

RESEARCH ARTICLE

High Accuracy Human Activity Recognition Based on Sparse Locality Preserving Projections

Xiangbin Zhu^{1*}, Huiling Qiu¹

College of Mathematics, Physics and Information Engineering, Zhejiang Normal University, Zhejiang, China

¹ These authors contributed equally to this work.

* zhuxb@zjnu.cn



Abstract

Human activity recognition(HAR) from the temporal streams of sensory data has been applied to many fields, such as healthcare services, intelligent environments and cyber security. However, the classification accuracy of most existed methods is not enough in some applications, especially for healthcare services. In order to improving accuracy, it is necessary to develop a novel method which will take full account of the intrinsic sequential characteristics for time-series sensory data. Moreover, each human activity may has correlated feature relationship at different levels. Therefore, in this paper, we propose a three-stage continuous hidden Markov model (TSCHMM) approach to recognize human activities. The proposed method contains coarse, fine and accurate classification. The feature reduction is an important step in classification processing. In this paper, sparse locality preserving projections (SpLPP) is exploited to determine the optimal feature subsets for accurate classification of the stationary-activity data. It can extract more discriminative activities features from the sensor data compared with locality preserving projections. Furthermore, all of the gyro-based features are used for accurate classification of the moving data. Compared with other methods, our method uses significantly less number of features, and the over-all accuracy has been obviously improved.

OPEN ACCESS

Citation: Zhu X, Qiu H (2016) High Accuracy Human Activity Recognition Based on Sparse Locality Preserving Projections. PLoS ONE 11(11): e0166567. doi:10.1371/journal.pone.0166567

Editor: Houbing Song, West Virginia University, UNITED STATES

Received: July 11, 2016

Accepted: October 31, 2016

Published: November 28, 2016

Copyright: © 2016 Zhu, Qiu. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Our paper only involves access to data stored on publicly available databases. The data set that we used was not collected by us. Our experiment data set exists in the public domain. There have been lots of published papers which used the data set. The data set link is: <https://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones>. All data have been robustly anonymized because there is no personal information in the data set. The individuals to whom the data relates cannot be identified because the researchers do not have access to any identifiable information. All researches using the

Introduction

With the rapid development of information technology, it has been observed there is an accelerated growth of smartphones, which incorporate a variety of sensors, such as high-resolution cameras, light sensors, gyroscopes, accelerometers, GPS, temperature sensors and so on [1]. In recent years, the people with smartphones have reached near 80% of the world adult population [2], and by 2020, it will reach 80% [3]. It can be envisioned that such powerful devices can provide a tool to automatically monitor activities of daily living (ADL) and enhance us the ability of making better decision regarding our future actions [4]. This is not only for its flexibility, convenience and availability but also for its easiness to use [5]. For these reasons, human activity recognition with smartphone sensor data has been a hot research topic.

data set need not to identify the individuals. This is the reason why there is no any personal information in the data set.

Funding: The author(s) received no specific funding for this work.

Competing Interests: The authors have declared that no competing interests exist.

There have been a lot of researches to identify human activity based on video data. Many of these results can be applied to human activity recognition of sensor data. But we need to take full account of the features of smartphone sensor data. Firstly, smartphone sensor data is multi-sensor data streams. The features of smartphone sensor data usually have more dimensionality than that of video data in human activity recognition. Secondly, not like HAR of video data, HAR of smartphone sensor data cannot directly recognize a human activity from shape features, which can be extracted from video data. HAR of smartphone sensor data need to indirectly infer the human activity from sensor data streams, such as accelerometers data, gyroscopes data and localization sensor data. Therefore, we should take advantage of the characteristics of human activity. For example, we can classify human activity based on the characteristics of human activity so that the common attribute of one category of human activity can be employed by HAR. Moreover, human activities can be hierarchically classified because human activities can be grouped into subsets as cluster analysis does. Thus, we can hierarchically recognize human activity with a multi-stage classifier.

Nowadays, the recognition of human activity using smartphones sensor has an upsurge of interest for many researchers and gained preliminary achievements. For example, there have been proposed some outstanding AI methods on human recognition. Wei *et al.* [6] used a hidden Markov model(HMM) to build the cloud resource allocation model, which is based on an imperfect information Stackelberg game. The HMM method was employed to predict the service provider's current bid. Li *et al.* [7] proposed a conditional random field based event detection method to analyze the human group behavior, and the human groups were identified by a temporal-spatial clustering approach. Lee and Cho proposed hierarchical hidden Markov models(HHMMs) [8], Sharma *et al.* used neural network [9], Anguita *et al.* used Multiclass Hardware-Friendly Support Vector Machine [10] and Ronao *et al.* introduced the technology of two-stage CHMMs [11]. However, these research results are not effective enough for some applications. So, we need a creative method to improve performance in several aspects including accuracy and speed. In dataset processing, for instance, dimensionality reduction, which aims to reduce the number of features, has still room for improvement. For example, Anguita *et al.* used 561 variables and Charissa Ann Ronao *et al.* selected 119 variables. Moreover, although many different classification methods have been used, the final recognition rate is still not accurate enough, especially for the easily confused activities. One reason is the different physical location of sensor which volunteer worn to perform the activities might cause different data collected [5], and the data may be similar with features so that models are difficult to distinguish them.

Our approach is combined with dimensionality reduction and hierarchical classifier. Dimensionality reduction includes feature selection and feature reduction. Since the time series data we used in the experiment is continuous and standardized, each data point is very closely related with adjacent data points. Random forests variable importance measures perform well with these data [12]. The obtained database is real data sets and randomly partitioned into two sets: testing set and training set. Moreover, these data are high correlation coefficient of local, particularly in the stage of feature reduction in which the data local correlation coefficient is higher than the data in feature selection stage. SpLPP(Sparse Local Preserving Projections) optimally preserves the neighborhood structure the data set. It can choose the number of neighborhood and calculates the weight value adaptively. Moreover, SpLPP can extract more discriminative activities features from the sensor data compared with locality preserving projections. The variables of gyroscope play a leading role for stationary activities and the acceleration variables play a minor role [11]. Thus, we choose random forests variable importance measures as the method of feature selection and sparse local preserving projection as the way of feature reduction.

The accelerometer and gyroscope sensory data are sequential nature and multivariate. This is advantageous to use the CHMM to classify because those characteristic are properly consistent with the nature of this method [13]. These activities are hierarchical, which means we can multiple utilize the method of “one divides into two” to recognition activities. In this paper, we use a three-stage CHMMs approach.

Materials

In this section, we present a three-stage continuous hidden Markov model (TSCHMM) approach, which use random forests variable importance measures for feature selection and sparse local preserving projection for feature reduction. Meantime the three-stage CHMM is advantageous to the sensory data which has typical temporal characteristics and the inherent hierarchical frame of activities.

