ISSN: 2472-3703 DOI: 10.6911/WSRJ.202101_7(1).0003

Application of Stochastic Gradient Boosting Algorithm in Human Behavior Recognition

Jia Yang[1, a], Guoliang Yuan[1]

[1]Department of Information Engineering, Shanghai Maritime University, China

[a]973460250@qq.com

Abstract

Aiming at the problem of human body's daily behavior recognition, a classification model based on machine learning algorithm is proposed. First, the acceleration data in three directions of the human body is collected through the IMU three-axis acceleration sensor, and the acceleration is processed by wavelet filtering, and then the acceleration is extracted. The time domain, frequency domain, and time-frequency domain total 27-dimensional features. Use machine learning classification algorithms to comprehensively compare the classification results of different models to determine the best classification model. Experimental results show that the stochastic gradient ascent algorithm has the best classification accuracy. The algorithm realizes the recognition of seven daily behaviors such as going upstairs, going downstairs, running, walking, jumping, sitting down, and falling. The average accuracy is up to 85%.

Keywords

IMU sensor; Behavior recognition; Wavelet filtering; Stochastic gradient boosting; Vehicle detection.

1. INTRODUCTION

As the proportion of middle-aged and elderly people in the total population increases, the incidence of various chronic diseases has also increased significantly.

In order to prevent the elderly from becoming a burden to the family, real-time care of the elderly is particularly important. The monitoring of the daily behavior of the elderly can provide the elderly with comprehensive care [1]. Therefore, the recognition of human behavior has become a hot issue in the scientific community. With the rapid development of machine learning and deep learning technology, its application to the field of human behavior recognition has become a development trend, and more and more The attention of researchers.

Looking at the research status in the field of behavior recognition in recent years, data collection methods can be roughly divided into two categories, namely behavior recognition methods based on computer vision [2] and behavior recognition methods based on wearable devices [3]. Behavior recognition methods based on computer vision are easily disturbed by the environment, and the recognition effect is not good in places with many people and complex backgrounds. Human behavior recognition based on acceleration sensor has attracted more and more researchers' attention due to its simple equipment and strong anti-interference ability. The essence of this method is to detect the human body acceleration signal to recognize its movement pattern. The main work of this algorithm is the feature extraction of acceleration signal and the selection of recognition algorithm. For feature extraction, it includes the mean and standard deviation in the time domain; the amplitude and energy of the frequency domain signal components [4-5]. In terms of classification methods, support vector machines (SVM) [6],

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202101_7(1).0003

random forest (RF) [7] neural networks (BP), classification by threshold, etc. are widely used. Literature [8] proposed a recognition algorithm based on Hidden Markov Model HMM. Literature [9] proposed a Kalman filtering algorithm for human body state recognition.

Based on the above-mentioned literature, the author adopts the mean, standard deviation, maximum, minimum, range, mode, slope mean, and slope standard deviation in the feature extraction in the feature extraction. A total of 8 features are adopted in the frequency domain. The shape mean, shape standard deviation, shape skewness, shape kurtosis, amplitude mean, amplitude standard deviation, amplitude skewness, amplitude kurtosis, DC component, the first 5 peaks, the frequency corresponding to the first 5 peaks, and the total energy is 15 A feature; in the time-frequency domain, four types of wavelet packet entropy are adopted, a total of four-dimensional features. The stochastic gradient ascent algorithm is adopted in the recognition algorithm, and the experiment proves that the algorithm has the best recognition accuracy. The flow chart of this experiment is shown in Figure 1.

