

# Recognition of Multiple Human Body Postures Based on Six-axis Sensor

Wei Li\* and Dongxiu Ou

The Key Laboratory of Road and Traffic Engineering, Ministry of Education, China School of Transportation Engineering, Tongji University 4800 Cao'an Road, Shanghai, China

\*Corresponding author

**Abstract**—In the existing methods of recognition of multiple human body postures, recognition of human body postures based on wearable sensors has recently become a research hotspot because of its advantages such as simple information acquisition, low cost, and fast transmission. Based on the monitoring data collected by the six-axis sensor, this paper performs Kalman filtering on the data, and then selects a Gradient Boosting Decision Tree model from the classification algorithms in machine learning to classify and recognize various human behavior postures. And it could be benefit to key crowds such as the elderly.

**Keywords**—wearable sensors; recognition of human postures; decision tree

## I. INTRODUCTION

At present, the market for wearable devices such as wristbands and watches have experienced rapid development in 2016 and has entered the stage of change on demands. How to develop multiple applications based on wearable devices is the competition point of the wearable device market in the future. And the automatic recognition of human posture based on wearable devices will be a successful application mode.

Most existing wearable devices use a three-axis sensor to achieve the original data and sense the linear acceleration information in the three directions of X, Y, Z respectively. This article uses the latest six-axis sensor, which integrates the functions of a three-axis accelerometer and a three-axis gyroscope so that it can sense full angular acceleration information at the same time to better monitor motions.

The human motion gesture recognition mainly includes data collection, data processing, feature extraction and classifier design, etc., thereby extending different research hotspots. The specific flow is: the data collected by the sensor transmits the acquired acceleration signal to the PC or the mobile device through wireless transmission or other transmission methods. The PC or mobile device obtains and stores the received data through software of the host computer, while maintaining continuous communication with the collector. After the stored data is preprocessed, feature extraction is performed according to the processing results and a classification algorithm is adopted for gesture recognition.

## II. RELATED WORKS

There are many methods of recognition of multiple human body postures obtained in the existing researches, which are mainly divided into two types.

One is the detection technology of human posture that fixes the external detection equipment, such as Bao Nan's distinction of falling, fall-like actions and daily actions based on the Kinect infrared sensor (1). However, this requires the installation of an external detection device, which has low operability and has no user specificity.

The other type is a human detection technology based on data collected by a wearable device. For example, Shi Xin et al. differentiates a fall and a walking posture using a support vector machine algorithm based on changes in pressure data that collected by a pressure sensor placed on an instep (2). Based on the triaxial acceleration, Ren Zhiling et al. established thresholds to distinguish standing, walking, and running behaviors using SVM(Support-Vector-Machine) algorithm (3). Zhang Junjian used SVM and the time average of acceleration vector to establish thresholds to distinguish between falling, going downstairs, walking, and Jogging behavior (4). Li Dong et al. differentiated the falling, walking, and sitting in different postures by using the acceleration amplitude combined with the acceleration and the angle of each axis angle (5). Ren Ziliang combined the data collected by acceleration sensor, magnetic sensor and pressure sensor on the basis of a single sensor and used the threshold and voting algorithm to identify the postures of falling, walking, and sitting down (6).

The main problems of the above studies are reflected in the following. First, the sensor data source is a single source. Only three-axis sensor data (X, Y, and Z-axis acceleration data) or pressure sensors are used to extract feature quantities that are confined to simple acceleration thresholds. Features cannot accurately reflect the characteristics of the action. Second, most detection methods only distinguish between a few postures such as falling and walking, and only a single action is detected and the convergence between normal operations did not been taken into account. Third, the wearable devices used in researches are often placed in the waist, chest, feet, which is inconsistent with the current wearable devices such as smart bracelets commonly used in every day, and the application environment is divorced from reality.

