Team Number :	85306

Problem Chosen :	A

2018 APMCM summary sheet

Real-time training model for elderly people balance ability

Abstract

Along with the decline of physical function, elderly people are more likely to fall than ordinary people. Once an accidental fall happens, it will has a tremendous impact on the quality of life of the elderly in their later years. Through the analysis of the row data given by the title, our group established a comprehensive body balance assessment model for the elderly and provide feasible suggestions accordingly. This process is of great help to the society to better understand the action ability of the elderly group and then take preventive measures against falls.

With regard to question 1, firstly, we extract realistic step, center of gravity and motion features by integrating and interacting the Annex 2 raw data filtered by Kalman algorithm in advance. Then we use these three characteristics to classify the subjects blindly. In the process of classification, we use K-means clustering analysis method. For the aggregated three groups, we use BP neural network to detect the raw monitoring point data reversely. Through this step, we obtain the weight of each detector's data on the impact of the three classified groups. Combined with relevant information and practical experience, we select the detector with high weight after arrangement for purposeful combination. Thus, we achieve the goal of extracting 25 balancing ability features.

Since each body balance feature is composed of multiple detector data, it has a considerable amount of information.When it comes to question 2, we directly use the 25 body balance features obtained from the previous question to establish a body balance assessment system.Through statistical analysis, we know that the body balance assessment system meets all the preconditions of factor analysis. Therefore, we use the principal component analysis method to get the principal component factors of the assessment system of individual balance characteristics and the value of risk assessment. At the same time, combined with factor analysis, we find that the two non-strong related factors, muscle strength and vestibular balance, have a great impact on the balance ability of the elderly. In this regard, we put forward reasonable suggestions for the elderly, such as taking part in more balancing exercises, strengthening muscle exercises and improving body shape.

For question 3, we test the actual data of the elderly in Annex 1 with the results of our assessment system. The simulation proves that our model works well. In the statistical process, the characteristics of the elderly with poor balance force are analyzed and reasonable suggestions are given accordingly, such as attitudes towards disease, companionship of relatives and friends, walking advice and so on.

Key words: Attention To The Elderly; Kalman Filter; K-Means Clustering Algorithm; BP Neural Network; Balance Ability Assessment Model; Factor Analysis

Contents

I. Overview

In order to indicate the origin of balanced assessment of the elderly, the following background is worth mentioning.

1.1 Background

Falling is one of the most important risks that can infect health to elderly people. Researches find out that among those are above 65 years old, every year 28-35% elderly people once fall [1]. However, among those are above 70 years old, the percentage rise to 32-42% [2]. Falling elderly people often lead to trauma, functional decline, decreased self-confidence, reduced social participation, and even endanger the lives of the elderly, which is an important cause of threatening the physical and mental health of the elderly. And to the whole society, falling makes fewer active people and more health care expenditures which leads to the waste of social resource meanwhile. [3-6]Thus, preventing elderly from falling is one of the methods to dealing aging population.

In developing countries, there has been little research on the fall of older people because they have not yet or have just entered an aging society [7]. In order to make the rapidly growing elderly population not only have a long life span, but also maintain a good health condition, the whole society should pay sufficient attention to the problems that may affect the health of the elderly, including the fall problem. Consequently, it is of great realistic importance to make a balance ability assessment for elderly people with a view to assisting them in mobility status, correcting postures and preventing accidental falls.

1.2 Restatement of the Problems

The balance ability of the human body refers to a stable state of the human body, that is, the ability to return to the original stable state and maintain it regardless of the position, movement or external force. As a means of detecting physical fitness, it has gradually become one of the important indicators for measuring physical fitness. Our body can not be separable from the maintenance of body balance during any exercise. If all the elderly can keep self balance well, the occurrence of falls will reduce in a sharp way. However, with the increase of age, the human body undergoes a series of degenerative changes in morphological structure and physiological functions, and the balance ability will also decrease, resulting in the occurrence of frequent falls in the elderly. Therefore, using the data in the attachments to build an efficient assessment model of elderly people balance ability can not only improve their awareness of self-protection, but also raise the society concern of the daily behavior of the elderly who are prone to fall.

Based on the actual behavior data given to the elderly, we need to accomplish the following three questions:

Question1: Analyze the balance features of elderly people based on the data given and establish a feature extraction model based on an analysis of steps, the center of gravity and motion.

Question2: Build a balance risk assessment system and give advice accordingly.

Question3: Make an analog computation and a comparative analysis of the body balance force based on thec provided and give effectual advice to elderly people with weak balance ability.

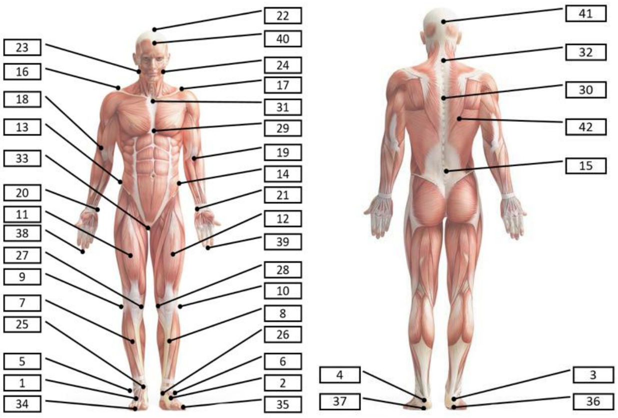


Figure 1: The layout of the monitoring points

1.3 Analysis of the Problems

The attachment contains real-time detection data from a random sampling test by deploying 42 monitoring points on the body of 76 the elderly subjects.The data table describes the real-time three-dimensional data of each subject in free walk state.The data table describes the real-time three-dimensional data of the elderly in the normal walking state, namely, the horizontal data (x-axis), the longitudinal data (y-axis), and the height data (z-axis) relative to the detection zero at each time point.We are required to use the real-time data in the appendix to establish a practical model for assessing the balance ability of the elderly.

Due to the difference of their body states, the track of gravity center, walking steps and motion characteristics of the elderly have obvious separability. Firstly, the the track of gravity center, walking steps and motion characteristics are obtained by integrating the real-time detection data of the elderly, and then the mobility of the elderly is classified as a feature extraction model according to these three characteristics.The original detection data are analyzed again by the feature extraction model of elderly people balance ability, so as to obtain the detection positions which have a greater impact on the balance of action, and transform the combination of these monitoring points into twenty-five behavioral characteristics in real life. Finally, using these behavioral characteristics, a comprehensive assessment model of the balance ability of the elderly is established.

- Analysis of problem 1

  Step 1 Analyze and process the original data. We use the proportional coefficient multiplication coefficient method to analyze and obtain the trajectory characteristics of the human body's center of gravity. Then, the steps characteristics of the test individuals are mainly extracted by the motion trajectory of the leg detector. Next, by analyzing the correlation information of the original data, the velocity, acceleration and angle of the specific part are calculated to obtain the dynamic characteristics of the individual.

Step 2 Use unsupervised learning to extract these three features from the original data. After learning, 76 elderly people were clustered into three categories to generate feature extraction model.

Step 3 Use the feature extraction model to learn the weights of the original data set, under BP network, and the influence proportion of different monitoring points under this model is obtained. In this paper, the array and combination of monitoring points with high impact ratio are transformed into twenty-five behavioral characteristics with practical application significance, so as to establish an old people's balance ability evaluation model with universal application value.

- Analysis of problem 2

We standardized the twenty-five behavioral characteristics extracted from the first question into one-dimensional data based on real life experience, and process them with factor analysis, so as to establish an old people's balance ability assessment model with universal application value.

In the process of factor Analysis, we can further mine the correlation information of the characteristics of balance ability measurement of the elderly through data analysis, and then put forward behavioral suggestions for the elderly to better maintain a certain balance ability.

- Analysis of problem 3

We quantify the real behavioral ability information and falls of the elderly in Annex 1, and compare it with the established balance ability assessment model. In the process of model evaluation criteria, the behavioral characteristics of the elderly who are apt to fall are emphatically analyzed. It should be popularized by means of promotion and propaganda, so as to raise the awareness and means of self-prevention of this group of people, and better arouse the attention of this group in need of great attention.