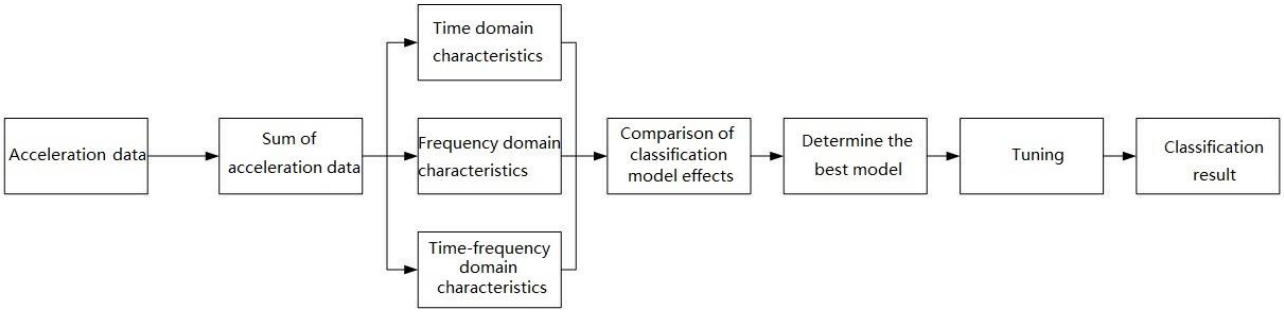


Figure 1. Experimental flowchart

2. DATA COLLECTION AND PREPROCESSING

2.1. Data Collection

The IMU device used in this article is based on MPU6050. The device integrates some acceleration and angular velocity measurement units. The module can be connected to the computer via Bluetooth to facilitate data collection and data processing, as shown in the figure Shown in Figure 2. Here, we set the sensor sampling frequency to 10HZ, set a sampling time of 25.6 seconds for four actions such as walking, running, and up and down, and set a sampling time of 6.4 seconds for jumping, sitting, and falling. Perform data collection as described above.

Figure 2. IMU module

Among the various parts of the human body, the waist has the advantages of small changes in the center of gravity and small changes in the collected data. Therefore, I chose to fix the IMU on the waist of the human body through a thin line, which constitutes a simple system for our experiment.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202101_7(1).0003

2.2. Pretreatment

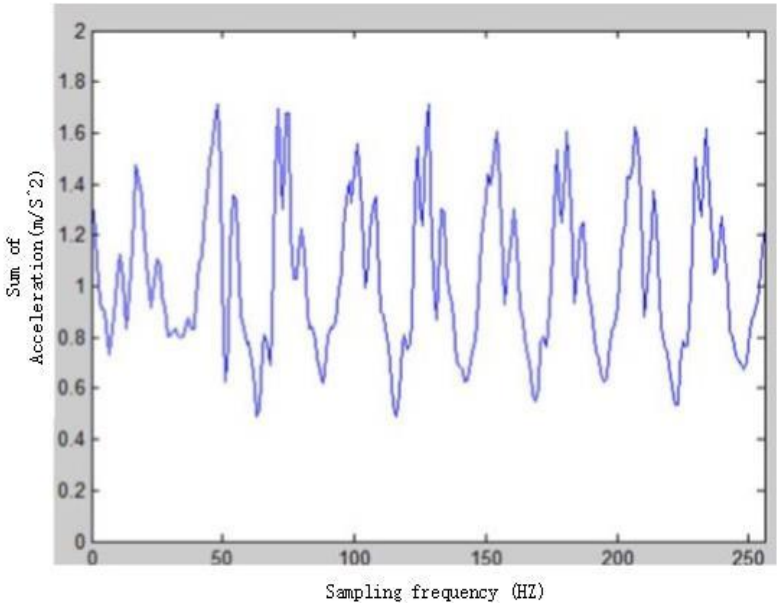
2.2.1 Sum Acceleration

Since the acceleration collected by the sensor is divided into three directions: X-axis, Y-axis, and Z-axis, in order to overcome the influence of gravitational acceleration and facilitate data processing, we calculate the acceleration in the three directions of X-axis, Y-axis and Z-axis sum acceleration:

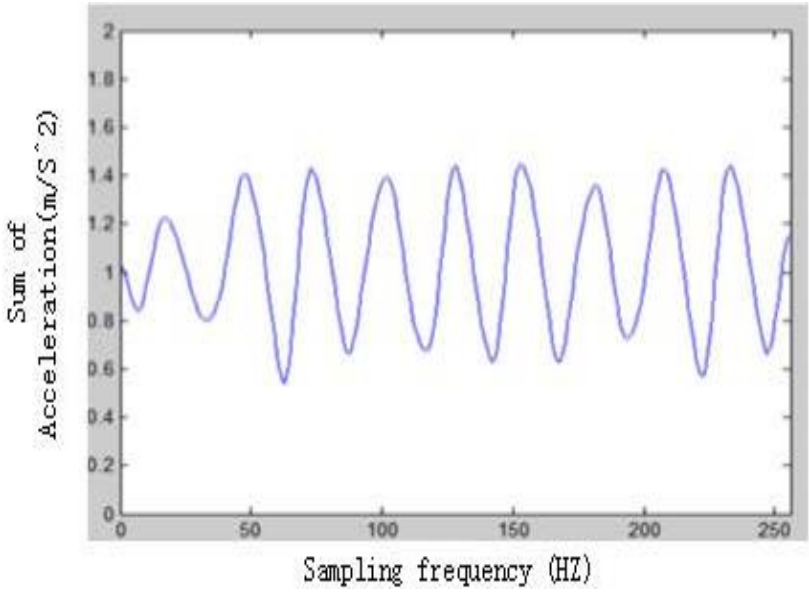
$$A_{sum} = \sqrt{a_x + a_y + a_z} \tag{1}$$