Our method is shown in Fig 1. The signal data is collected by the acceleration and gyroscopic sensor in a smartphone. In this paper, for introducing our work, we focus on six basic human daily activities, including Walking, Upstairs, Downstairs, Sitting, Standing and Laying. The first step is feature extraction. We totally extracted 561 features to describe each activity window, including the standard measures: means, correlation, signal magnitude area (SMA),

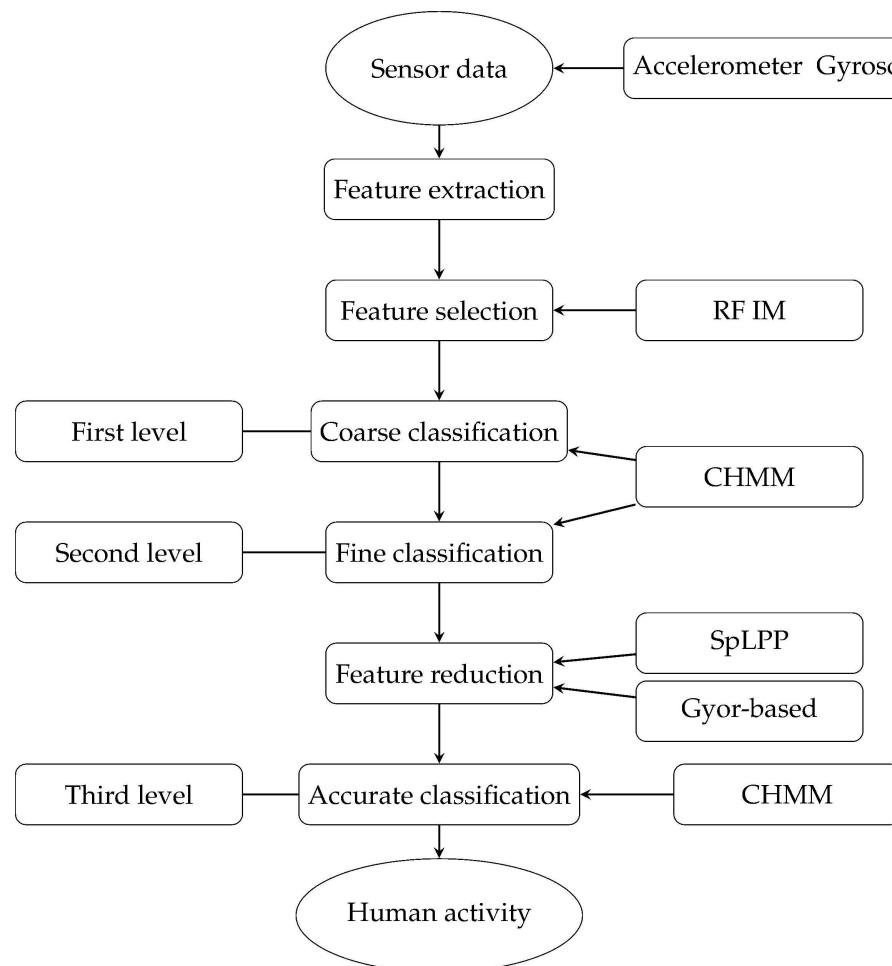


Fig 1. Human Activity Recognition Process Hierarchy.

doi:10.1371/journal.pone.0166567.g001

autoregression coefficients and the new features: energy of different frequency bands, frequency skewness, angle between vectors. The next is feature selection, using the RF to compute their importance scores. At coarse classification stage, TSCHMM divides the feature subset into moving subclass and stationary subclass. Moving subclass includes Walking, Upstairs and Downstairs. Stationary subclass includes Sitting, Standing and Laying. Different subclasses correspond to their respective feature subsets. At the second classification stage, the fine classifier classifies the moving subset into Walking subclass and Up-Downstairs subclass, the stationary activity is divided into Laying sub-activity and Sitting-Standing sub-activity. Then, the next step is feature reduction, using the SpLPP and Gyro-based to reduce the number of features. TSCHMM for accurate classification divides the final feature subsets into subclasses of Sitting, Standing, Upstairs and Downstairs.

Random Forest Importance Measures(RFIM)

Dimension reduction need not only to reduce the amount of data but also to find out which features are more important for each activity. The key technology is to quantitatively evaluate the importance of each feature. This paper employs the random forest importance measures method, which can effectively estimate each variable importance.

Random forests is a powerful ensemble method for classification, introduced by L. Breiman (2001). It builds a set of de-correlated trees [14]. RF(random forests) is a classifier composing of simple decision trees:

$$k(x, \Theta_h), h = 1, \dots, n; \quad (1)$$

where $\{\Theta_h\}$ are i.i.d random vectors and each tree casts a unit vote for the most popular class at input x .

This algorithm combines the idea of Bootstrap aggregation and bagging technology, and it has been proved, when dealing with a lot of redundant features problems, that bagging is worse than random subspace selection [15]. This method also has other important features such as out-of bag (OOB) error, variable importance and correlation.

In this paper, we emphasize two different objectives about variable selection: (1) to find out important variables which highly related to the response variable for explanation purpose; (2) to find out a small number of data volume for a good prediction of the response variable [16]. Let us briefly show the experiment framework:

1. Get the RF variable importance scores $a_i, i = 1, \dots, n$;
2. Order the n variables in decreasing;
3. Compute the averages scores b of RF variable importance scores;
4. Remain the variables $a_i, a_i \geq b, i = 1, \dots, n$, which scores are above the average;

Sparse Locality Preserving Projections(SpLPP)

According to the data type, the data dimension reduction methods can be divided into linear and nonlinear dimensionality reduction [17]. Typical linear dimension reduction algorithms include Principal Component Analysis (PCA), Multidimensional Scaling (MDS) and Linear Discriminant Analysis (LDA). Similarly, there are several nonlinear dimensionality reduction (Manifold Learning) methods, such as Isometric map (Isomap), Locally Linear Embedding (LLE), Laplacian Eigenmaps (LE), Local Tangent Space Alignment (LTSA) and Locality Preserving Projections (LPP). Some algorithms such as PCA and LDA have solid theoretical

foundations and they are easy to analyze. However, these methods also ignore the intrinsic geometry structural features of high dimensional space and finally lead to the low recognition rate. The manifold learning methods such as LLE, Isomap and LE exhibit good performance on some artificial data set, but on the real data set, they cannot get better results and often lead to “out-of-sample” problem.

LPP is a popular and efficient algorithm for linear manifold learning. It preserves the advantages of the nonlinear manifold learning and linear dimensionality reduction method [18]. Moreover, LPP has overcome the “out-of-sample” problem and develops the performance of manifold learning on real data. LPP establishes the data mapping base on neighborhood graph and keeps the local structure of data, then makes the high dimensional data X mapped to a low dimensional data Y , such that the points which are close in the higher dimensional space will be still mapped as close in lower dimensional space [19]. LPP aims to seek a transformation matrix A to implement the dimensionality reduction of data; that is $y_i = A^T x_i$. The objective function:

$$J = \sum_{i=1}^n \sum_{j=1}^n \|y_i - y_j\|^2 W_{ij}; \tag{2}$$

which W_{ij} are the projection weights between samples x_i and samples x_j .

We can get the transformation matrix A by minimizing the function as following algebraic steps:

$$\begin{aligned} A_{opt} &= \operatorname{argmin} \left(\sum_{i=1}^n \sum_{j=1}^n \|y_i - y_j\|^2 W_{ij} \right) = \\ & \operatorname{argmin} \left(\sum_{i=1}^n \sum_{j=1}^n \|A^T x_i - A^T x_j\|^2 W_{ij} \right) = \\ & \operatorname{argmintr} (A^T X D X^T A - A^T X W X^T A) = \\ & \operatorname{argmintr} (A^T X L X^T A) \end{aligned} \tag{3}$$

where D is the diagonal matrix, $D_{ii} = \sum_{j=1}^n W_{ij}$; $L = D - W$ is the Laplacian matrix, W is the projection weight matrix comes from W_{ij} .

In order to remove the arbitrary scaling factor, we impose a constraint condition $Y D Y^T = 1$. The transformation matrix can be obtained by solving the generalized eigenvalue problems:

$$X L X^T A = \lambda X D X^T A \tag{4}$$

Let the solutions of the Eq 2 be the column vector a_1, a_2, \dots, a_d , sorted according to their eigenvalue $\lambda_1, \lambda_2, \dots, \lambda_d$. Then we can get $A = (a_1, a_2, \dots, a_d)$. Thus, the embedding is as follows:

$$x_i \rightarrow y_i = A^T x_i, A = (a_1, a_2, \dots, a_d); \tag{5}$$

where y_i is the d -dimensional vector and A is a matrix of $n \times d$.

