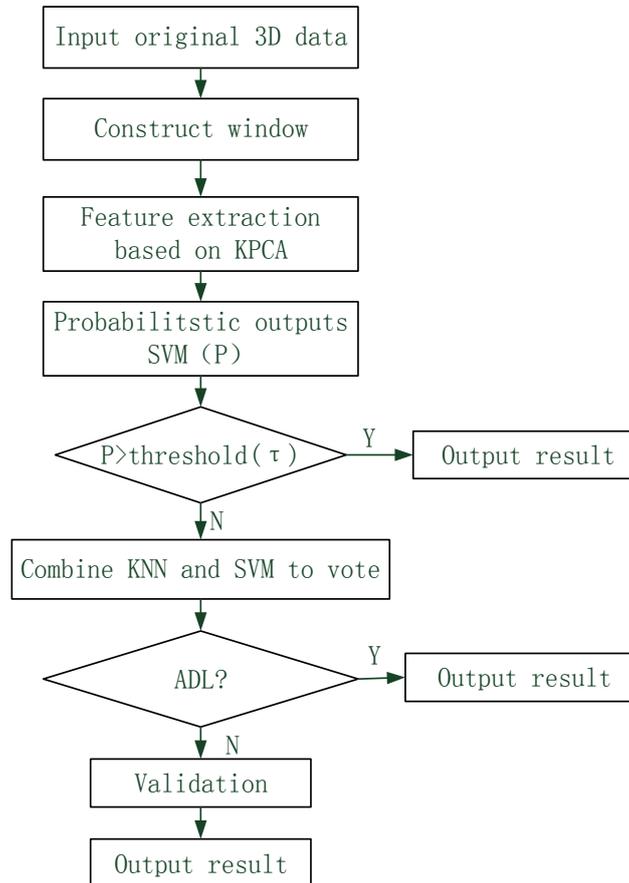

Figure 2. The Framework of our Method

A threshold-based method is used for the preliminary prediction to construct time window. Then the statistical features are extracted from the samples, finally the probabilistic outputs SVM and KNN are combined to predict. Considering some samples are hard to predict which always present low probabilities near to 0.5 from the probabilistic outputs SVM. So the KNN and SVM is combined to vote for the samples with low probabilities, here we set a threshold for probabilities. In order to improve the accuracy, the validation is used for the samples which are predicted as falls with the combined classifier.

2.1. Data Acquisition

At present, most studies use the data collected by themselves. However the difficulties degree of recognizing the different types of falls and ADL is different, in order to evaluate the performance of our method, a public database "Localization Data for Person Activity Data Set" of UCI machine learning is adopted in this paper. Data contains recordings of five people performing different activities. Each person wore four sensors (tags) while

performing the same scenario five times. People used for recording of the data were wearing four tags (ankle left, ankle right, belt and chest). This article chooses three dimensional acceleration time series from the chest, because the data from chest change has more obvious trend when fall occurs compared with other parts.

2.2. Preliminary Prediction

A piece of original three axis acceleration time series and the resultant acceleration are shown in Figure 3. The black flagged part is falling state, it can be seen that the resultant acceleration presents a peak and the value is greater in the time series. As the original resultant acceleration series contains noise,we use the wavelet transform to denoising , the results shows in Figure 4.
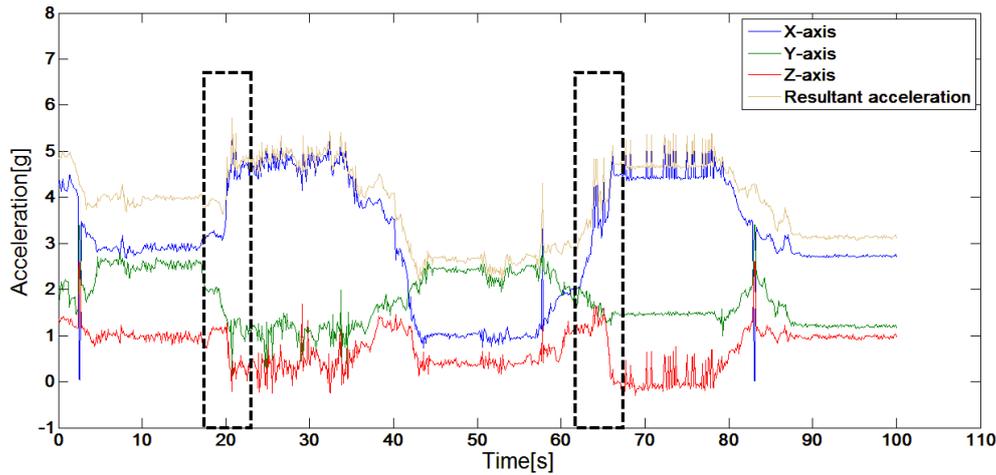