### III. DATA COLLECTING AND PROCESSING

This article selects the six-axis sensor (JY901). The measurement dimensions include three-dimensional acceleration, three-dimensional angular acceleration and three-dimensional magnetic field. It can fully meet basic data acquisition requirements.

In order to obtain raw data that is closer to reality, data collection is divided into two phases. Experimental individuals with different heights and weights are performed in stages according to their personal habits. The first stage: acquire the data of the fixed movements: The scene is indoors, and the basic posture of the human body includes squatting, standing, walking, running, and sitting down. The main purpose of this stage of the experiment is to identify the feature values that can be used after the initial recognition and posture recognition. The second stage: acquire the data of continuous motions: The scene is indoors. The postures include five kinds of postures: walking, going up and down the elevator, descending the stairs, going up stairs and running.

Firstly, Kalman filter is used to process the collected raw data. In the host computer, then six data curves with less interference can be obtained. The Figure I are screenshots of the part of the host computer that captures some of the fixed motions. Kalman filtering is a mature method of application, and the specific process is not described here.

The characteristics of acceleration changes and angular acceleration data of a single action are all obvious, and the acceleration characteristics of each gesture also do not have similar situations. Therefore, a single action is easily

thresholded based on the data characteristics of the acquired data, and the accuracy of discrimination is high.

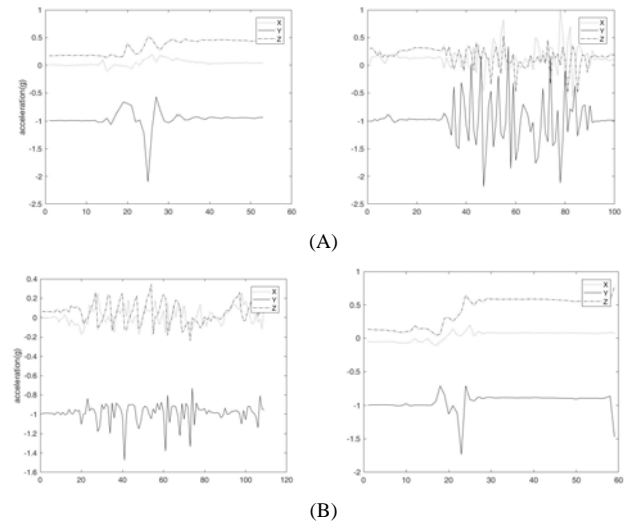


FIGURE I. (A). LINEAR ACCELERATION CURVE OF SQUATTING AND RUNNING; (B). LINEAR ACCELERATION CURVE OF WALKING AND SITTING

After the data of a single action is acquired, the test data of the continuous actions is collected immediately. The continuous actions includes five types of postures: walking, going up and down the elevator, going down the stairs, going up the stairs and running. The data curves displayed by the host computer for continuous operation are shown in the Figure II below (taking a series of continuous actions as an example).