The drawback of local preserving projection algorithm is the number of samples which are contained in the local neighborhood cannot be adaptively selected. Moreover the neighborhood graph is very sensitive to data noise. However, SpLPP can effectively combine the LPP with the Sparsity representation theory. It gets the optimal solution of sparse representation by using l_1 norm method in the LPP algorithm [20] and solves the parameters selection. More importantly, it achieves the purpose of keeping the reconstruction relationship of input data.

SpLPP algorithm aims to use the sparsity representation theory to construct the projection weight W_{ij} . For each sample x_i , we compute the l_1 norm problem to find out its sparsity reconstruction weight coefficient W_i .

$$\min_{w_i} \|x_i - Xw_i\|_2 + \lambda \|w_i\|_1; w_{ij} \geq 0, j = 1, 2, \dots, m; \tag{6}$$

where $[w_{i,1}, \dots, w_{i,i-1}, 0, w_{i,i+1}, \dots, w_{i,m}]^T$ is a m -dimensional column vector and the i -th is 0, the other elements $w_{i,j}, j \neq i$ denote the contribution of each sample x_j to reconstruction the sample x_i .

By calculating the weight vector of each sample, we finally get the sparse reconstruction weight matrix $W = (w_{i,j})_{m \times m}$, which W_i is the optimal solution of sparse representation.

With the reduced dimension and local variance information preservation, the extracted features y will be used as the new input features for accurate recognition [21].

Continuous Hidden Markov Models(CHMM)

Hidden Markov models(HMM) arises out by a simple Markov chain. It is a statistical model for sequence of data items named observation vectors and system states cannot be observed but obtained through observation vector. It is an important method for both automatic speech recognition (ASR) and signal processing [22]. HMMs are based on a hidden Markov chain $\{S_n\}$; the state transition of the system. A sequence of hidden state $\{S_n\}$ and the observed variables $\{O_n\}$ are conditionally independent. Each of the distribution O_n only relies on the corresponding state S_n . Besides, HMM is a double stochastic procedure: the transitions between states; the statistical relationship between state and observation symbol. Since state-conditioned observation distributions of sensory data is continuous, we use the continuous HMMs in this paper. CHMMs are composed of two state sets and three probability matrices, the model as following equation:

$$\lambda = (A, B, \pi); \tag{7}$$

1. Hidden state S : the state is satisfied with Markov property and it is the actual state of the hidden Markov model;
2. Observation state O : the state is associated with the hidden state and it is observable;
3. Initial state distribution π : the distribution denote the hidden state probability matrix when $t = 1$,

$$\pi_i = P(S_0 = i), 1 \leq i \leq N; \tag{8}$$

where N is the probability distribution number;

4. Transition probability matrix A : the matrix describe the transition probability between hidden states,

$$a_{ij} = P[S_{t+1} = j | S_t = i], 1 \leq i, j \leq N; \tag{9}$$

5. Observation probability matrix B : the matrix is represent the transition probability between hidden state and observation state,

$$b_i(o) = p_o(o; \theta_i) = f(o; \theta_i); \tag{10}$$

where $p_o(o; \theta_i)$ indicate the emission density of state i ,

$$p_o(o; \theta_i) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp - \frac{(o - \mu_i)^2}{2\sigma_i^2}; \tag{11}$$

Eq 11 is a Gaussian distribution, the mean is μ and variance is σ^2 .

Three-Stage Continuous Hidden Markov Models(TSCHMMs)

Three-Stage Continuous Hidden Markov Models (TSCHMMs) exploit the hierarchical nature of activities and the arithmetical of CHMM. The recognition process starts with the acquisition of the sensor data. It contains coarse, fine and accurate stage classification which are described in the following section.

The First-Stage CHMMs for Coarse Classification. The sensor data, including six daily living activities, was gathered from the sensors built in a smartphone. In the first stage, we classify six activities into two classes (moving and stationary). As Alg. 1 shows, the sensor data needs to be preprocessed, be processed for features extraction by RFIM and then be supplied to CHMMs for first-level training and testing. Moving train data is used to train moving CHMM and stationary train data is fed to train stationary CHMM. The feature subsets are pre-selected and used in each step.

Algorithm 1 First-stage CHMMs for Classification.

```

Input:  $X_{train}$ : train data;  $y_{train}$ : train label;  $X_{test}$ : test data;  $y_{test}$ :
test label;  $O$ : Number of coefficients in a vector;  $Q$ : Number of states;  $M$ :
Number of mixtures;
Output: Stationary, Moving
1: Compute the RFVI scores  $a_i$  and the averages scores  $b$ 
2: while  $i \leftarrow 1$  to  $t$  do
3:   if  $a_i \geq b$  then
4:      $i1 \leftarrow i1 + 1$ 
5:      $data(:, i1) \leftarrow X_{train}(:, i)$ 
6:   else
7:     delete  $X_{train}(:, i)$ 
8:   endif
9: endwhile
10: while  $j \leftarrow 1$  to  $n$  do
11:   if  $y_{train}(j)$  in {1 6} then
12:      $X_{train1}(j, :) \leftarrow data(1, j1, :)$ 
13:   else
14:      $X_{train2}(j, :) \leftarrow data(1, j2, :)$ 
15:   endif
16: endwhile
17: Training Laying CHMM or Training Walking CHMM
18: for  $k \leftarrow 1$  to  $m$  do
19:    $X_{test}(k, :) \leftarrow data(1, 1, :)$ 
20:   compute the model log-likelihood
21: endfor
22: Selection through max probability

```

When train data has been trained moving CHMM and stationary CHMM, test data is supplied to the two CHMMs. The result can get two probability values to assess the classes. Thus if moving CHMM probability value is higher than stationary CHMM probability value, we can classify that it is the class of moving, vice versa. Hence, setting a subclass $r \in R$, we can build a

CHMM λ' , then evaluate the initial state distribution π , transition probability matrix A and observation probability matrix B , finally optimize the corresponding training state likelihood.

Based on the three basic problems and the corresponding algorithms of HMM [22] [23], in the first-stage testing, we use the feature subset O to estimate the observation likelihood of the subclass $r \in R$. We use the forward-backward algorithm to compute all of CHMMs for each subclass by $P(O|\lambda')$ and then get the CHMM of each subclass. This evidently requires a method to allow comparability between each CHMM. For this reason, we have opted to compute log-likelihood using a (mixture of) Gaussians HMM for each subclass and to choose the activity which corresponding to the highest probability, the relationship as:

$$r' = \operatorname{argmax}_{r \in R} P(O | \lambda'); \tag{12}$$

The Second-Stage CHMMs for Fine Classification. Each activity has its unique property. The Laying activity can be very efficiently classified from stationary class [24]. But for the subclass of moving, the state of Walking can be recognized as upstairs or downstairs with high probability. As Alg. 2 shows, the purpose of the second-stage CHMMs, combined with the activity hierarchical nature, is to recognize the Walking and Laying. In order to achieve this, based on the first-stage data, the process goes on running either towards the Laying subclass or the Sitting-Standing subclass. That means Laying train data is fed to the Laying HMM, Sitting-Standing train data is fed to Sitting-Standing HMM. As the same way, data of Walking is used to be trained Walking HMM, the same as data of Up-Downstairs. Finally, stationary and moving subclass will be classified into 2 activities.

Algorithm 2 Second-stage CHMMs for Classification.

Input: X_{train} : train data; y_{train} : train label; X_{test} : test data; y_{test} : test label; O : Number of coefficients in a vector; Q : Number of states; M : Number of mixtures;

Output: Laying, Sitting-Standing, Walking, Up-Downstairs

```

1: while  $j \leftarrow 1$  to  $n$  do
2:   if  $y_{train}(j)$  in {1 6} then
3:      $X_{train1}(j, :) \leftarrow data(1, j1, :)$ 
4:   else
5:      $X_{train2}(j, :) \leftarrow data(1, j2, :)$ 
6:   end if
7: endwhile
8: Training Laying CHMM or Training Walking CHMM
9: for  $k \leftarrow 1$  to  $m$  do
10:   $X_{test}(k, :) \leftarrow data(1, 1, :)$ 
11:  compute the model log-likelihood
12: end for
13: Selection through max probability

```

At this stage, we can build CHMM λ'^s for each stationary activity. Same as the first level, we get the highest probability r'_s for stationary subclasses, individually. We also build CHMM λ'^m for each moving activity and get the highest probability r'_m for moving subclasses. Noteworthy, the number of mixtures M is changed.