Among them, is the acceleration in the three directions of X axis, Y axis and Z axis.

2.2.2 Wavelet Denoising



(a)



(b)

Figure 3. (a)Raw data before denoising; (b)Data after denoising

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202101_7(1).0003

Take walking as an example, as shown in Figure 3(a), the image shows the image of walking data before denoising. There are many glitches in this data. This is caused by body shaking when collecting signals. It can be predicted that if you directly The above signals are directly put into the classification model training without processing, which will inevitably cause the classification effect to be greatly reduced, so it is particularly important to choose a suitable denoising method. Traditional denoising methods mainly include linear filtering and nonlinear filtering, such as median filtering and Wiener filtering. The shortcomings of the traditional denoising methods are that the entropy of the signal transformation and the signal can be increased, the non-stationarity of the signal can not be characterized and the correlation of the signal can not be obtained, and the wavelet transformation can overcome the above shortcomings. As the collected signal contains many singularities and the detailed characteristics of the signal need to be preserved, we choose the wavelet transform modulus maximum method to filter, and the filtering result is shown in Figure 3(b).

3. ACCELERATION FEATURE EXTRACTION

3.1. Windowing

In order to realize real-time processing of data and reduce the amount of data, it is necessary to perform windowing processing on the denoised data, so as to extract features and facilitate the sending of the data to the machine learning model for training. For the four actions of walking, running, up and down stairs, the length of the sliding window is 32, and for

Jumping, sitting down, falling down, because the duration is shorter and there are fewer data points that can be processed, the selected window length is 16, and the repetition rate of the two windows is set to 50%.

3.2.1 Time Domain Characteristics

After being windowed, the mean value is expressed as:

$$u = \frac{1}{N}\sum_{i=1}^{N} A_i \tag{2}$$

Which represents the window size, represents the acceleration value. The standard deviation indicates the degree of data deviation, which is calculated as:

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(A_i - u)^2} \tag{3}$$

Sort the data from small to large to get the maximum, minimum, difference-range, mode. The acceleration slope represents the slope of the line between two adjacent points and reflects the rate of change of acceleration. The calculation method is:

$$k = \frac{y_i - y_j}{i - j} \tag{4}$$

Among them, y_i , y_j respectively represent the acceleration value between two adjacent points. The mean value and variance of the calculated slope can be calculated respectively to obtain the mean value and variance of the acceleration slope.

3.2.2 Frequency Domain Features

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202101_7(1).0003

First, use the Fast Fourier Algorithm (FFT) to convert the time domain signal into thefrequency domain signal. The calculation method is:

$$X(k)=\sum_{n=0}^{N-1}x(n)W_N^{nk} \tag{5}$$

Among them $W_N=e^{-j\frac{2\pi}{N}}$.Direct Current (DC) is the first component after Fourier transform, which is the mean value of these signals.