1.4 General Assumptions

We make the following assumptions to complete our model through this paper.

- Assume that the subjects are tested under the condition of free walking.
- Consider only the effects of steps, center of gravity and motion, ignoring the effects of other factors on balance ability.
- Assume that the data given in the title are true and reliable.
- Assume that the locus of the monitoring points of the subjects is regular.
- Ignoring random errors, all the errors mentioned in this paper are systematic errors.

1.5 Notions and Symbol Description

We will define the following variables here as they are used throughout our paper.

Table 1: Notions and symbol

Symbols	Definitions	Symbols	Definitions
η	step size	λ	eigenvalue
I	loss function	A	factor load matrix,
w	weight	a_{ij}	actor load

$E_q(t)$	corresponding objective function	Z(t)	total objective function
E (x)	expectation	D (x)	variance
ε	special factor	F	common factor
μ	corresponding eigenvectors	C	cumulative variance contribution rate
$P_{something}$	position in three-dimensional space	f(x)	auchy distribution membership functions

II. Model Establishment and Solution of Problem 1

2.1 Problem 1 Solving Process

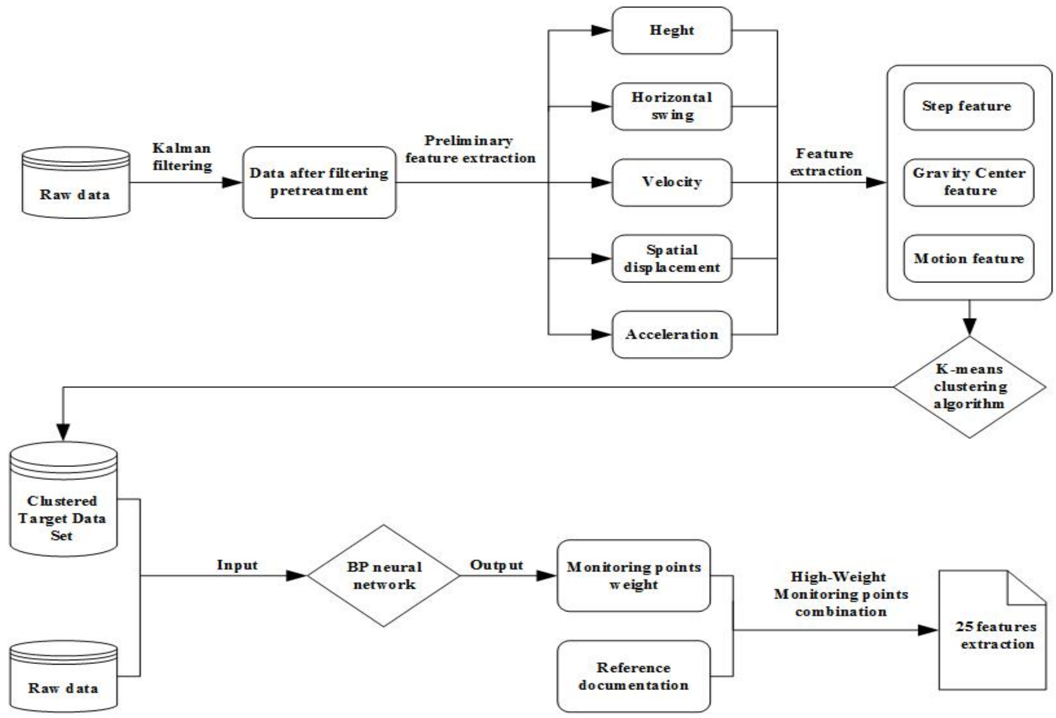


Figure 2: Flow chart of Problem 1

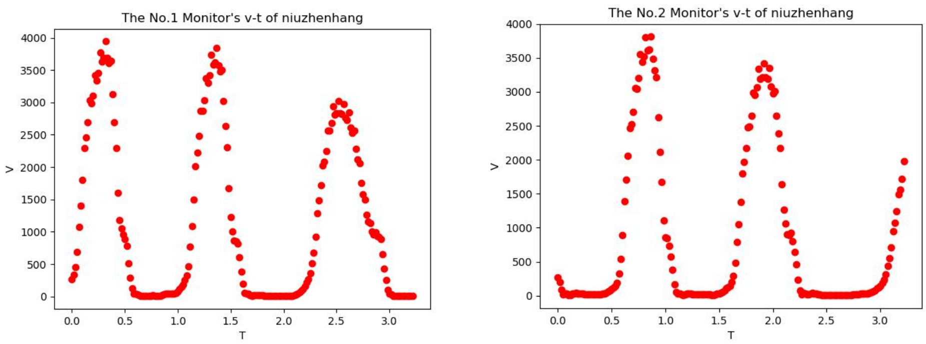
Fig.2 shows the whole brief process of how our group solve the Problem 1.

2.2 Preliminary Feature Extraction from Raw Data

First, we use the SQL server software to import the data of Annex 2 into the database and establish the data set of the subjects. Then Kalman filtering algorithm is used to pre-process the data set to filter the data with large missing values and low importance of attributes. Then, based on step, center of gravity and motion, the data of subjects are randomly selected as samples for analysis. Then 42 monitoring points' data corresponding to each sample are plotted in batches using Python language and MATLAB software.

2.2.1 Steps Feature

In this section, our goal is to extract the steps feature. From the large amount of raw data captured by the monitoring points, we make targeted selection combining with the existing life experience. Then the step features are extracted from the filtered data for further aggregation analysis. When discussing human step features, we mainly analyze the following parts of the human hip, while the monitoring points distributed below the human hip are monitoring point 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 25, 25, 27, 28, 34, 35, 36, 37. We use the images of the velocity of these monitoring points varying with time and the projection of their trajectories in the horizontal direction for further analysis.



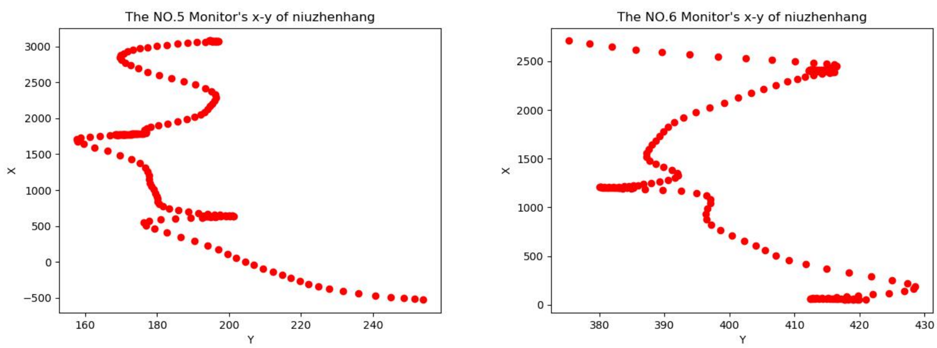
(a)Captured by No.1 Monitor (b)Captured by No.2 Monitor

Figure 3: V-T Correlation Graphs of niuzhenhang

Taking raw data from niuzhenhang.trc as an example, the V-T correlation graphs of monitoring point 1 and monitoring point 2 below the hip were analyzed.

It is known that monitoring point 1 and monitoring point 2 are located on the left and right feet of the subjects. By comparing the two graphs, it can be found that when the speed of monitoring point 1 is not zero, the speed of monitoring point 2 approaches zero, and when the speed of monitoring point 1 approaches zero, the speed of monitoring point 2 is not zero. So we can conclude that the subjects'left and right feet can move normally.

Next, we select the X-Y displacement Graphs of monitoring point 5 and 6 for further analysis.