The Third-Stage CHMMs for Accurate Recognition. The last stage is, as shown in Alg. 3 and Alg. 4, aims to recognize Sitting and Standing from the subclass of Sitting-Standing and to classify Upstairs and Downstairs from Up-Downstairs data. The training and testing process will continue to run as the first and second level. We use two states for CHMMs and each of the subclass will be classified into two activities.

Algorithm 3 Third-stage CHMMs for Sitting-Standing classification.

Input: X_{train} : train data; y_{train} : train label; X_{test} : test data; y_{test} : test label; O : Number of coefficients in a vector; Q : Number of states; M : Number of mixtures;
Output: Standing, Sitting
1: $y_i = A^T x_i$
2: **while** $j \leftarrow 1$ **to** n **do**
3: **if** $y_{train}(j)$ in {4 5} **then**
4: $X_{train1}(j, :) \leftarrow data(1, j1, :)$
5: **else**
6: $X_{train2}(j, :) \leftarrow data(1, j2, :)$
7: **end if**
8: **end while**
9: Training Standing CHMM or training Sitting CHMM
10: **for** $k \leftarrow 1$ **to** m **do**
11: $X_{test}(k, :) \leftarrow data(1, 1, :)$
12: compute the model log-likelihood
13: **end for**
14: Selection through max probability

Algorithm 4 Third-stage CHMMs for Up-Downstairs classification.

Input: X_{train} : train data; y_{train} : train label; X_{test} : test data; y_{test} : test label; O : Number of coefficients in a vector; Q : Number of states; M : Number of mixtures;
Output: Upstairs, Downstairs
1: **while** $i \leftarrow 1$ **to** t **do**
2: **if** feature is Gyro-based data **then**
3: $i1 \leftarrow i1 + 1$
4: $data(:, i1) \leftarrow X_{train}(:, i)$
5: **else**
6: delete $X_{train}(:, i)$
7: **end if**
8: **end while**
9: **while** $j \leftarrow 1$ **to** n **do**
10: **if** $y_{train}(j)$ in {2 3} **then**
11: $X_{train1}(j, :) \leftarrow data(1, j1, :)$
12: **else**
13: $X_{train2}(j, :) \leftarrow data(1, j2, :)$
14: **end if**
15: **end while**
16: Training Downstairs CHMM or training Upstairs CHMM
17: **for** $k \leftarrow 1$ **to** m **do**
18: $X_{test}(k, :) \leftarrow data(1, 1, :)$
19: compute the model log-likelihood
20: **end for**
21: Selection through max probability

As previously mentioned, we build CHMM λ^{rss} for each Sitting-Standing activity and λ^{mmm} for each Up-Downstairs activity, then estimate the model parameter $(A, B, \pi)_{ss}$ and $(A, B, \pi)_{mm}$, finally optimize the corresponding training states likelihood. Just as the previous levels, we get the highest probability r'_{ss} and r'_{mm} for each subclass, individually.

It is worth noting that the feature of subclass in this level will be processed. We use SpLPP for feature reduction to get the new feature subsets which are fed to both Sitting and Standing CHMMs. For another group, basing on the original sensor data, we remain the gyroscope

variables and remove the acceleration variables to obtain new subset which are used to train Upstairs CHMM and Downstairs CHMM.

Results

In this paper, the experiment data is from the public domain UCI HAR data set, including accelerometer and gyroscope XYZ data [25]. The related sensors incorporate in a Samsung Galaxy SII smartphone, which was worn by each volunteer. Each person performed the six activities (WALKING, UPSTAIRS, DOWNSTAIRS, SITTING, STANDING, LAYING) twice. Every time the phones were worn on different location. One was on the left-side of the belt and the other was free setting by users. The sensor signals (accelerometer and gyroscope) were pre-processed by applying noise filters and then sampled in fixed-width sliding windows of 2.56 sec and 50% overlap(128 readings/ window). Thus, 17 signals were totally obtained by calculating variables in the time and frequency domain (e.g. mean, standard deviation, signal magnitude area, entropy, signal-pair correlation, ect.). Each activity window was described by 561 features and normalized between -1 and 1. Through the randomly partition, 70% of the volunteer dataset was selected as training data and 30% of the dataset was selected as testing data [25].

The idea of using RFIM feature selection proposed by the Genuer, et. al [16]. We ran repeat RF tests ($n_{tree} = 500; m_{try} = \sqrt{\text{size}(X, 2)}$) and got all of the variable importance scores. Then we ranked the scores in descending order and computed the average of all the variable importance scores, finally remained the variables which scores are above the average. After these processes, we kept 132 variables as the first stage feature subset. Moving and stationary feature subsets in the second stage are from the first stage feature subset.

We adopt the method of SpLPP feature reduction similar to the paper of Zheng Z, et. al. [20]. For subclass of Up-Downstairs on the second level feature subset, the number of the related features could be reduced from 132 to 13 by SpLPP technology. The 13 features would be used in the third stage. Based on original data, we selected the gyroscope variables and gave up the acc-based variables. Thus, the number of the related features was reduced from 561 to 218 and the 218 features would be used as third level subset.

For evaluating the performance of the RFIM-SpLPP, a set of experiments were carried out using the HAR dataset as mentioned before. They consisted of learning CHMMs with different feature selection techniques and then comparing their performance in terms of test data error. In order to test the performance of our method, in the experiments, we used many dimension reduction algorithms, such as RFIM, PCA, LPP, Correlation and SpLPP. As we can see from Fig 2, in the first stage, we used RFIM (132 features), LPP (118 features), PCA (118 features) and Correlation (232 features) to reduce the dimension of sensor data. The result is that the RFIM performed well than others. At the second stage, we compared the methods of RFIM (132 features), PCA (118 features), SpLPP (97 features) and LPP (118 features). From the Fig 3 and we can be easy to see that RFIM (132 features) lead to the better performance. At the third stage, for the different activities, we selected the different dimension reduction methods (Gyro-based(218 features), RFIM-SpLPP(13 features), RFIM-LPP(25 features), RFIM-PCA(25 features)) and we can see that the accuracy rate of the Gyro-based and RFIM-SpLPP are significantly higher than that of other methods. The detailed result of these comparison are depicted in Figs 4 and 5.