Figure 3. Acceleration Plots for the Three Axis and the Resultant Acceleration

Then we set a threshold τ^P on the value of peaks, if the peak present in resultant acceleration series and it's value is greater than τ^P, a 10s time window centered on the peak will be constructed from the original three axis acceleration. Here we choose the original three axis acceleration data without filtering, so the sample could keep the complete information. Additionally it is very important to choose the threshold, high threshold could lead to miss some fall data, but if the threshold is too low, it will increase the amount of calculation. The selection of threshold mainly through the statistics of the values of all falls data from the resultant acceleration time series, then the smallest fall data will be the threshold τ^P .
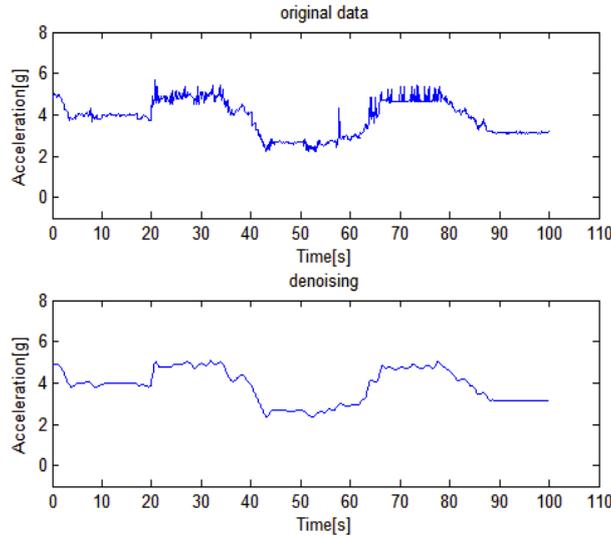
Figure 4.The Original Resultant Acceleration Data and the Denoising Result

2.3. Statistical Feature Extraction

In this paper, the method of KPCA is used to extract the statistical features from the samples directly. KPCA is a nonlinear principal component analysis(PCA) method, which through nonlinear mapping kernel function map the input data space(x_i, x_j) to a high dimensional feature space F as present in Eq.1, and then the PCA is used in F space. The nonlinear statistical features can extracted effectively. The principle of KPCA shows in Figure 4, Φ(x) is the nonlinear function which is complex while the kernel method K(x_i, x_j) is more simple. So the kernel method is used in KPCA. This paper use the RBF-kernel function as present in Eq.2. In this method the KPCA are used in three axis acceleration separately.
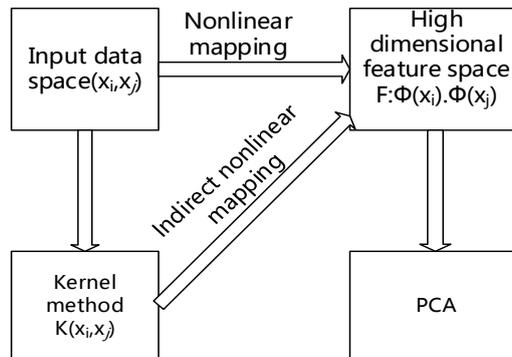

Figure 4. The Principle of KPCA

$$(x_i, x_j) \rightarrow K(x_i, x_j) = \Phi(x_i) \cdot \Phi(x_j) \qquad (1)$$

$$K(x, x_i) = \exp(-\frac{\|x - x_i\|^2}{\sigma^2}) \qquad (2)$$

2.4. Classifier

The KNN algorithm is a non-parametric density estimation function for classification [10]. The distance between the testing sample and training samples are evaluated, then

find the nearest K neighbors, the class of the testing sample determined by the K neighbors' classes. And the class belong to the class which appear most times of the K samples. In this paper the euclidean distance is used, and we set the K-value=5.