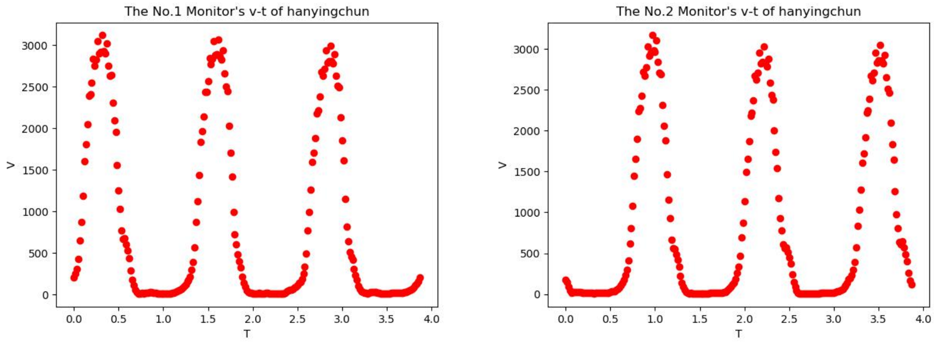


(a)Captured by No.5 Monitor (b)Captured by No.6 Monitor

Figure 4: X-Y displacement Graphs of niuzhenhang

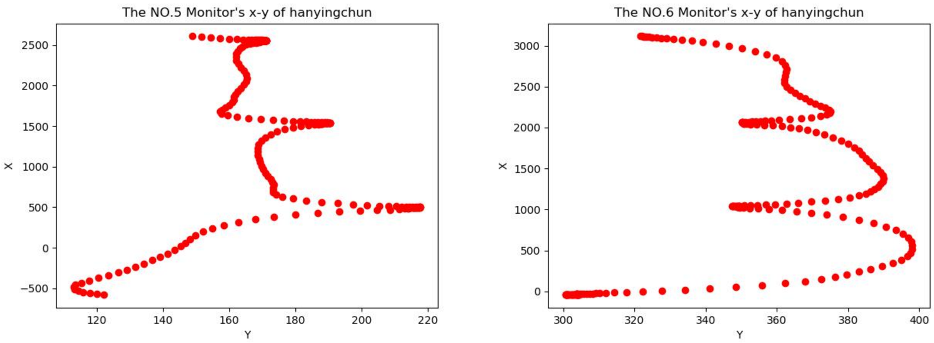
The dense point set in the graph indicates that the moving distance is shorter in the time interval, while the sparse point set indicates that the moving distance is longer.As can be viewed from Fig. 4, when the subjects walked alternately with their left and right feet, the subjects tended to walk in roughly one direction. Therefore, we can roughly judge that the whole body tends to move in a certain direction and the experimenter can keep his body dynamically balanced.

Then, we do the same analysis of hanyingchun's raw data.



(a)Captured by No.1 Monitor (b)Captured by No.2 Monitor

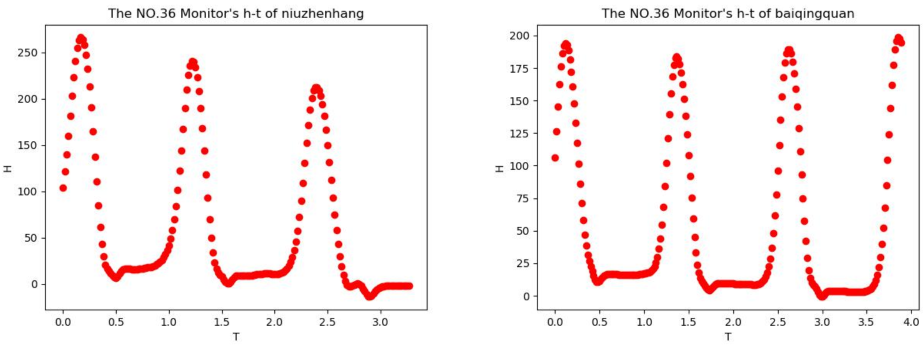
Figure 5: V-T Correlation Graphs of hanyingchun



(a)Captured by No.5 Monitor (b)Captured by No.6 Monitor

Figure 6: X-Y displacement Graphs of hanyingchun

From Fig. 5, based on previous inferences, we can analyze that the subjects'left and right feet can move normally. However, from Fig. 6, we can see that when the subjects walk alternately with their left and right feet, the landing point of the subject's left and right feet sometimes deviates to the right and sometimes to the left. Therefore, we can conclude that the subject's body was unbalanced while walking.



(a) Niuzhenhang (b)baiqingquan

Figure 7:H-T Correlation Graphs of niuzhenhang and baiqingquan

Finally, we process the data obtained by the 36 monitoring point of different individuals, and then analyze it according to H-T correlation.From Fig.7, the maximum height of niuzhenhang is obviously larger than that of baiqingquan. But the maximum height distribution of baiqingquan in each action cycle is more balanced than that of niuzhenhang.

After preliminary processing and analysis of the data obtained by the monitoring point, we can easily get the speed of human walking, the position of left and right feet and the height of the legs raised when walking. Our steps feature is mainly composed of these characteristics.

2.2.2 Center-of-gravity Feature

In this section, our goal is to extract the feature data of the center of gravity. The detector data related to the center of gravity are processed by multiplying coefficient method to obtain the feature data of the center of gravity. The analysis of the change of the center of gravity of the subjects in different frames is helpful to the analysis of the balance of the subjects.

By looking up the relevant information, we know that the position of the center of gravity of the human body changes with the movement of the human body. Firstly, the original data of Cuizhenhua is taken as a sample, and the data information of his 42 monitoring points is drawn into a three-dimensional moving image, as shown in Fig. 8.

Figure 8:3-D moving image of 42 detection positions from Cuizhenhua

Because there are 42 moving pictures, and each image corresponds to different positions of the body, it is difficult to find out the changing regulation of the human body's center of gravity from these 42 pictures. Therefore, we select the main joint nodes from the body.

According to the influence of these joint nodes on the center of gravity of the human body, we multiply the coordinate values of the center of each joint of the body by the corresponding proportional coefficient, and the sum of the product is the coordinate values of the total center of gravity of the human body. The formula is as follows:

$$
\begin{aligned}
P_{gravity} &= 0.0154 * P_{foot} + 0.0647 * P_{ankle} + 0.1630\, P_{knee} \\
&\quad + 0.0192 * P_{wrist} + 0.0168 * P_{hand} + 0.0706 * P_{head} \\
&\quad + 0.3176 * P_{hip} + 0.2747 * P_{shoulder} + 0.0580 * P_{cubital}
\end{aligned} \tag{1}
$$

$P_{something}$ represents the coordinates of a part of the body in three-dimensional space.

From the sketch given by the title, the monitoring points corresponding to each part from the formula are:

- Foot Detector: 34, 35
- Ankle Detector: 5, 6, 25, 26
- Knee Detector: 9, 10, 27, 28
- Hip Detector: 13, 14
- Shoulder Detector: 16, 17
- Elbow Detector: 18, 19
- Wrist Detector: 20, 21
- Hand Detector: 38, 39
- Head detector: 22

The spatial coordinates X, Y and Z of the center of gravity are obtained by processing cuizhenhua's data with multiplication coefficient method, as shown in the Tab. 2 below:

Table 2: Spatial position of center of gravity

Time(s)	Coordinate X	Coordinate Y	Coordinate Z
0	-231.482	541.8904	944.3832
0.017	-186.57	546.8961	942.0415
0.033	-141.422	552.0665	940.2291
0.05	-95.9752	557.4332	939.4749
0.067	-50.4131	562.8399	940.1686
0.083	-5.07546	568.0015	942.3987
……	……	……	……

We plot the data in Table 1 into three-dimensional images. The drawn image is shown in Fig. 9:

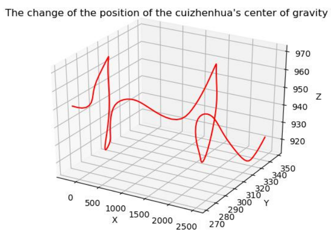


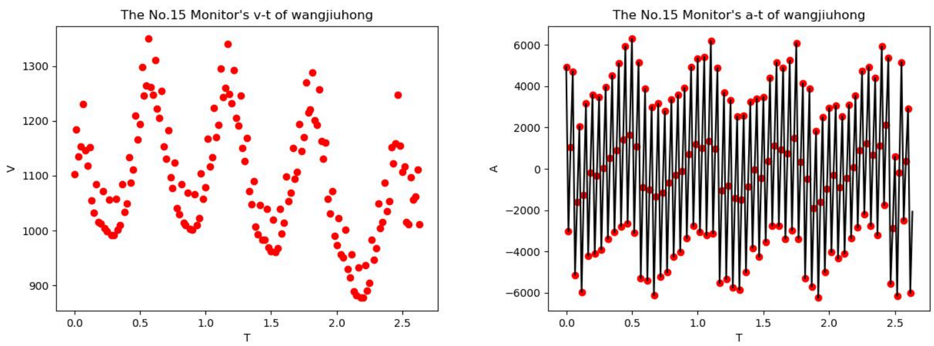
Figure 9: The Position Map of the Center of Gravity Analyzed by Cuizhenhua's data

Using the center of gravity position map, we can intuitively see the changing trend and regularity of the subject's center of gravity position. From the analysis of Figure 8, it can be concluded that the difference of center of gravity height change is less than 50 and the difference of left and right swing is less than 80 during the whole experiment.The extraction of center-of-gravity features is helpful to the next step of data analysis.

2.2.3 Motion Feature

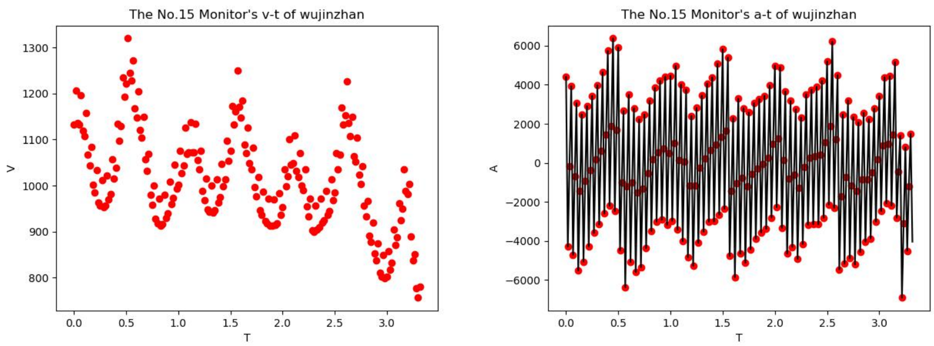
In this section, our goal is to extract dynamic feature. In the analysis of human motion, we mainly consider the influence of human walking speed and acceleration, arm swing and so on. Firstly, we use the speed feature obtained in the first section to calculate the acceleration feature, and analyze the influence of the speed acceleration of human walking on human motion. Secondly, we select the monitoring points installed on the arm and use the horizontal projection image of its trajectory to analyze the influence of arm swing on human motion.

Take the data of wangjiuhong and wujinzhan for example. First, the velocity and acceleration are obtained by processing the two sets of raw data.Then we picture the velocity and acceleration data for preliminary analysis.Since monitoring point 15 is approximately the same as the position of the gravity-center when the human body moves, we select the velocity and acceleration changes of monitoring point 15 to analyze.



(a) V-T Correlation (b)A-T Correlationr

Figure 10: V-T, A-T Correlation Graphs of wangjiuhong

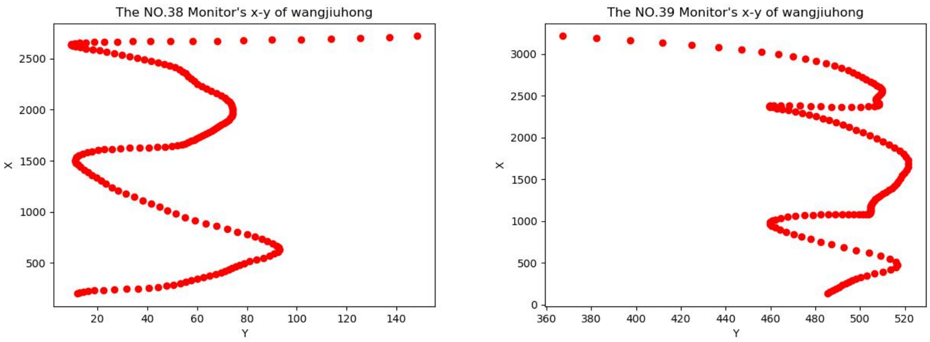


(a)V-T Correlation (b)A-T Correlationr

Figure 11: V-T, A-T Correlation Graphs of wujinzhan

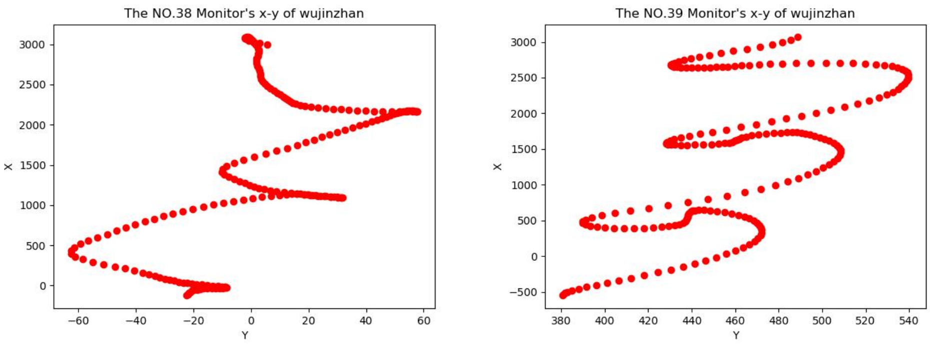
By analyzing drawn V-T, A-T correlation graphs, we find that the mean value of velocity of wangjiuhong is larger than that of wujinzhan. In addition, the change rate of Wangjiuhong's A-T data is gentler than that of wujinzhan's A-T data. It can be concluded that the equilibrium state of wangjiuhong is stronger than that of wujinzhan.

Because monitoring point 38 and monitoring point 39 are located at the end of the arm of the subject, we select the X-Y displacement graphs of these two monitoring points for arm swing analysis.



(a)No.38 Monitor Point (b)No.39 Monitor Point

Figure 12: Swing Trajectory Graphs of wangjiuhong



(a)No.38 Monitor Point (b)No.39 Monitor Point

Figure 13: Swing Trajectory Graphs of wujinzhan

From the analysis of the four pictures above, the swing of wangjiuhong's left and right arms is basically symmetrical, while that of wujinzhan's left and right arms is not so symmetrical. Further analysis shows that the symmetry of wangjiuhong's arm swing can make his body maintain a relatively good balance, while the asymmetry of wujinzhan's arm swing makes his balance worse. It is concluded that the more symmetrical the swing information of the left and right arms is, the better the balance of the subjects is, and vice versa.

Based on the above inferences, the mean value of velocity, the severity of acceleration change and the symmetry of left and right arm swing are used to extract Motion features.

2.3 K-means Clustering Algorithm in Unsupervised Learning

2.3.1 Unsupervised Learning

Unsupervised models are trained using data that consists only of input vectors, with no specific target output in mind. Instead of telling an unsupervised algorithm what it should be looking for in the data, the algorithm does the work itself, in a sense independently finding structure within the data.

2.3.2 K-Means Clustering

K-Means clustering is an unsupervised learning algorithm that, as the name hints, finds a fixed number (k) of clusters in a set of data. A cluster is a group of data points that are grouped together due to similarities in their features. When using a K-Means algorithm, a cluster is defined by a centroid, which is a point (either imaginary or real) at the center of a cluster. Every point in a data set is part of the cluster whose centroid is most closely located. To put it simply, K-Means finds k number of centroids, and then assigns all data points to the closest cluster, with the aim of keeping the centroids small.

K-Means starts by randomly defining K centroids. From there, it works in iterative steps to perform two tasks:

Step1: Assign each data point to the closest corresponding centroid, using the standard Euclidean distance.

Step2: For each centroid, calculate the mean of the values of all the points belonging to it. The mean value becomes the new value of the centroid.