We show the confusion matrix of the classification results on test data using TSCHMMs in Table 1. The table evidently reveals that for this classifier of TSCHMMs is sufficient for achieving optimum performance. The test error remains stable and our method for the Walking activity and the Laying activity has high-performing, especially for Laying activity for which

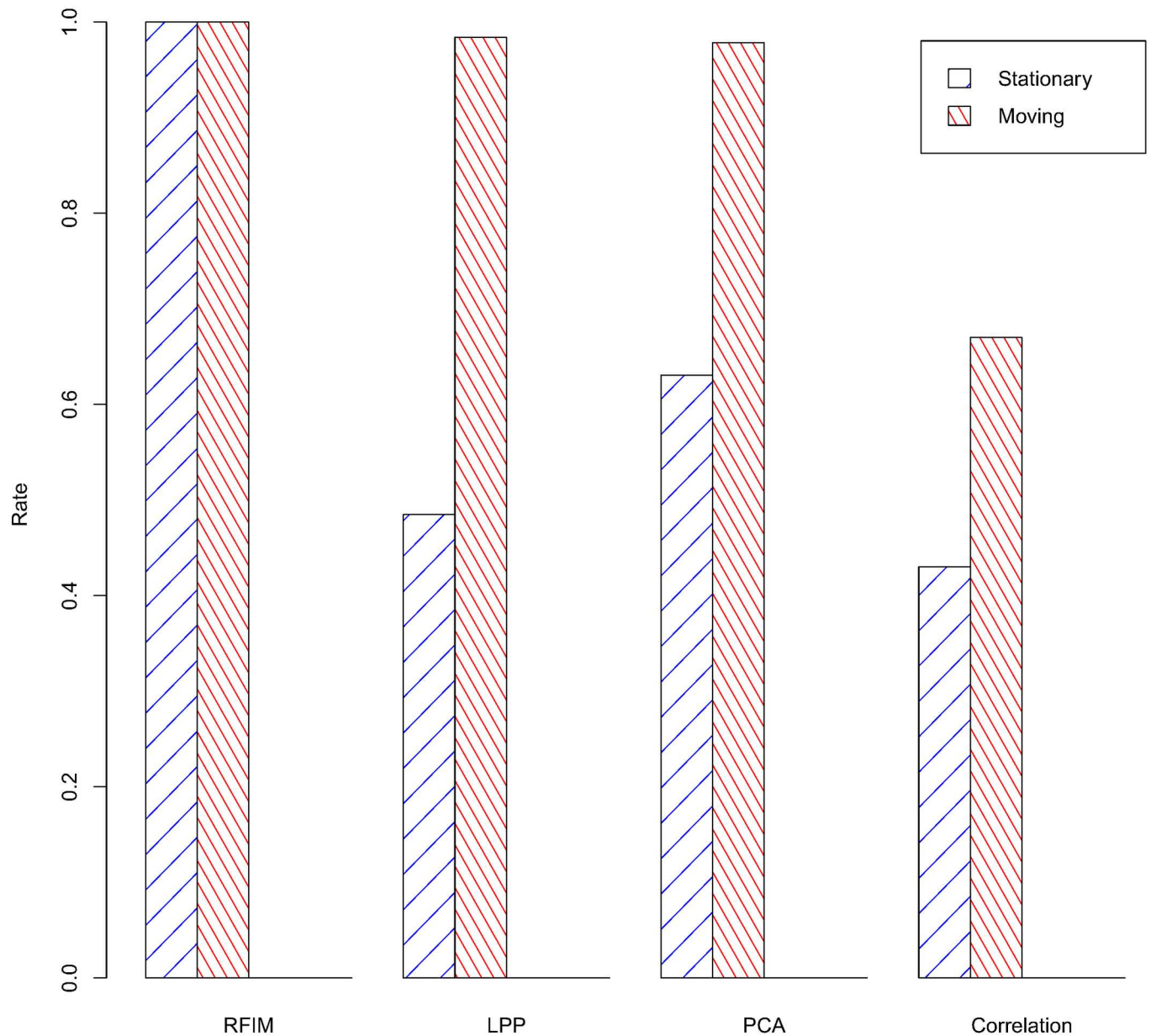


Fig 2. Coarse Classification Rate of Different Feature Selection Techniques.

doi:10.1371/journal.pone.0166567.g002

the performance achieved complete proper recognition. Moreover, it is also seen from the table that Sitting was the most difficult activity for classification because the physical location of Sitting is between the position of Laying and Standing, thus it is easy to be recognized as other two activities. It also shows that latter activities are more easily confused and the main cause might be the physical location of sensor where volunteer worn to perform the activities.

The Fig 6 shows the recognition results with different number of features obtained from the SpLPP. We can see that there is a large difference in the accuracy rate of different features number, and the reason is: SpLPP optimally preserves the geometric features of the original data, if d is large, the mapping result will contain too much noise, similarly, if d is small, the

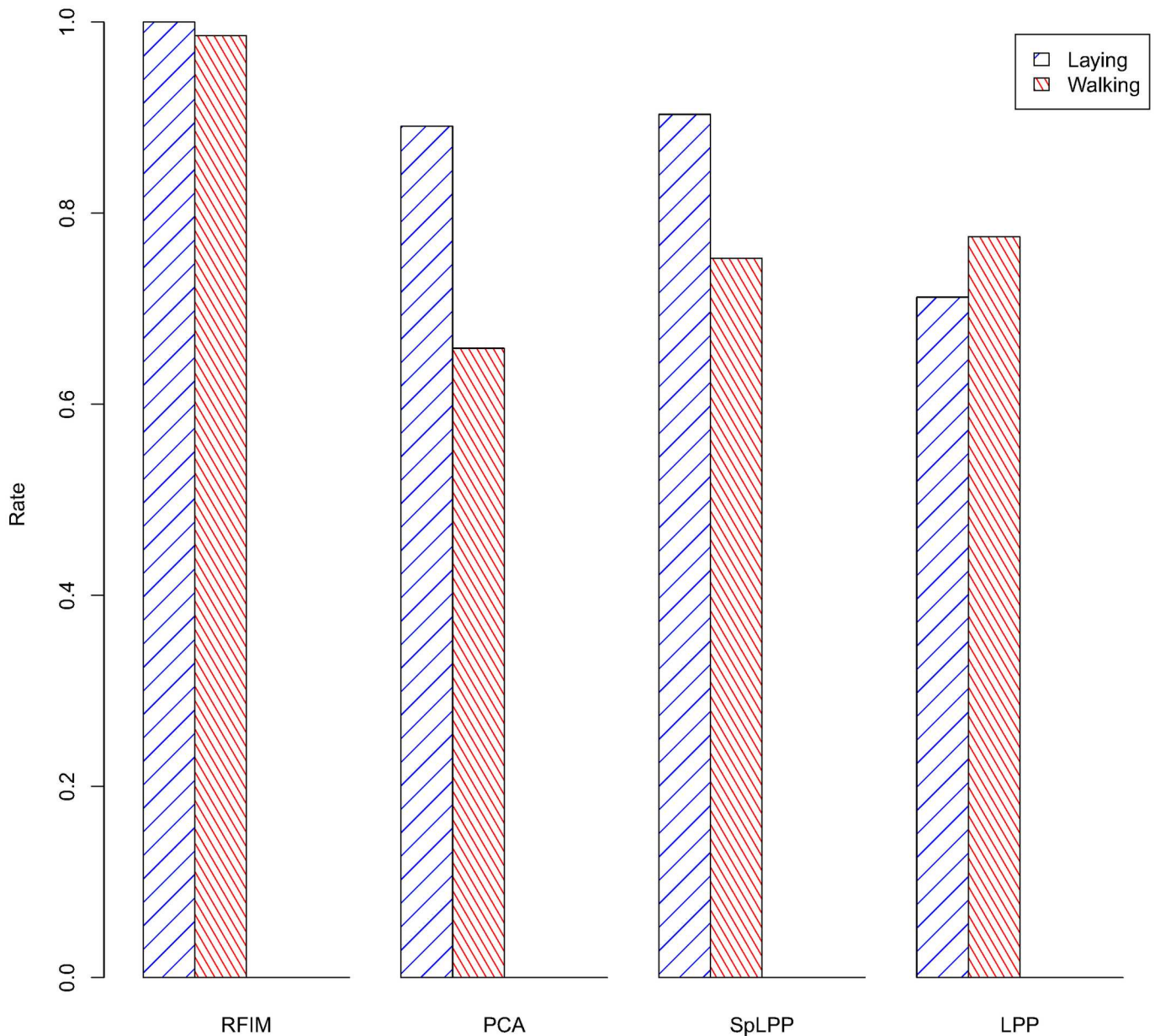


Fig 3. Fine Classification Rate of Different Feature Selection Techniques.

doi:10.1371/journal.pone.0166567.g003

different points may overlap in the low dimensional space. So in our simulation experiment, 13 is the optimal feature number.