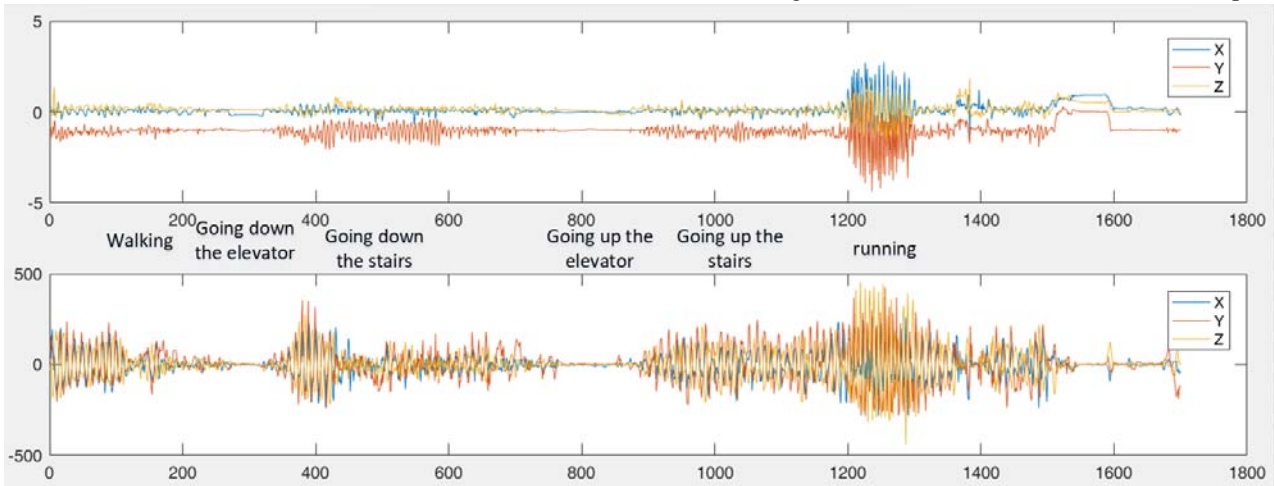


FIGURE II. A SET OF CONTINUOUS ACTIONS

The characteristics of the data curves presented by continuous motions are also obvious. However, how to build sample data and extract characteristics of samples in discrete data points is the focus of this study. It is known that each set of data includes six types of data: linear acceleration in the X, Y, and Z axis directions, and angular acceleration in the X, Y, and Z axis directions. Three features are extracted for each type of data: variance of the data, the maximum value and the

minimum value. In this way, each sample owns  $3 \times 6 = 18$  features.

If all the data within a time span of 50 data points is taken as a sample, that is, a set of  $6 \times 50$  raw data points constitutes a sample of raw data, and 10 samples of each posture are to be extracted, which will lead to the low sampling interval. So the variances, maximum and minimum values in multiple samples will duplicate, which is not benefit to the learning model.

Therefore, all data within a time span of 15 data points is taken as a sample, that is, a set of 6\*15 raw data points constitute a sample of raw data. 15 consecutive data points lead that the coincidence rate of the raw data between the samples is 1/3, and the amount of data is not too small to reflect the action characteristics of a single gesture. Then use the six kinds of features mentioned above, that is, the size of one sample is 6\*15=90 original data values, and each of the 15 values is further extracted 3 data features, and there are 18 feature values in total. Therefore, a sample contains 3 \* 6 = 18 characteristic values. The sample interval takes 7 to 15 unequal data points. In the same manner, 10 samples are taken from each pose to form a training sample set containing 50 samples for training the model.

Then, 6\*15 consecutive data are randomly extracted from the original data set, and a 3\*6 sample is constructed as a test sample according to the method above. After 10 times of extraction, a group of test samples containing 10 test samples is generated to verify accuracy of the model.

#### IV. ALGORITHM AND VALIDATION

##### A. Algorithm

The decision tree model constructed in this paper adopts the Gradient Boosting Decision Tree (GBDT), also known as the Multiple Additive Regression Tree (MART). It is an iterative decision tree algorithm. The algorithm consists of multiple decision trees. The conclusions of all the trees will add up together to make the final answer. Different from the traditional Boosting algorithm (such as AdaBoosting), the basic classifier of the GBDT algorithm is a regression tree. The purpose of the iteration is to improve the model by calculating the negative gradient of the last model, and then building a new decision tree in the direction of the decreasing gradient of the residual error.