The K-means algorithm takes the minimum mean square error within the cluster as the loss function. Assuming that the sample set is divided into K clusters and the i-th cluster is denoted as, the loss function can be written as:

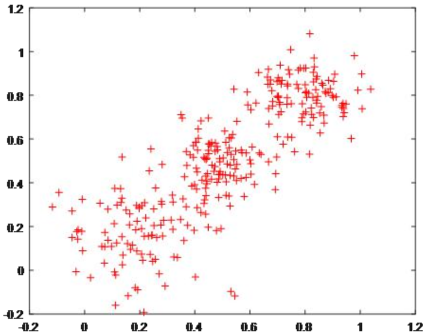
$$I=\sum_{i=1}^{k}\sqrt{\sum_{x\in c_i}(x-u_i)^2} \tag{2}$$

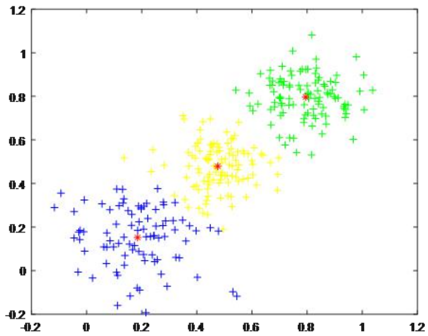
Where is the mean within the cluster, which is also the meaning represented by the mean in the k-means algorithm.This loss function describes how closely the samples in the cluster are clustered around the mean, that is, how similar the samples in the cluster are.

The clustering process of k-means algorithm is as follows:

- Randomly select K sample points in the sample set as the initial value of the mean
- Calculate the distance between each sample and each mean
- Divide the sample into the cluster with the closest mean point
- Update each mean point according to the samples in the cluster
- Repeat steps 2-4
- Get K clusters

Next, we illustrate the clustering process of k-means algorithm with real data.We generate 126 samples of gaussian white noise with standard deviation of 1 around (0.2,0.3), (0.4,0.6) and (0.8,0.8) points in the coordinate plane respectively, as shown in the Fig. 14(a).We use the k-means algorithm to cluster this sample set. Ideally, the k-means algorithm can find our initial sample center points (0.2,0.3), (0.4,0.6) and (0.8,0.8) only through these samples.The sample set is iterated four times, and the clustering of the first iteration is shown in Fig. 14(a).





(a): 126 samples of gaussian white noise (b):Aggregatedata sets into three categories

Figure 14: Iterative process

As is shown in Fig. 14(b), after 4 iterations, the coordinates of the mean points of the three clusters are (0.2,0.2), (0.5,0.5) and (0.8,0.8) respectively. It can be seen that the data have been clustered into three categories.

2.4 Feature Extraction of The Elderly People Balance Ability

2.4.1 Weight Analysis based on BP Neural Network

BP neural network was selected as the model for feature extraction.It can learn and store a large number of input-output mapping relationships without revealing the mathematical equations that describe the input-output mapping relationship beforehand. In BP neural network, the steepest descent method is often adopted as the rule of network learning. The predicted output can approach the expected output continuously, which is realized by updating the weight of the network through the feedback adjustment of the error obtained through signal forward propagation.

The model structure of BP neural network constructed in this study can be divided into three layers. The first layer is the input layer, the second layer is the hidden layer, and the third layer is the output layer. Fig. 15 shows the three-layer BP neural network.If the output result calculated by the forward propagation of signal does not reach the expected output value, the error value of feedforward propagation will be calculated.In order to make the predicted output approach the expected output continuously, the prediction error is calculated and propagated in the reverse direction during the feedforward propagation of the signal, and the weight in the network is adjusted appropriately according to the magnitude of the error value until the pre-set error function or the specified training number is reached.

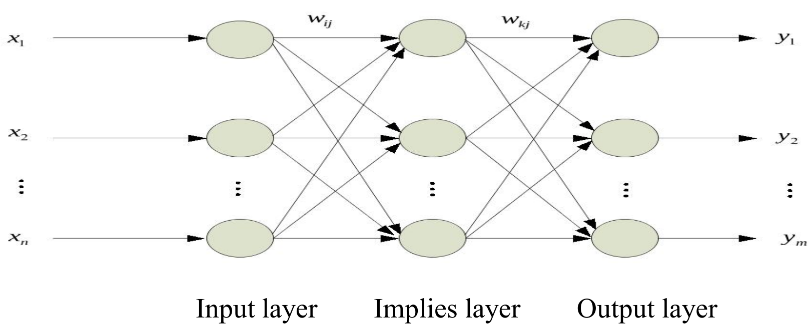


Figure 15:BP neural network structure

The specific implementation steps of BP learning algorithm are as follows:

Step1: Set the initial weight w and threshold θ.

Step2: The packaged sample set was extracted into the training of BP neural network for signal forward propagation, and the predicted output of the network was obtained, namely formula (3) :

$$y_{iq}(t) = f\left[x_{iq}(t)\right] = f\left[\sum_j w_{ij}(t) I_{jq}\right] \tag{3}$$

When group q is input, the network output of the k_{th} node in the output layer after weight djustment t times is $y_{kq}(t)$.

Step3: Calculation of network objective function.

The corresponding objective function is shown in formula (4) :

$$E_q(t) = \frac{1}{2}\sum_k \left[d_{kq}(t) - y_{kq}(t)\right]^2 \tag{4}$$

Where $E_q(t)$ represents the corresponding network objective function when the sample input of group p is input. Z(t) is the total objective function:

$$Z(t) = \sum_p E_p(t) \tag{5}$$

Step4: Discriminate whether the termination conditions are satisfied. If the prediction error function set by Z(t) in formula (5) is less than or equal to, the termination condition is satisfied;Otherwise, go to **Step5**.

Step5: Feedback of sample error. According to the process of reverse propagation of Z(t) error from output to input, the adjustment formula of iteration algorithm for connecting weight t+1times from neuron j to neuron i is as follows:

Where η is the step size.

The modified weights are used for the forward propagation of BP neural network training and the target function is calculated. When the termination condition is satisfied: the total target function of the network is less than the function of the set of the error function, and reach the prescribed number of training.

$$w_{ij}(t+1) = w_{ij}(t) - \eta \frac{\partial J(t)}{\partial w_{ij}(t)} = w_{ij}(t) - \eta \frac{\partial E_p(t)}{\partial w_{ij}(t)} \tag{6}$$

Where η is the step size.

The modified weights are used for the forward propagation of BP neural network training and the target function is calculated. When the termination condition is satisfied: the total target function of the network is less than the function of the set of the error function, and reach the prescribed number of training.

2.4.2 Body Balance Features

We put the data of 42 monitoring points of each person into the training model and record their weights respectively. The top 15 weights with the highest contribution rate were selected. Then we make up 25 features which can affect the body's balance by using these weights.

Next, we list 25 characteristics as follows:

Table 3: Composition Analysis of Equilibrium Characteristic

	Balance features	**Monitoring points**	**Description**
1	Step length	36,37	Step length means that the distance between one side of the foot following the ground and the next side following the ground, have individual difference, basically concern with the mobile range of leg length, hip joint and knee joint, have the cent of right and left lower limbs
2	Stride length	34,35,36,37	In one step cycle, the distance from the coordinate position of a unilateral footprint to the forward and backward direction of the next footprint on the same side
3	Step width	34,35,36,37	Step width refers to the distance between the center lines of the feet on both sides in walking. The calculation method is the distance from the center of the footprint on one side to the center of the footprint on the opposite side along the X-axis in a gait cycle
4	Step speed	34,35,36,37	Step speed refers to the distance traveled per unit of time, which measures the speed of walking. The test method is to measure the distance traveled and the time by the footpath, and then calculate the ratio
5	Acceleration	34,35,36,37	It's the ratio of the change in velocity to the time it takes for that change to occur
6	Step frequency	36,37	Refers to the number of steps per minute
7	Walking cycle	36.37	Walking cycle refers to the period from the moment of landing of one leg to the next landing moment
8	Left support phase	35,37	The left supporting phase is the time when the left lower limb contacts the ground in a gait cycle. The calculation method is that the time when the left lower limb leaves the ground at the NTH step minus the time when the left lower limb lands
9	Right support phase	34,36	The right supporting phase is the time when the right lower limb contacts the ground in a gait cycle. The calculation method is that the time when the right lower limb leaves the ground at the NTH step minus the time when the right lower limb lands
10	Double supported phase	34,35,36	Double support phase refers to the time when both lower limbs touch the ground at the same time in a gait cycle. Single support phase and swing phase each account for 40% of the walking cycle, and the other time is the period of double support phase
11	Oscillating phase	35,37	Swing refers to the period of time in which a single lower limb swings in the air in a gait cycle. The calculation method is that the moment of landing of a single lower limb minus the moment of leaving the ground of the previous step
12	Height of gravity center	15,33	The center of gravity is the point of action of the resultant force of gravity on each part of an object
13	average swing	4,11,34,36	Mean swing amplitude refers to the average distance of COP from the swing center point
14	The length of the swing track	34,35,36,37	The length of the swing track, namely, the swing distance, refers to the total length of the trajectory of the pressure center in a certain period of time, which is the sum of the distance between adjacent COP points.
15	Swing speed	7,9,11,34,36	The swing rate refers to the distance of COP swing in unit time, which indicates the speed of body swing and is also related to the direction.

16	Foot Angle	35,37	It refers to the Angle between the center line of the foot print and the foot line of the footprint, which is related to the ankle joint of the lower limb and can be divided into left and right
17	Symmetric oscillations	35,37	The ratio of the size of the oscillating phase parameters
18	Weight of support	35,37	Ratio of support phase parameters
19	Step size symmetry	36,37	Ratio of step size parameters
20	Brain history	40,41	Patients with vestibular disorders, CVD, head trauma, etc.
21	Visual impairment	40	If suffer from vision obstacle cannot adjust through lens wait for a disease
22	Nervous system	29,31	Such as drug/alcohol withdrawal symptoms, neurological disorders, hypoxia and other diseases
23	Muscular dystrophy	7,8,11,12	Such as suffering from skeletal myopenia, muscular atrophy and other diseases
24	Bone damag	7,8,29,34	Such as the history of osteoporosis fractures and other diseases
25	Cardiovascula r	22,40	Such as suffering from hypotension, hypertension, cardiovascular disease and other diseases

III. Model Establishment and Solution of Problem 2

Based on the twenty-five behavioral characteristics extracted from the previous chapter, we constructed an assessment model for the balance ability of the elderly.

3.1 Problem 2 Solving Process

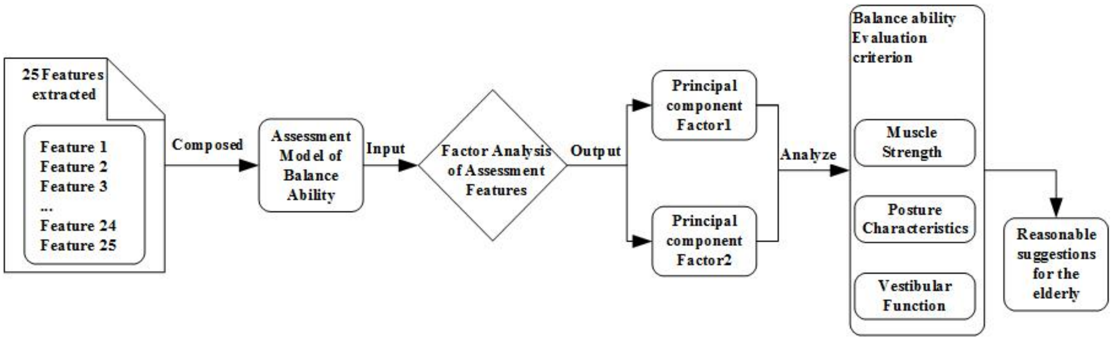


Figure 16: Flow chart of Problem 2

Fig.16 shows the whole brief process of how our group solve the Problem 1.

3.2 Standardize the Extracted Features

Because the computer can't deal with the specific characteristics of real concepts, it is necessary to digitalize all the extracted features, and then abstract the extracted concept of behavioral characteristics into digital signals for processing.

- Firstly, we quantify the behavioral features described by language. For example, a behavioral feature can be summarized into four states: {A, B, C, D} with corresponding values of 5, 4, 3 and 2, respectively. The Cauchy distribution membership functions with membership constraints f(5)=1, f(3)=0.8 and f(1)=0.01 are used to quantify the membership functions.

$$f(x)=\begin{cases}[1+\alpha(x-\beta)^{-2}]^{-1}, & 1\le x\le 4\\ a\ln x+b, & 4\le x\le 7\end{cases} \tag{7}$$

The quantized values of states A, B, C and D are {1, 0.9, 0.8, 0.5}.

- In order to eliminate the influence of dimension and range differences between indicators, standardization is needed.Since the feature data will be analyzed by factor analysis afterwards, we need to use distance measure similarity to reduce data dimension, Z-Score standardization is used to process data features.
  1. Calculate the mathematical expectation E (x) and variance D (x) of each variable.
  2. Z-Score standardize the original data x:

$$x'=\frac{x-E(x)}{\sqrt{D(x)}} \tag{8}$$

3.3 Factor Analysis

Factor analysis is an extension of principal component analysis. In factor analysis, the intrinsic correlation structure between the original variables is grouped, and the strong correlation is grouped into a group.

3.3.1 Mathematical Model

Assuming that there are *p* original variables, which are expressed by $\{x_1,x_2,...,x_p\}$ respectively. The mean value of each variable is 0 and the standard deviation is 1, the linear combination of *k* (*k* < *p*) factors $\{f_1, f_2, ..,f_k\}$ is used to represent each original variable:

$$\begin{aligned}x_1&=a_{11}f_1+a_{12}f_2+......a_{1k}f_k+\mu_1\\ x_2&=a_{21}f_1+a_{22}f_2+......a_{2k}f_k+\mu_2\\ &...\\ x_p&=a_{p1}f_1+a_{p2}f_2+......a_{pk}f_k+\mu_p\end{aligned} \tag{9}$$

Representation by matrix method:

$$X=AF+\varepsilon \tag{10}$$

3.3.2 Prerequisites for Factor Analysis

- Calculate the correlation coefficient and make statistical test.If most of the correlation coefficients in the correlation coefficient matrix are less than 0.3, then these variables are not suitable for factor analysis.
- Computing the relevance matrix of reflection image:

$$MSA_i = \frac{\sum_{i \neq j} r_{ij}^2}{\sum_{i \neq j} r_{ij}^2 + \sum_{i \neq j} p_{ij}^2} \tag{11}$$

- Bartlett's sphericity test

  Starting from the correlation coefficient matrix of the original variable, assuming that the correlation coefficient is the unit matrix, if the corresponding *P* value of the test is less than the given significance level *a*, the original hypothesis should be rejected and the original variable should be considered suitable for factor analysis.

- KMO test

  The value range of this statistic should be between 0 and 1. The closer the value is to 1, the stronger the correlation between variables, so the original variables are suitable for factor analysis. More than 0.9 is appropriate; 0.8-0.9 is appropriate; 0.7-0.8 is general; 0.6-0.7 is acceptable; 0.5-0.6 is not appropriate; 0.5 below is extremely inappropriate.

3.3.3 Factor Extraction and Solution of Factor Load Matrix

Principal Component Analysis transforms the original P variables into another set of irrelevant variables y by means of coordinate transformation.

$$\begin{aligned} & y_1 = u_{11}x_1 + u_{12}x_2 + \ldots\ldots u_{1p}x_p \\ & y_2 = u_{21}x_1 + u_{22}x_2 + \ldots\ldots u_{2p}x_p \\ & \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ & y_1 = u_{p1}x_1 + u_{p2}x_2 + \ldots\ldots u_{pp}x_p \\ & u_{i1}^2 + u_{i2}^2 + \ldots\ldots u_{ip}^2 = 1(i = 1、2、\ldots、p) \end{aligned} \tag{12}$$

The variables determined according to the above principles are in turn the first, the seond, ... , the *p-th* principal component of the original variable.The first principal component accounts for the largest proportion of the total variance, while the other principal components in the total variance decreases in turn.