We have compared the proposed method with many other statistical techniques such as RF, conventional HMM and Two-level HMM. The classification result of those approaches is depicted in Table 2, where estimates of the precision is given. 132 features were used to evaluate the performance of those methods. The figure illustrates that the accuracy of third-stage CHMMs and two-stage CHMMs vary slightly for Laying activity, Standing activity and Walking activity. But for the Downstairs activity, the Upstairs activity and Sitting activity, the third-stage CHMMs perform better. Therefore, our proposed method of TSCHMM is more competitive with the other algorithms.

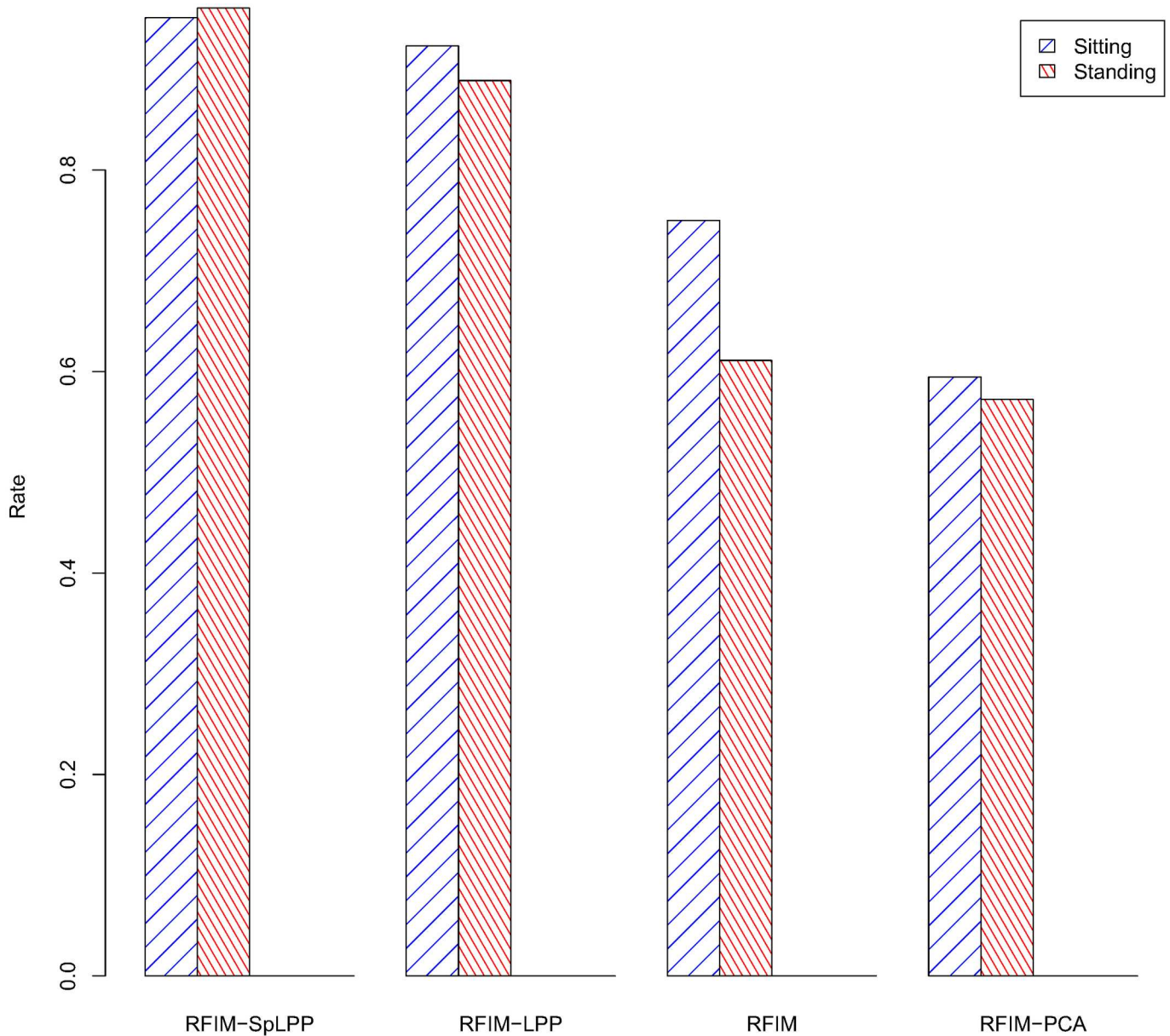


Fig 4. Accurate Classification Rate of sitting and standing activities.

doi:10.1371/journal.pone.0166567.g004

Discussion

It is novel and effective for the three-stage continuous hidden Markov model method to recognize the human daily activities based on smartphone sensor data. We have taken into account of the human activity's characteristics and propose the innovative combination method for human activity recognition. The TSCHMMs use less number of features and obtain evident high accuracy. From our simulation study, it is observed that as the strength of the relationship between the dimensionality reduction and hierarchical classifier, the performance of our method combined with dimensionality reduction and hierarchical classifier is considerably and consistently better than the ones only using classifier. In our method, we have chosen the RFIM-SpLPP technology. In our experiments, RFIM-SpLPP was compared with the method

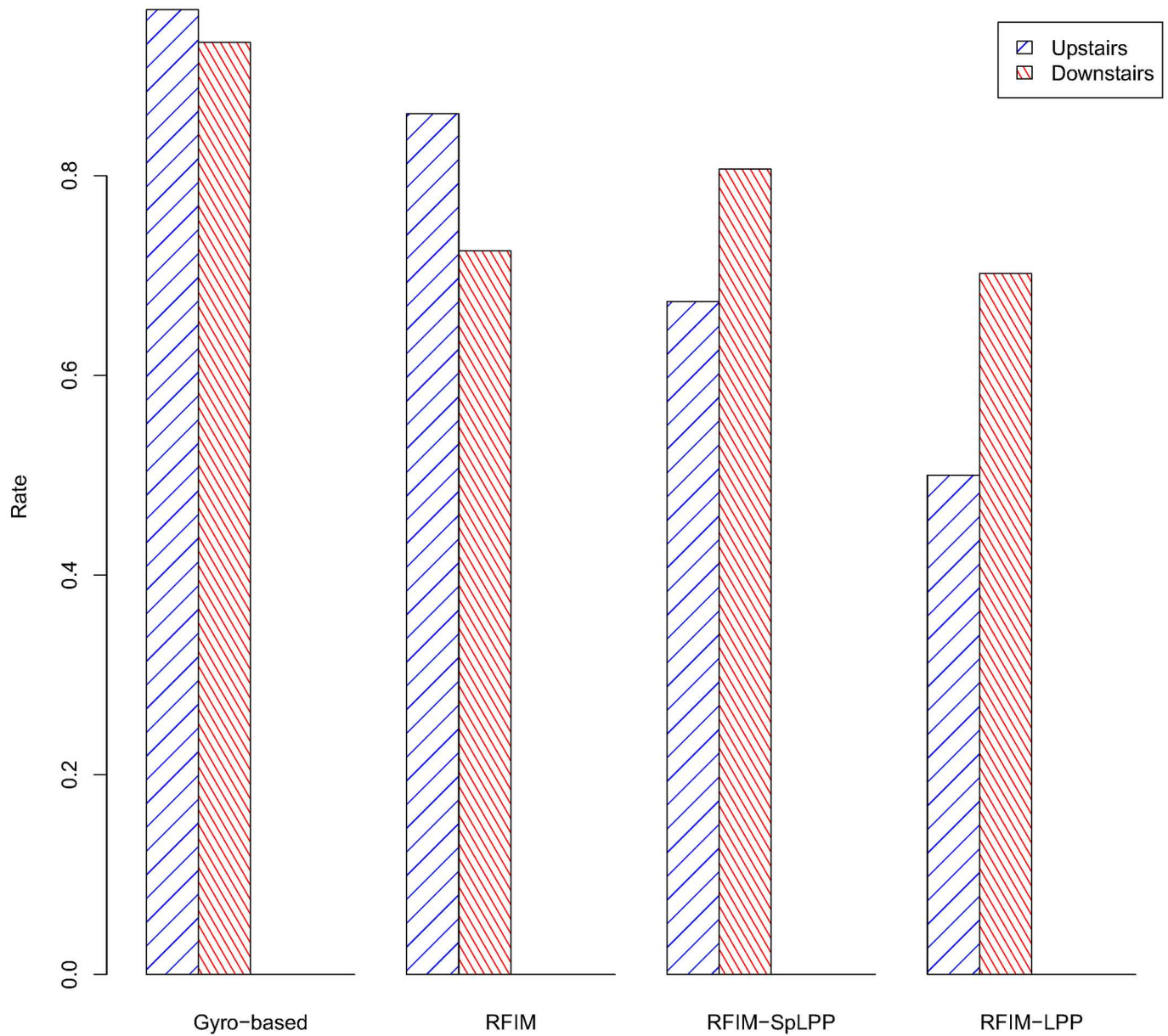


Fig 5. Accurate Classification Rate of upstairs and downstairs activities.