Let C_i be the first amplitude value of a window, which means the number of sampligpoints in a window, The calculation methods of several values of the shape feature are:

a. Mean Value

$$u=\frac{1}{S}\sum_{i=1}^{N}iC_i$$

b. Standard deviation

c. Skewness $\sigma=\sqrt{\frac{1}{S}\sum_{i=1}^{N}(i-u)^2C_i}$

$$\gamma=\frac{1}{S}\sum_{i=1}^{N}(\frac{i-u}{\sigma})^3C_i$$

d. Kurtosis

$$\lambda=\frac{1}{S}\sum_{i=1}^{N}(\frac{i-u}{\sigma})^4C_i-3$$

In the same way, 4 characteristic values of amplitude mean value, amplitude standard deviation, amplitude skewness and amplitude kurtosis can be obtained. After sorting the amplitude values in ascending order, the first 5 peaks and the corresponding frequencies can be obtained, and the energy can be:

$$E=\sum_{i=1}^{N}|F_i|^2 \tag{6}$$

3.2.3 Time-frequency Domain Characteristics

Wave transform can effectively extract information from the signal, and perform multi-scale refinement analysis on functions or signals through operations such as scaling and translation, which solves the problem that Fourier transform cannot solve. The wavelet energy entropy reflects the energy distribution of the signal at different scales. Based on the above considerations, Hannon entropy, L norm, logarithmic energy entropy, and SURE entropy are the four characteristics of the time-frequency domain. To represent the projection coefficient of the signal on an orthogonal wavelet basis, it can be obtained by the following formula:

a. Shannon entropy

$$E(s)=-\sum_i s_i^2\log(s_i^2)$$

b. Norm

$$E(s)=\sum_i|s_i|^P$$

c. Log energy entropy

$$E(s)=\sum_i\log(s_i^2)$$

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202101_7(1).0003

d. SURE entropy

$$E(s) = n\sum_i \min(s_i^2, p^2)$$

After the above-mentioned feature extraction process, a total of 3382 samples were obtained, including 555 downstairs, 630 upstairs, 413 down, 231 sitting down, 350 jumping, walking 663, and running 540.

4. DETERMINE THE CLASSIFICATION MODEL

After the collected data is feature extracted, 27-dimensional features are obtained. As can be seen from the data distribution diagram in Figure 4, the original data roughly presents a Gaussian distribution, so the original data is normalized.

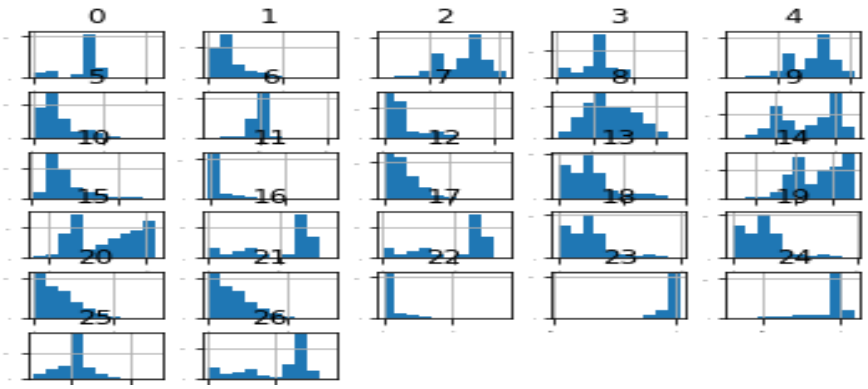


Figure 4. Data distribution histogram

Now use non-integrated algorithm and integrated algorithm for model training. The results obtained are shown in Table 1 and Table 2:

Table 1. Non-integrated algorithms

Model name	Score	Standard deviation
Logistic regression	0.769319	0.014846
Linear discriminant analysis	0.718292	0.018237
K neighbors	0.813299	0.018288
Decision tree	0.794830	0.015191
Support Vector Machines	0.793713	0.019642
Naive Bayes	0.623298	0.015191