In SVM model, a best hyperplane is found which can separate two classes of data points right as possible, and make the distance between the hyperplane and the separated data point be the furthest.

A two class probabilistic output SVM and KNN are combined based on the features mentioned. The outputs of the probabilistic output SVM is the probabilities of the samples belong to each class [2]. Here KNN is combined with SVM by the simple type of voting when the probabilities are lower than the threshold τ. In order to keep the rate of the fall samples detected correctly higher, the type of voting is the sample predicted to a fall sample when one or two classifiers predicted the sample is a fall.

2.5. Validation

The type of voting mentioned may lead to some ADL samples are false recognized, here the validation used for the samples which predicted as fall with the combined classifier. The validation utilizes the character that when a fall occurs the time series with the state of sitting or lying will last for some time [2]. So samples by the auxiliary sampling of these samples is used.Then match the data with training the samples whose state is lying or sitting, if the samples match these two state, it will be predicted as fall samples.

3. Experiment and Result

The experiment based on a public database of UCI is set to better evaluate the algorithm, which can be measured by the success criteria of Accuracy, Sensitivity, Specificity and the AUC (area under the ROC curve): .

Accuracy is defined as the ratio of samples that recognized accurately to all the test samples as in Eq.3.

$$Accurac = \frac{tp + tn}{tp + tn + fp + fn} \tag{3}$$

Sensitivity is defined as the ratio of the falls that correctly detected to total number falls in the samples as in Eq.4.

$$Sensitivity = \frac{tp}{tp + fn} \tag{4}$$

Specificity is defined as the ratio of ADL that correctly recognized to all the ADL in the samples as in Eq.5

$$Specificit y = \frac{tn}{tn + fp} \tag{5}$$

The tp is the number of samples that falls identified correctly, tn is the number of samples that ADL recognized correctly and fp is the number of samples that ADL misjudged as the falls ,fn is the number of samples that fall misjudged as the ADL. AUC is the measurement criteria of a classification model, it corresponds to the area under the ROC curve. Figure 5, shows the ROC curve of our method as follows.
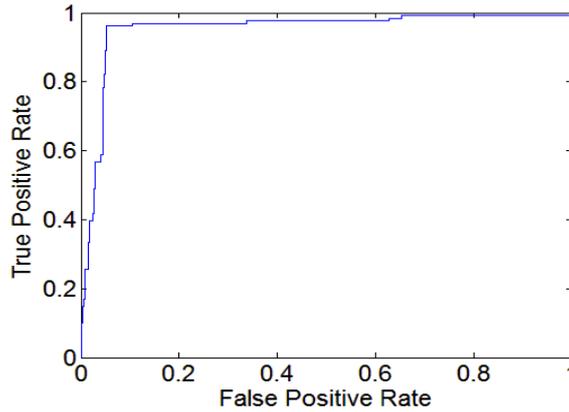
Figure 5. The ROC Curve of our Method

Classification based on discrete features and wavelet based features by wavelet transform is experimented to contrast with statistical features by KPCA, and the results are shown as Table 1. We can certainly confirm that the result based on the KPCA is better than others obviously,with the Accuracy and the Sensitivity of the KPCA is larger than the others. So we can conclude that method based on statistical features to fall detection is effective, competitive and feasible.