doi:10.1371/journal.pone.0166567.g005

Table 1. Confusion Matrix of Classification Results on Test Data Using TSCHMMs.

	walking	upstairs	downstairs	sitting	standing	laying	precision
walking	489	7	0	0	0	0	98.59%
upstairs	0	455	16	0	0	0	96.60%
downstairs	0	28	392	0	0	0	93.33%
sitting	0	0	0	467	24	0	95.11%
standing	0	0	0	21	511	0	96.07%
laying	0	0	0	0	0	537	100%

doi:10.1371/journal.pone.0166567.t001

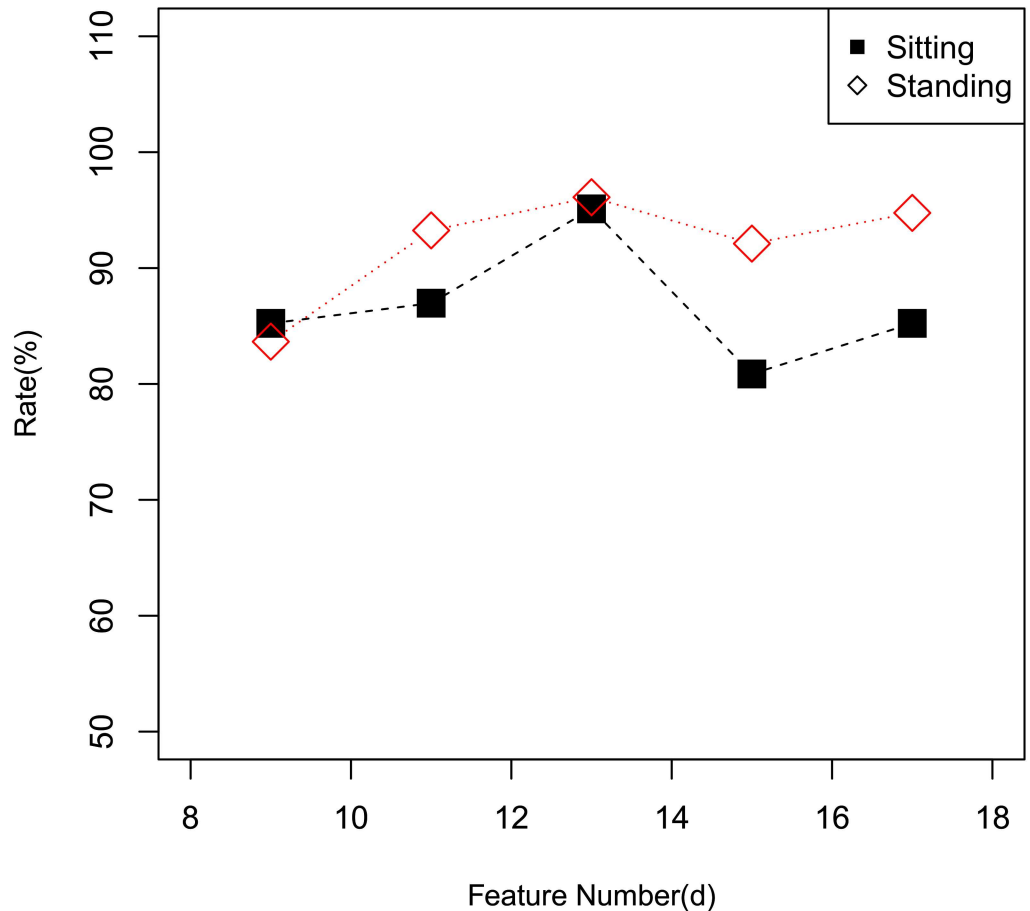


Fig 6. Accuracy Rate of SpLPP with Different Feature Number.

doi:10.1371/journal.pone.0166567.g006

only using RFIM or SpLPP. The result is that RFIM-SpLPP is indeed the better. Hence, with appropriate reduced feature set, the TSCHMMs can obtain the equivalent classifier performance and outstanding recognition rate.

We have examined the result of the classification performance for each class. This is one key guarantee that the probability of moving activity appearing in the stationary subset or stationary activity existing in the moving subset is 0% when differentiating these activities at the first stage. In addition, there is an obvious misclassification for Sitting and other activities attributed to the physical location of the device and our result of coarse classification is the essential steps to discrimination of these non-moving activities. At the second-stage CHMMs,

Table 2. Resulting Recognition Rates of Different Classifiers.

	Third-stage CHMM	Two-stage CHMM	HMM	RF
Laying	100%	99.33%	94.79%	100%
Standing	96.07%	96.43%	77.07%	90.41%
Sitting	95.11%	90%	44%	85.34%
Down-stairs	93.33%	86.24%	81.43%	84.52%
Up-stairs	96.6%	92.36%	77.71%	89.81%
Walking	98.59%	96.13%	87.1%	97.18%

doi:10.1371/journal.pone.0166567.t002

we distinguished the special activities of walking and laying. This reduced the recognition burden for the third-stage. Based on the studies at the previous two stages, the accurate recognition is more simple which are only two activities. Through the seamless convergence of this three stages, the recognition rate for each activity has been significantly improved.

Although the use of accelerometer and gyroscope sensor could enhance the performance of the combination algorithm, it is unrealistic to believe that the general public will uninterruptedly perform the different kind of activities because of the the complex patterns of daily life, which is the daily behavior habits intergrating the moving activity and stationary activity. In our experiment, we only used Gyro-based data to recognize the activities attributed to the property of Sitting-Standing subclass. Thus, we repute that the method can reduce the complexity of the experiment and improve the accuracy of recognition. Since the smartphone is always sharing the information and services with other applications, this is advantageous so that we can perform the same study with the single specialized device in the future.

Conclusion

In recent years, HAR of smartphone sensor data has received much attention. Generally, HAR of smartphone sensor data has its distinct advantages: firstly, it can continuously record information of the subjects; secondly, it is cheap and convenient for ordinary people to use this type of HAR. We think that HAR of smartphone sensor data could be helpful for physical health management of elderly people living alone. Furthermore, we can imagine that it could also be helpful for psychological health management. However, these applications need to get high accuracy of HAR. If HAR of sensor data has low accuracy, these high level applications based on human activities would have higher error rate, such as psychological health management. The three-stage continuous hidden Markov model (TSCHMM) approach want to recognize the general activities on smartphone sensor data with high accuracy. As we know, CHMMs are professional to deal with the time-series data, such as accelerometer and gyroscope sensor data. On the other hand, human activities can usually be processed as structure objectives, which can help us design a multi-level classifier. As has been argued, the TSCHMM includes these ideas. In the experiments mentioned above, the third-stage structure obviously reduced the number of features, thereby reduced the time complexity and saved the memory space. We also showed that the feature reduction technique of sparse local preserving projections combined with the feature selection method of random forest variable importance measures and other domain knowledge is effective in discovering the most available features.