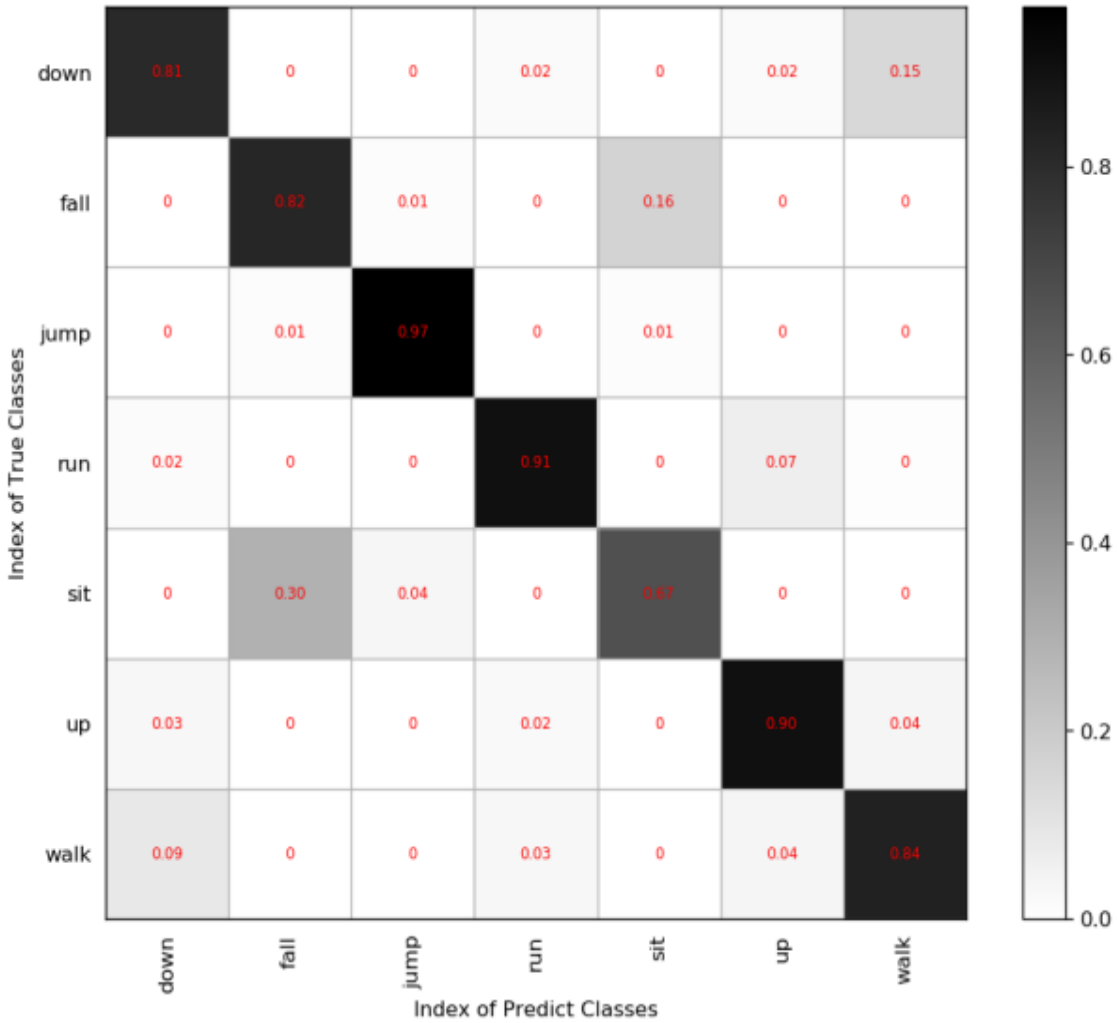
Table 2. Integration algorithm

Model name	Score	Standard deviation
Random forest	0.837349	0.014584
Extreme random tree	0.846944	0.014794
Lifting algorithm	0.390764	0.027937
Stochastic gradient boosting algorithm	0.847328	0.013237
Random forest	0.837349	0.014584

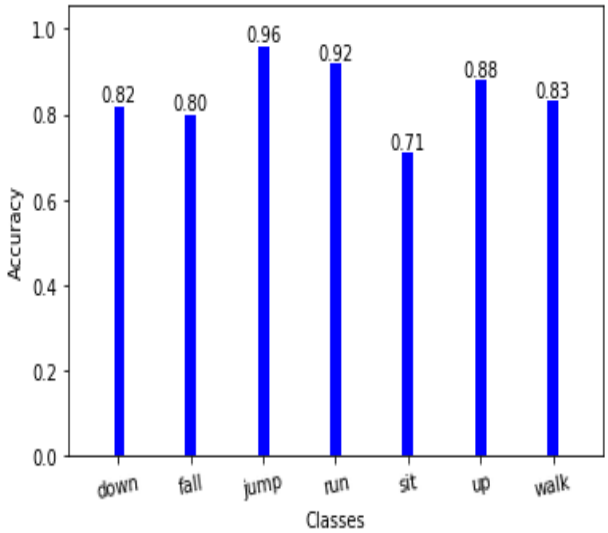
ISSN: 2472-3703 DOI: 10.6911/WSRJ.202101_7(1).0003

Comprehensively comparing the above results, the stochastic gradient boosting algorithm shows the highest performance and the smallest standard deviation, so the stochastic gradient boosting algorithm is used as the training model.