3.3.4 Specific Solution Steps

- **Step1:** Standardizing all the original variables.
- **Step2:** Calculating the correlation coefficient matrix of variables.
- **Step3:** Finding the eigenvalues $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p$ and corresponding eigenvectors u_1、u_2、...、u_p of the correlation coefficient matrix.

Since the purpose of factor analysis is to reduce the number of variables, when calculating factor load matrix, we usually do not select all eigenvalues. By choosing only the first *k* eigenvalues and eigenvectors, the following factor load matrix containing *k* factors is obtained:

$$A=\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mm} \end{bmatrix}=\begin{bmatrix} u_{11}\sqrt{\lambda_1} & u_{12}\sqrt{\lambda_2} & \dots & u_{1m}\sqrt{\lambda_m} \\ u_{21}\sqrt{\lambda_1} & u_{22}\sqrt{\lambda_2} & \dots & u_{2m}\sqrt{\lambda_m} \\ \dots & \dots & \dots & \dots \\ u_{m1}\sqrt{\lambda_1} & u_{m2}\sqrt{\lambda_2} & \dots & u_{mm}\sqrt{\lambda_m} \end{bmatrix} \tag{13}$$

Generally, the number of characteristic roots whose cumulative variance contribution rate is more than 85% is chosen as the number of factors.

Formula for calculating cumulative variance contribution rate:

$$C=\frac{\sum_{i=1}^{m}\lambda_i}{\sum_{i=1}^{p}\lambda_i} \tag{14}$$

3.3.5 Analysis of the Results

Table 4: Prerequisite test results of factor analysis

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.754
Bartlett's Test of Sphericity	Approx. Chi-Square	363.159
	df	38
	Sig.	.000

It can be seen from this table that the KMO statistic is 0.754, which is larger than the minimum standard. At the same time, P < 0.001 under the Bartlet spherical test,, shows that it is suitable for factor analysis.

Table 5: Factor Analysis Result

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	5.780	83.462	82.148	5.780	83.462	82.148	3.108	45.261	42.263
2	.553	7.543	90.702	.653	7.543	90.702	3.045	44.482	90.702
3	.249	3.459	94.462						
4	.122	1.483	96.259						
5	.058	.542	99.054						
...	...	...	...						
25	.001	.003	100.000						
Extraction Method: Principal Component Analysis.									

It can be seen from this table that the first two principal components have larger eigenvalues and their cumulative contribution rate reaches 90.702%. Therefore, the first two common factors are selected. Through the analysis of the first two characteristic principal components, it can be concluded that the balance ability of the elderly is mainly evaluated by the sense of balance of sense and the strength of walking muscles.

3.4 Suggestions to Elderly People with Weak Balance Ability

According to the factor analysis of the model in the previous section, two main components in the evaluation of balance ability are obtained, which can be obtained by further analysis.

3.4.1 The Influence of Vestibular Organs and Related Suggestions

In balance control, vestibule is mainly reflected in the regulation of its internal instability. The improvement of vestibular function can better ensure the overall coordination of the body and make the balance ability of the body more stable.

Therefore, we suggest that the elderly should take effective measures to improve vestibular function. In China, most elderly people go to parks to play Tai Chi and dance. Because related studies have shown that these exercises can improve the function of the vestibular organs of the brain and improve the balance ability of the elderly.

3.4.2 The Influence of Muscle Strength and Related Suggestions

Human trunk control ability is closely related to balance ability, so muscle strength plays an important role in maintaining balance ability. Increasing lower limb muscle strength in the elderly can delay the decline in balance.

Therefore, we propose that the elderly should maintain muscle strength training. Muscle strength training is one of the most effective ways to improve balance ability, which can improve the postural control of the elderly. And the elderly muscle strength training should follow the gradual principle. Because the muscle strength will decrease after the training stops, the muscle strength training of the elderly should be regular and persevere for a long time.

3.4.3 The Influence of Body Shape and Related Suggestions

Body features can be obtained from the data of all detection points at a certain time. Because of the larger weight itself and the distribution ratio of body mass, the change of body gravity center in exercise is more complex, so the balance ability of human body will be reduced. Older people's obesity may lead to a greater reduction in balance ability and increase the probability of falling events.

Therefore, we suggest that the elderly should pay attention to the prevention of obesity. Older people can lose weight through lifestyle changes, such as diet and exercise.

IV. Comparative Analysis and Solution of Problem 3

4.1 Comparative Analysis of Actual Data

According to the actual data, the comparative analysis of body balance force is as follows:

Table 6: The kinds of diseases of the elderly at different ages

Age	Brain disease	Visual impairment	Neuropathic disorder	Anoxia	Muscular dystrophy	Bone damage	Cardiovascular& Blood pressure
50-59	50%			25%	25%	50%	50%
60-69	2%	25%	20%	13%	11%	60%	70%
70-79	4%	11%	21%	21%	11%	50%	64%
80-89		28%	15%			57%	71%

According to Tab. 6, we can see that osteoporosis, cardiovascular disease, hypertension and hypotension are the main causes of poor balance in the elderly people.

Table 7: Falls of elderly people of different ages

Age	Number of elderly people	Fall times in one year	Percentage	Fall times in 2015	Percentage
50-59	4	0	0%	0	0%
60-69	44	5	11.3%	10	22.7%
70-79	25	10	40%	5	20%
80-89	7	1	14.2%	4	57.1%

According to Tab. 7, we can see that the number of falls increases with age, indicating that the balance ability is gradually decreasing.

Table 8: The ratio of falls corresponding to different heights and weights

Height	Fall times/number of people	Weight	Fall times/number of people
140-150	7%	50-60	24%
150-160	32%	60-70	20%
160-170	17.6%	70-80	16.6%
170-180	10%	80-119	14.2%

From Tab. 8, we can see that people with height between 150 and 160 or weight between 50 and 60 are more likely to fall down.

4.2 Suggestions to Elderly People with Weak Balance Ability

We used the established model to predict and evaluate the balance ability of the elderly, and the obtained comprehensive index was compared with the real data in the above table, and the correlation was shown in Fig. 14.

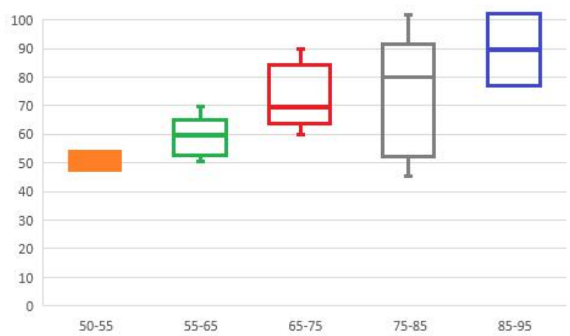


Figure 17: Correlation analysis between real data and model data

It can be seen that the average score of the comprehensive indicators of each group of subjects is positively correlated with the indicators affecting the balance ability reflected by the real data.This shows that the comprehensive evaluation index is effective and can reflect the level of balance ability.

Through observation and comparison of the data, we will put forward several suggestions for the elderly with weak balance ability:

First, to actively use the method of food therapy to prevent and treat diseases, such as radish and kiwi for hypertension food, both safe and no side effects.

Secondly, with the increase of age, the elderly should be accompanied or take auxiliary equipment when they go out to reduce the chance of falling down.

Finally, elderly people should walk slowly and avoid strenuous exercise.

V. Strengths and Weaknesses

5.1 Strengths

- We use recursive Kalman filter algorithm to pre-process data. It is suitable for the estimation of multidimensional stochastic processes, and only the statistical characteristics of the system state at the current time need to be considered in the estimation process.
- Data aggregation is implemented by K-means classification algorithm. This method has high accuracy in data processing and is insensitive to outliers in data, so it can avoid the influence of error data on clustering results.
- In training, BP neural network can automatically extract the "reasonable rules" between output and output data by learning, and self-adaptively memorize the learning content in the weights of the network.