Given a training sample  $\{(x_i, y_i)\} (i = 1, 2, \dots, n)$ . Since recognition of postures is a classification process, a logarithmic loss function is used, which is:

$$L(y, F(x)) = 2 \sum_{i=1}^q \log(1 + \exp(-2y_i p_i)) \quad (1)$$

$x_i = (x_{1i}, x_{2i}, \dots, x_{qi})$ ,  $n$  is the number of samples,  $q$  is the number of eigenvalues in the sample,  $y_i$  is the actual sample label and  $p_i$  is the prediction label of the sample. The detailed steps of the GBDT algorithm are as follows:

1) Initialize the model and estimate the constant value  $\beta$  that minimizes the loss function

$$F_0(x) = \arg \min \sum_{i=1}^n L(y_i, \beta) \quad (2)$$

2) Build the model in the direction of the gradient descent of the last model loss function, from  $m = 1$  to  $M$  ( $M$  is the number of iterations):

a) Calculate the negative gradient of the loss function in the current model and use it as an estimate of the residual.

$$r_{im} = -\left[\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)}\right]_{F_m(x)=F_{m-1}(x)}; i = 1, 2, \dots, n \quad (3)$$

b) Take the estimated residual obtained in a as input and fit a regression tree and find the leaf node region  $R_{j,m} (j = 1, 2, \dots, J)$  of the regression tree.

c) To minimize the loss function, find the optimal step size in the descending direction  $\beta_{j,m}$ :

$$\beta_{j,m} = \arg \min_{\beta} \sum_{x \in R_{j,m}} L(y_i, F_{m-1}(x) + \beta) \quad (4)$$

d) Update the model  $F_m(x)$ :

$$F_m(x) = F_{m-1}(x) + \sum_{j=1}^J \beta_{j,m} I; x \in R_{j,m} \quad (5)$$

3) End the process of iteration and then get the model  $F_M(x)$ :

$$F_M(x) = \sum_{m=1}^M \sum_{j=1}^J \beta_{j,m} I; x \in R_{j,m} \quad (6)$$

4) According to the obtained model, estimate the probability  $p_+(x)$  that the sample belongs to positive class, and the probability  $p_-(x)$  that the sample belongs to negative class:

$$p_+(x) = p_r(y = 1|x) = \frac{1}{1 + e^{-2F_M(x)}} \quad (7)$$

$$p_-(x) = p_r(y = -1|x) = \frac{1}{1 + e^{2F_M(x)}} \quad (8)$$

5) Predict the label  $y(x)$  of the sample label according to the following criteria, where  $c(-1,1)$  is the price function, which represents the cost when the real category is 1 and the prediction category is -1:

$$y(x) = 2 * \lfloor \{c(-1,1)p_+(x) > c(1,-1)p_-(x)\} - 1 \rfloor \quad (9)$$

Among them,  $\lfloor \cdot \rfloor$  is the function that converts Boolean values to  $\{0,1\}$ .

##### B. Validation

First, we use all 18 eigenvalues of the sample as input to the model to train a gradient decision tree, then randomly select

test samples from three consecutive sets of actions. Each test samples set contains ten random samples and each set of continuous actions contains four sets of test samples. The accuracy after using the test set are shown in the following table:

TABLE I. THE ACCURACY WHEN USING 18 EIGENVALUES

ACCURACY	<i>The first set of test samples</i>	<i>The second set of test samples</i>	<i>The third set of test samples</i>	<i>The fourth set of test samples</i>
From first set of continuous actions	0.6	0.6	0.7	0.6
From second set of actions	0.5	0.9	0.4	0.9
From third set of actions	0.7	0.9	0.5	0.5

From the results above, if all 18 eigenvalues are used, the accuracy of the model is not high. Considering that the information contained in the eigenvalues of the x, y, and z-axis overlaps with the information contained in the x, y, and z-axis acceleration, which is easy to interfere with the classification. From the data analysis and the weights of the decision tree, it can also be seen that the linear acceleration and angular acceleration in the y-axis direction play the largest role in the classification of the postures. Therefore, the 9 eigenvalues of the x, y, and z-axis accelerations and the 3 eigenvalues of y-axis angles are selected to be the input of the model, and the test results under the twelve eigenvalues are shown in the following table, and the classification accuracy has been greatly improved.

