Available online at www.sciencedirect.com
ScienceDirect
April 2017, 24(2): 31–37
www.sciencedirect.com/science/journal/10058885
The Journal of China Universities of Posts and Telecommunications
http://jcupt.bupt.edu.cn

Compressed sensing method for human activity recognition using tri-axis accelerometer on mobile phone

Song Hui (✉), Wang Zhongmin

School of Computer Science and Technology, Xi'an University of Posts and Telecommunications, Xi'an 710121, China

Abstract

The diversity in the phone placements of different mobile users' dailylife increases the difficulty of recognizing human activities by using mobile phone accelerometer data. To solve this problem, a compressed sensing method to recognize human activities that is based on compressed sensing theory and utilizes both raw mobile phone accelerometer data and phone placement information is proposed. First, an over-complete dictionary matrix is constructed using sufficient raw tri-axis acceleration data labeled with phone placement information. Then, the sparse coefficient is evaluated for the samples that need to be tested by resolving L1 minimization. Finally, residual values are calculated and the minimum value is selected as the indicator to obtain the recognition results. Experimental results show that this method can achieve a recognition accuracy reaching 89.86%, which is higher than that of a recognition method that does not adopt the phone placement information for the recognition process. The recognition accuracy of the proposed method is effective and satisfactory.

Keywords activity recognition, compressed sensing, mobile phone accelerometer, phone placements

1 Introduction

Smart mobile phones in which an accelerometer is embedded can sample the tri-axis acceleration data of mobile users in real time, which allows instantaneous recognition of their activities, such as standing, walking, and running. As compared to an activity recognition method that uses specified wearable body sensors, that which uses a mobile phone's accelerometer has many advantages. For example, it does not require users to wear additional devices and can record and process activity data within mobile phones for a wide range of applications, including daily activity monitoring, intelligent health assistance, falling detection for elderly people, etc. Therefore, mobile user activity recognition has become an active topic in the mobile computing and ubiquitous computing research field [1–3]. In recent years, researchers have proposed different methods that use the tri-axis acceleration data of mobile phones to recognize human activities. Most of these methods extract features from the captured acceleration signals and then build a classification model to recognize various activities. For example, Parviainen et al. proposed a Bayesian model [4], Lee et al. proposed a mixture-of-experts model [5], Deng et al. proposed a reduced kernel extreme learning machine model [6], Büber et al. proposed a *k*-nearest neighbor (KNN) model [7], and Zeng et al. proposed a convolutional neural networks model [8] to recognize activity.

It should be noted that the place in which different users habitually carry their mobile phone differs, and therefore, the phone placements are diverse during sampling of the tri-axis acceleration data. This uncertainty increases the difficulty of activity recognition, because the data sampled may be completely different when the user is performing the same activity but carrying the phone in different body locations. Most recognition methods calculate the synthetic acceleration data by combing tri-axis acceleration data, and use synthetic acceleration data to

Received date: 19-09-2016
Corresponding author: Song Hui, E-mail: songhui@xupt.edu.cn
DOI: 10.1016/S1005-8885(17)60196-1

avoid the problem caused by the varying placement of mobile phones. To address this problem, various methods have been proposed using different approaches. For example, Wang et al. proposed a method to transfer raw tri-axis acceleration data into a different coordinate system to obtain higher recognition accuracy [9]. Zhu et al. proposed a method that uses the similarity of activities to achieve placement-independent results [10]. Wang et al. proposed a method that uses the fast Fourier transformation (FFT) curve to achieve a result that is placement-independent [11]. However, according to the results of our study and experiments, we argue that phone placement information, when used appropriately, can be a factor that facilitates the recognition of human activity. Further, in our method the additional operations that in order to achieve activity recognition to exclude the placement information are not necessary.

In this paper, we propose a solution for recognizing human activity by means of a compressed sensing method using both acceleration data and phone placement information. Compressed sensing theory was originally used to reconstruct a signal using limited or incomplete samples if the signal is sparse in a certain transformation domain [12]. It can also be applied to pattern recognition fields, such as image, face, and speech recognition. Studies have been conducted on human activity recognition in which compressed sensing theory was applied. Zhang et al. proposed a sparse representation method to recognize human activity that uses wearable sensors data [13]. AKimura et al. proposed a compressed sensing method for human activity sensing that uses mobile phone accelerometers data [14]. Xu et al. proposed a compressed sensing method to recognize human activity in wearable body sensor networks [15]. In our previous study, we developed a compressed sensing method to recognize human activity that uses mobile phone acceleration data, which achieved satisfactory results [16]. However, we did not introduce phone placement information into this recognition method. In the present study, we took advantage of phone placement information for activity recognition and achieved a method that yields an even better recognition rate than that of the former method. Our experimental results show that by using the proposed method five human activities (standing, walking, running, walking upstairs, and walking downstairs) can be recognized with an accuracy rate of up to 89.86% when the mobile phones are carried in three different places (in the hand, trouser pocket, and handbag).