5. EXPERIMENTAL RESULTS AND ANALYSIS



(a)



(b)

Figure 5. (a)Confusion matrix; (b)Accuracy of recognition of different human behaviors

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202101_7(1).0003

The 3382 samples collected are divided into 2 parts: the first part is used to train the algorithm to generate the model, the second part is to predict the result through the model, and compare it with the known result, and use the 10-fold cross validation method to evaluate the accuracy of the algorithm model Here, 80% of the data is used as the training set, and 20% of the data is used as the test set. The confusion matrix and recognition accuracy are shown in Figure 5 respectively.

It can be seen from Table 3 that the recognition rate of jumping and running has reached more than 90%, the recognition rate of going up and down, falling, and walking has reached more than 80%, while the recognition rate of sitting down is only more than 70%, which is not very effective. ideal. The average correct rate of these seven behavior recognition has reached 85%, and a good recognition effect has been achieved.

Table 3. Classification report

Serial number	Behavioral state	Accuracy	Recall rate	F1 value
1	down	0.82	0.81	0.82
2	up	0.88	0.90	0.80
3	fall	0.80	0.82	0.81
4	sit	0.71	0.67	0.69
5	jump	0.96	0.97	0.96
6	walk	0.83	0.84	0.84
7	run	0.92	0.90	0.91
Mean		0.85	0.86	0.85

6. CONCLUSION

This paper uses the 27-dimensional data features of the time domain, frequency domain, and video domain to extract acceleration signals, and uses gradient boosting algorithm to establish a classification model to realize 7 daily human behaviors such as going up and down, walking, running, jumping, falling, and sitting. The recognition accuracy can reach 85%, which achieves a good recognition effect. However, due to the small number of selected features, the accuracy of recognition has not yet reached a particularly good effect, and more features can be considered in the extraction of model features in the future. In the selection of classification models, more advanced models can be considered.

REFERENCES

[1] Chen Lingquan. On the mental health problems of the elderly and their social nursing [D].2013.

[2] Lin Haibo, Li Yang, Zhang Yi, Luo Yuan. Recognition and Application of Human Movement Patterns Based on Time Series Analysis [D]. Chongqing University of Posts and Telecommunications, 2014.K. Kang, H. Li, J. Yan, X. Zeng, B. Yang, T. Xiao, C. Zhang,Z. Wang, R. Wang, X. Wang et al., "T-cnn: Tubelets with convolutional neural networks for object detection from videos," IEEE Transactions on Circuits and Systems for Video Technology, vol. 28, no. 10, pp. 2896–2907, 2018.

[3] Dimitriors S Alexiadis,Dimitrios Zarpalas.Real-Time,Full 3-D Recon-struction of Moving Foreground Objects From Multiple[J].IEEE Transactions on Multimedia,2013,15(2)

[4] Jennifer R K, Gary MW, Samuel AM. Activity Recogniti Using Cell Phone Accelerometers[J]. ACM SIGKDD Exploration Newsletter, 2010, 12(2): 74-82.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202101_7(1).0003

[5] Preece SJ, Goulermas JY, Kenney LPJ, et al. A Comparison of Feature Extraction Methods for the Classification ofDynamic Activities from Accelerometer Data[J]. IEEE Transactions on Biomedical Engineering, 2009, 56(3): 871-879.

[6] Liang Weijie, Zhang Yinghong, Jing Hui, Huang Bo, Zheng Ji. Research on fall detection method based on support vector machine [D]. Guilin University of Electronic Technology, 2014.

[7] Cai Yiheng, Wang Xueyan, Ma Jie, Kong Xinran. 3D human pose estimation based on random forest misclassification processing [D]. 2019

[8] Lu Xianling, Wang Hongbin, Wang Yingying, Xu Baoguo. A human fall recognition method based on acceleration sensor [D]. Jiangnan University, 2013.

[9] Song Lingling, Li Peng, Wang Ruifeng. Real-time human body state recognition algorithm based on acceleration sensor [D]. Yantai Vocational College, 2013.