TABLE II. THE ACCURACY WHEN USING 12 EIGENVALUES

ACCURACY	<i>The first set of test samples</i>	<i>The second set of test samples</i>	<i>The third set of test samples</i>	<i>The fourth set of test samples</i>
From first set of continuous actions	0.9	0.8	0.7	0.9
From second set of actions	0.9	0.9	1.0	0.9
From third set of actions	0.7	0.9	0.9	0.8

In the analysis of test samples failed to classify in recognition continuous motions, it was found that there were other interference actions in a series of continuous actions, such as raising hands suddenly, waving elbows, etc. In the classification model, the abnormal actions are not possible to match any type of posture so that misjudgment occurs. Moreover, in the selection of training samples, misjudgment can also occur when the data at the joints of the two types of gestures are selected. These two types of misjudgments interfere with recognition of postures.

Therefore, this paper proposes the following judgement logic: the recognition result of one sample is used as a small identification unit, and the final result depends on recognition result of 4 consecutive samples, namely 60 data points. That is,

a result that aggregates recognition result of four small units is called a large unit. The specific logic is as follows:

1. When the results of the four identified small units are the same, the result is the final result;

2. When the results of the four small units are inconsistent, the recognition results with the highest number of occurrences is selected as the final result; when the situation with 2:2 occurs, the results of the first two small units are selected as the final result.

Under the recognition logic, the output result can be synthesized through voting when identifying multiple types of continuous actions, which can avoid misjudgment caused by abnormal samples and increase the accuracy of the recognition algorithm. The recognition results for the first set of consecutive actions using this recognition logic shows that only 3 of the 37 actions were recognition errors compared with the actual situation and the miscarriage rate is less than 10%.

## V. CONCLUSION

Based on the existing researches on recognition of human postures, this paper uses a six-axis sensor to collect the raw data and the wearing position for the sensor is on the wrist, which is in line with the actual environment. The integration of six-axis sensors can also improve the performance of existing wearable devices. The various data features extracted in this paper can cover some characteristics human postures. The Gradient Boosting Decision Tree used in this paper belongs to a simpler algorithm and has high real-time performance. It has high accuracy from the test results and can basically meet the needs of differentiating daily movement gestures.

However, in terms of application scenarios, this article only considers an indoor scene, the types of gestures involved are few, and the accuracy of the algorithm also has room for optimization and improvement. Therefore, in the future work, the data under more scenarios should be collected for analysis and the algorithm should be optimized.

## ACKNOWLEDGMENT

This work is supported by the Research Program of Shanghai Science and Technology Committee (16511104902) and the Research Program of Shanghai Science and Technology Committee(18DZ2202600). The authors are grateful to the reviewer for helpful comments and suggestions on initial drafts.

## REFERENCES

- [1] Bao Nan, Jiang Bowen, Li Yongze, et al. Research on Fall Detection Based on Kinect Sensor[J]. *Electronic Design Engineering*, 2017, 25(12): 149-152.
- [2] Shi Xin, Xiong Qingyu, Lei Yining. Research on Fall Detection System Based on Pressure Sensor[J]. *Journal of Instrumentation*, 2010, 31(3): 715-720.
- [3] Ren Zhiling, Zhang Bingqian, Zheng Liyuan. Fall Detection and Alarm System Design Based on Acceleration Sensor[J]. *Computer Measurement & Control*, 2013, 21(6):1428-1430.
- [4] Zhang Junjian, Zhao Jie, An Xiaojing, et al. Research on Fall Detection Based on Triaxial Acceleration Sensors[J]. *Progress in Biomedicine*, 2014, 14(18):3585-3588.

- [5] Li Dong, Liang Shan. Design of Fall Detection Device for Elderly Based on Acceleration Sensor[J]. *Transducer and Microsystem*, 2008, 27(9):85-88.
- [6] Ren Ziliang, Li Yong. Wireless multi-sensor fall monitoring system based on secondary judgment[J]. *Computer Engineering and Design*, 2012, 33(5): 1785-1789.