Table 1. Comparison of the Results with Different Features

	accuracy	sensitivity	specificity	AUC
discrete features	87.66%	81.395%	90.003%	0.8728
wavelet based features	92.128%	89.147%	93.26%	0.9206
statistical features	95.106%	96.124%	94.72%	0.9525

In order to better illustrate the advantages of the KPCA method, some comparison with the two other literature, which are also based on the public database of UCI is done. And the result is shown as Table 2. We can see that the result based on the KPCA is better than others .

Table 2. Comparison of the Results with other People

	accuracy	sensitivity	specificity
Lustrek *et. al.,*[3]	91.33%	84.00%	98.67%
Shuai Huang *et. al.,*[2]	90.50%	96.67%	84.32%
Our method	95.106%	96.124%	94.72%

4. Conclusion

In this paper we proposed a statistical features based method, first a threshold based

method used for the preliminary prediction to get the time window, then the features are extracted by KPCA directly from the window, and then the classifiers of combining the probabilistic outputs SVM and KNN are applied to classify, finally the validation is used to improve the accuracy. The results are improved compared with other literature used the same database. As a future work, we are planning to improve the accuracy by other method and we will study different methods of fall detection.

References

[1]   X. Yu. "Approachs and principles of fall detection for elderly and patient", HEALTHCOM, Singapore, **(2008)** July 7-9.
[2]   S. Huang, Y. Yang and W. Liu, "An Enhanced Fall Detection Approach Based on Cost Sensitivity Analysis", ACIS International Symposium on Software and Network Engineering, ; Seoul, Korea, **(2011)** Decemeber 19-20.
[3]   B. Kaluža, V. Mirchevska, E. Dovgan, M. Luštrek and M. Gams, "An Agent-based Approach to Care in Independent Living", Lecture Notes in Computer Science, **(2010)**.
[4]   A. K. Bourke and G. M. Lyons, "A threshold-based fall-detection algorithm using a bi-axial gyroscope sensor", Medical Engineering & Physics, vol. 30, no. 1, **(2008)**.
[5]   L. Palmerini, F. Bagalà, A. Zanetti, J. Klenk, C. Becker and A. Cappello, "A Wavelet-Based Approach to Fall Detection", Sensors, vol. 15, no. 5, **(2015)**.
[6]   F. Bagalà, C. Becker, A. Cappello, L. Chiari, K. Aminian, J. M. Hausdorff, W. Zijlstra and J. Klenk, "Evaluation of Accelerometer-Based Fall Detection Algorithms on Real-World Falls", PLoS ONE, vol. 7, no. 5, **(2012)**.
[7]   M. Yuwono, B. D. Moulton, S. W Su, B. G Celler and H. T. Nguyen, "Unsupervised machine-learning method for improving the performance of ambulatory fall detection systems", Bio Medical Engineering On Line, vol. 11, no. 9, **(2012)**.
[8]   J. He, C. Hu and X. Wang, "A Smart Device Enabled System for Autonomous Fall Detection and Alert", International Journal of Distributed Sensor Networks, vol. 10, **(2016)**.
[9]   N. Pannurat, S. Thiemjarus and E. Nantajeewarawat, "Automatic fall monitoring: a review," Sensors. vol. 14, no. 7, **(2014)**.
[10] R. M. Gibsona, A. Amiraa, N. Ramzana, P. Casaseca-de-la-Higueraa and Z. Pervezaa, "Multiple comparator classifier framework for accelerometer-basedfall detection and diagnostic", Applied Soft Computing, vol. 39, **(2016)**.
[11] A. T. Özdemir1 and B. Barshan. Detecting Falls with Wearable Sensors Using Machine Learning Techniques.Sensors , vol. 14, no. 6, **(2014)**.
[12] J. R. Kwapisz, G. M. Weiss and S. A. Moore, "Activity Recognition using Cell Phone Accelerometers", SIGKDD Explorations, vol. 12, no. 2, **(2010)**.
[13] C. Catal, S.Tufekci, E. Pirmit and G. Kocabag, "On the use of ensemble of classifiers for accelerometer-based activity Recognition", Applied Soft Computing, vol. 37, **(2015)**.