5.2 Weaknesses

- If there is a large observation error in Kalman filter. Then it will affect the filtering of erroneous and unreasonable data.
- Because BP neural network algorithm is essentially gradient descent method. If the objective function to be optimized is very complex, there will inevitably be a "zigzag phenomenon".
- The research only focuses on steps, center of gravity and motion, but lacks the influence of other factors on balance ability.
- The initial stage of the whole model is only from gait, center of gravity and movement. There is a serious lack of other factors affecting the balance ability of the whole model.

VI. Future Work

At the beginning of its establishment, the model was limited in the research of gait, center of gravity and movement, without considering the effects of vision, drugs and some physical diseases on human balance ability. This makes the feature extraction model in the input data analysis, the analysis of information will be relatively limited. Next, we will take into account the situation of vision, medicine and physical diseases, improve the model to extract features affecting the ability of balance characteristics, and better make a comprehensive body balance evaluation for the elderly.

With the development of the aging of the social population, the proportion of the elderly in the social group will gradually increase. In the future, the model can be widely used in the elderly population, accurately evaluate the balance ability of the elderly, and give systematic evaluation suggestions. Older people adjust their lifestyle according to the suggestions given by the system, so as to reduce the probability of falls in the elderly group.

References

[1] Blake A J , Morgan K , Bendall M J , et al. FALLS BY ELDERLY PEOPLE AT HOME: PREVALENCE AND ASSOCIATED FACTORS[J]. Age and Ageing, 1988, 17(6):365-372.
[2] Stalenhoef P A, Diederiks J P M, Knottnerus J A, et al. A risk model for the prediction of recurrent falls in community-dwelling elderly: A prospective cohort study[J]. Journal of Clinical Epidemiology, 2002, 55(11):1088-1094.
[3] Kalula S Z . A WHO Global Report on Falls among Older Persons Prevention of falls in older persons[J].
[4] Dargentmolina P , G Bréart. [Epidemiology of falls and fall-related injuries in the aged].[J]. Revue Dépidémiologie Et De Santé Publique, 1995, 43(1):72.
[5] Northridge M E , Nevitt M C , Kelsey J L , et al. Home hazards and falls in the elderly: the role of health and functional status[J]. American Journal of Public Health, 1995, 85(4):509-15.
[6] Steinmetz H M , Hobson S J G . Prevention of Falls Among the Community-Dwelling Elderly:[J]. Physical & Occupational Therapy In Geriatrics, 1995, 12(4):17.

Appendix

Programs and codes

```
#2.1 K-means code
function kmeans()
k=3;
x = 0.8 + sqrt(0.01) * randn(100,2);
y = 0.2 + sqrt(0.02) * randn(100,2);
z= 0.5 + sqrt(0.01) * randn(100,2);
D=[x;y;z]; u=randperm(size(D,1),k);
u=D(u,:);
c=zeros(size(D,1),1);distance=zeros(k,1);
while 1
    mark=0;
    for i=1:size(D,1)
```

```
            for j=1:k
                  distance(j)=sqrt((D(i,1)-u(j,1))^2+(D(i,2)-u(j,2))^2);
            end
            [~,m]=min(distance);
            c(i)=m;
      end
      u1=zeros(k,2);
      for i=1:k
            u1(i,1)=sum(D(find(c(:)==i),1))/size(find(c(:)==i),1);
u1(i,2)=sum(D(find(c(:)==i),2))/size(find(c(:)==i),1);
            if u(i,1)~=u1(i,1)||u(i,2)~=u1(i,2)
                  mark=1;
                  u(i,1)=u1(i,1);
                  u(i,2)=u1(i,2);
            end
      end
     if mark==0
           break;
     end
end
x=D(find(c(:)==1),:);
y=D(find(c(:)==2),:);
z=D(find(c(:)==3),:);
plot(x(:,1),x(:,2),'+y',y(:,1),y(:,2),'+b',z(:,1),z(:,2),'+g',u(1,1),u(1,2),'*r',u(2,1),u(2,2),'*r',u(3,1),u(3,
2),'*r');
axis([0,1,0,1]);xlabel('red');ylabel('acc');title('');
end

#2.2 function [ W,theta ] = BP_tranning( X, levels, step )
      if nargin < 3
           step = 0.01;
      end
            n_levels = size(levels, 2) - 1;
      n_input = levels(1);
      W = create_w(levels);
      theta = create_theta(levels);
            [n_data,col] = size(X);
      n_label = col - n_input;
            tranning_data = X(:,1:n_input);
      label = X(:,n_input+1:end)';
      f = @sigmoid;
      eps = 10e-5;
      old_error = 0;
      while true
```

```
        for k=1:n_data
            [output, Y] = BP_predict(tranning_data(k,:), W, theta);
            delta = label(:,k) - output;
                        for l=n_levels:-1:1
                net = W{l}*Y{l} + theta{l};
                if l == n_levels
                    S = diag(f(net).*(1-f(net)))*delta;
                else
                    S = diag((f(net).*(1-f(net))))*W{l+1}'*S;
                end
                                new_W{l} = W{l} + step*S*Y{l}';
                new_theta{l} = theta{l} + step*S;
            end                W = new_W;
            theta = new_theta;

        end                    y = BP_predict2(tranning_data, W, theta);
        delta = y - label;
        error = sum(sum(delta.^2))
        if abs(error - old_error) < eps;
            break;
        else
            old_error = error;
        end
    end
end
```

#3.1 z-score code

```
function [normalized_data] = zscorenormalize(source_data, kind)
[m, n]    = size(source_data);
normalized_data = zeros(m, n);
for i = 1:n
        mea = mean( source_data(:, i) );
        st = std( source_data(:, i) );
        normalized_data(:, i) = ( source_data(:, i)-mea ) / st;
end
```

#3.2 Factor analyze code:

```
function [F,v0,A,v,w]=Factor analyze(x,number)
[x]=zscorenormalize(x);
[n,m]=size(x);
x1=x;
s=x1(1,:)'*x1(1,:)/(n-1);
for i=2:n
     s=s+x1(i,:)'*x1(i,:)/(n-1);
```

```
end
[v,d]=eig(s,'nobalance');
  for i=1:m
        e=1/norm(v(:,m-i+1));
                v1(:,i)=e.*v(:,m-i+1);
  end
for i=m:-1:1
        d1(m-i+1)=d(i,i);
end
d=d1;
for i=1:m
        t(1,i)=d(i)/sum(d(:));
        t(2,i)=sum(d(1:i))/sum(d(:));
end
  for i=1:m
        if (t(2,i))<number
                v0(:,i)=sqrt(d(i)).*v1(:,i);
                t(2,i)=z(i);
        end
  end
  A=v0;
  [nx_A,ny_A]=size(A);
  for j=1:ny_A
        for k=j+1:ny_A
uu=(A(:,j).^2-A(:,k).^2);
vv=2*A(:,j).*A(:,k);
for i=1:nx_A
        uu(i)=uu(i)/(A(i,:)*A(i,:)');
        vv(i)=vv(i)/(A(i,:)*A(i,:)');
end
a_A=sum(uu);
a_B=sum(vv);
a_C=sum(uu.^2-vv.^2);
a_D=2*sum(uu.*vv);
kk_s=(a_D-2*a_A*a_B/nx_A)/(a_C-(a_A^2-a_B^2)/nx_A);
s_angle=0.25*atan(kk_s);
if (a_D-2*a_A*a_B/nx_A)>0
        a=abs(cos(s_angle));
        b=abs(sin(s_angle));
else
        a=abs(cos(s_angle));
        b=-abs(sin(s_angle));
end
FF=eye(ny_A);
```

```
FF(j,j)=a;
FF(j,k)=-b;
FF(k,k)=a;
FF(k,j)=b;
A=A*FF;
          end
  end
R = corr(x);
F=A'*R\x;
y=z*F;
[w,v]=sort(y,'descend');
```