Overall, our algorithms have achieved great performance. But for the future research, we still need to do more. There is much room to find a more effective method for feature selection and feature reduction. Analyzing the essential attribute of the obtained feature thoroughly can help to reduce the space and time complexity. What's more, under the rapid development of cyber-physical systems and the Internet of things, there are lots of new research topics with respect to body sensor data [26]. For instance, Lin *et al.* [27] applied the dynamic noise threshold technology to privacy protection, which may be more suitable for big data in body sensor network. Furthermore, it can also be a valuable research topic to develop the application which can online run on the smartphone platform with multimedia technology, such as touch-less interactive augmented reality game [28] and mobile health application based on virtual reality technology [29].

Author Contributions

Conceptualization: XZ.

Data curation: HQ.
Formal analysis: XZ HQ.
Methodology: XZ HQ.
Project administration: XZ.
Resources: XZ.
Software: HQ.
Supervision: XZ.
Visualization: HQ.
Writing – original draft: HQ XZ.

References

1. Kwapisz JR, Weiss GM, Moore SA. Activity recognition using cell phone accelerometers. *SIGKDD Explorations*. 2010; 12(2):74–82. doi: [10.1145/1964897.1964918](https://doi.org/10.1145/1964897.1964918)
2. Ekholm J, Fabre S. Forecast: Mobile data traffic and revenue, worldwide, 2010–2015. In: *Gartner Mobile Communications Worldwide*; 2011.
3. McIlroy T. Mobile strategies for digital publishing: a practical guide to the evolving landscape. *The Future of Publishing*; 2015.
4. Cook DJ, Das SK. Pervasive computing at scale: Transforming the state of the art. *Pervasive and Mobile Computing*. 2012; 8:22–35. doi: [10.1016/j.pmcj.2011.10.004](https://doi.org/10.1016/j.pmcj.2011.10.004)
5. Anguita D, Ghio A, Oneto L, Parra X, Reyes-Ortiz, JL. A public domain dataset for human activity recognition using smartphone. *European Symposium on Artificial Neural Networks(ESANN)*. 2013.pp. 437–442.
6. Wei W, Fan X, Song H, Fan X, Yang J. Imperfect information dynamic stackelberg game based resource allocation using hidden Markov for cloud computing. *IEEE Transactions on Services Computing*. 2016; PP(99):1–1. doi: [10.1109/TSC.2016.2528246](https://doi.org/10.1109/TSC.2016.2528246)
7. Li S, Qin Z, Song H. A temporal-spatial method for group Detection, locating and tracking. *IEEE Access*. 2016; 4: 4484–4494. doi: [10.1109/ACCESS.2016.2600623](https://doi.org/10.1109/ACCESS.2016.2600623)
8. Cho SB, Lee YS. Activity recognition using hierarchical hidden Markov models on a smartphone with 3D accelerometer. *Hybrid Artificial Intelligent Systems*, IOS Press. 2012; 6678: 460–467.
9. Sharma A, Lee YD, Chung WY. High accuracy human activity monitoring using neural network. *International Conference on Convergence and Hybrid Information Technology*. 2008.pp. 430–435.
10. Anguita D, Ghio A, Oneto L, Parra X, Reyes-Ortiz JL. Human activity recognition on smartphones using a multiclass hardware-friendly support vector Machine. *International Conference on Ambient Assisted Living and Home Care*.2012.pp.216–223.
11. Ronao CS, Cho SB. Human activity recognition using smartphone sensors with two-stage continuous hidden Markov models. *International Conference on Natural Computation (ICNC)*, 2014.pp.681–686.
12. Archer K, Kimes R. Empirical characterization of random forest variable importance measures. *Computational Statistics & Data Analysis*. 2008; 52: 2249–2260. doi: [10.1016/j.csda.2007.08.015](https://doi.org/10.1016/j.csda.2007.08.015)
13. Natarajan P, Nevatia R. Hierarchical multi-channel hidden semi Markov models. *International Joint Conference on Artificial Intelligence*. 2007.pp.2562–2567.
14. Breiman L. Random forests. *Machine Learning*. 2001; 45(1): 5–32. doi: [10.1023/A:1010933404324](https://doi.org/10.1023/A:1010933404324)
15. Hastie T, Tibshirani R, Friedman JH. *Random forests*. *The Elements of Statistical Learning*, Springer; 2003.
16. Genuer R, Poggi JM, Tuleau-Malot C. Variable selection using random forests. *Pattern Recognition Letters*. 2010; 31: 2225–2236. doi: [10.1016/j.patrec.2010.03.014](https://doi.org/10.1016/j.patrec.2010.03.014)
17. Lee JM. *Introduction to topological manifolds*. Springer-Verlag; 2003.
18. He X, Niyogi P. Locality preserving projections. *Advances in neural information processing system*, 16. 2003.
19. Zhang Y, Zhang P. Feature dimensionality reduction of wear particle images based on semi-supervised locality preserving projection. *Journal of Central South University, China*. 2015; 46: 2937–2943

20. Zheng Z, Huang X, Chen Z, He X, Liu H, Yang J. Regression analysis of locality preserving projections via sparse penalty. *Information Sciences*. 2015; 303: 1–14. doi: [10.1016/j.ins.2015.01.004](https://doi.org/10.1016/j.ins.2015.01.004)
21. Yu J. Bearing performance degradation assessment using locality preserving projections and Gaussian mixture models. *Mechanical Systems & Signal Processing*. 2011; 25: 2573–2588. doi: [10.1016/j.ymssp.2011.02.006](https://doi.org/10.1016/j.ymssp.2011.02.006)
22. Bilmes J. What HMMs can do. U. Washington Tech Report. 2002. Available: <https://www.ee.washington.edu/techsite/papers/documents/+UWEETR-2002-0003.pdf>
23. Rabiner L. A tutorial on hidden Markov models and selected applications in speech recognition. *Proc. of IEEE*. 1989; 77(2):257–286. doi: [10.1109/5.18626](https://doi.org/10.1109/5.18626)
24. Huynh N. Prediction model for smartphone activity behaviour. Available: <http://vodkhang.com/wp-content/uploads/2013/04/Prediction-Model-for-Smartphone-Behaviour.pdf>
25. Lichman M. Human activity recognition using smartphones data set. UCI machine learning repository. Irvine, CA: university of California, school of information and computer science. 2013. <https://archive.ics.uci.edu/ml/datasets/Human+Activity+Recognition+Using+Smartphones>
26. Song H, Rawat D, Jeschke S, Brecher C, editors. *Cyber-physical systems: foundations, principles and applications*. Boston: Academic Press; 2016
27. Lin C, Song Z, Song H, Zhou Y, Wang Y, Wu G. Differential privacy preserving in big data analytics for connected health. *Journal of medical systems*. 2016; 40(4): 1–9. doi: [10.1007/s10916-016-0446-0](https://doi.org/10.1007/s10916-016-0446-0) PMID: [26872779](https://pubmed.ncbi.nlm.nih.gov/26872779/)
28. Lv Z, Halawani A, Feng S, Ur Rehman S, Li H. Touch-less interactive augmented reality game on vision-based wearable device. *Personal and Ubiquitous Computing*. 2015; 19: 551–567. doi: [10.1007/s00779-015-0844-1](https://doi.org/10.1007/s00779-015-0844-1)
29. Lv Z, Chirivella J, Gagliardo P. Bigdata oriented multimedia mobile health applications. *Journal of medical systems*. 2016; 40:1–10. doi: [10.1007/s10916-016-0475-8](https://doi.org/10.1007/s10916-016-0475-8) PMID: [27020918](https://pubmed.ncbi.nlm.nih.gov/27020918/)