The remainder of this paper is organized as follows. In Sect. 2, the theory of compressed sensing, and the existing work on human activity recognition by using a compressed sensing method are introduced. In Sect. 3 the human activity framework are described. The experimental results and analysis are presented in Sect. 4. Finally, the conclusions are shown in Sect. 5.

2 Related work

2.1 Compressed sensing theory

Compressed sensing theory exploits the fact that many natural signals are sparse and compressible in the sense that their representations are concise when expressed in an appropriate basis. Random observation matrixes are used to project raw data into the required transformation domain. Suppose that $\boldsymbol{\alpha}$ is a vector of unknown, $\boldsymbol{y}$ denotes the available observed measurements, and $\boldsymbol{A}$ is the data matrix to describe the relation between $\boldsymbol{\alpha}$ and $\boldsymbol{y}$. Then, we have

$$\boldsymbol{y} = \boldsymbol{A\alpha} \tag{1}$$

where $\boldsymbol{y} \in \mathbb{R}^{N\times 1}$, $\boldsymbol{A} \in \mathbb{R}^{N\times M}$ and $\boldsymbol{\alpha} \in \mathbb{R}^{M\times 1}$.

For applications where the number of measurements is much smaller than the number of unknowns ($N<<M$), data matrix $\boldsymbol{A}$ is also called an over-complete dictionary matrix. In this case, Eq. (1) represents an underdetermined system and $\boldsymbol{\alpha}$ cannot be uniquely reconstructed from matrix $\boldsymbol{A}$ and measurements $\boldsymbol{y}$. However, in situations where $\boldsymbol{\alpha}$ is sufficiently sparse, we can reconstruct $\boldsymbol{\alpha}$ with the L0 sparsity formulation to obtain the approximate solution of $\boldsymbol{\alpha}$

$$\left.\begin{aligned} &\hat{\boldsymbol{\alpha}} = \arg\min \|\boldsymbol{\alpha}\|_0 \\ &\text{s.t.} \\ &\quad \boldsymbol{y} = \boldsymbol{A\alpha} \end{aligned}\right\} \tag{2}$$

Eq. (2) represents a determined system and its solution is stable. However, it is intractable, because it is an NP hard problem. The traditional heuristic to approximate the sparsity L0 is to use the minimal energy L2 instead. It is well-known that L2 is a least square formation and can be efficiently resolved. As the energy minimization L2 is not necessarily equivalent to the sparsity L0 in most cases, with high probability the solution of Eq. (2) is the same as the L1 minimization:

$$\left.\begin{aligned} &\hat{\boldsymbol{\alpha}} = \arg\min \|\boldsymbol{\alpha}\|_1 \\ &\text{s.t.} \\ &\quad \boldsymbol{y} = \boldsymbol{A\alpha} \end{aligned}\right\} \tag{3}$$

It has been proved that this L1 minimization can be formulated as a convex optimization problem [12]. In this case, the optimization problem is well-posed and can be resolved in polynomial time.

2.2 Compressed sensing method for human activity recognition

The existing compressed sensing based human activity recognition methods can be described as follows. Tri-axis acceleration data a_x, a_y, and a_z are sampled at a given sampling rate. For each moment, the synthetic acceleration $\boldsymbol{a}$ is calculated as $|\boldsymbol{a}|=\sqrt{a_x^2+a_y^2+a_z^2}$. A sample is composed of a set of synthetic acceleration data during a period of time, from which certain features are extracted. Consider that K different human activities need to be recognized. Each activity has M_i (i=1,2,...,K) training samples. For each sample, N features are extracted. All these features are used to construct a matrix $\boldsymbol{A}$ with N rows and M columns:

$$\boldsymbol{A}=[\boldsymbol{A}_1,\boldsymbol{A}_2,...\boldsymbol{A}_i,...,\boldsymbol{A}_K]\in\mathbb{R}^{N\times M} \tag{4}$$

where $\boldsymbol{A}_i$ is a sub-matrix for activity i, $\boldsymbol{A}_i\in\mathbb{R}^{N\times M_i}$, and $M=M_1+M_2+...+M_K$.

According to the compressed sensing theory described in the previous section, matrix $\boldsymbol{A}$ is the over-complete dictionary matrix for the K activities. Following Eq. (1), any given unknown input $\boldsymbol{y}$ ($\boldsymbol{y}\in\mathbb{R}^{N\times 1}$) that needs to be recognized can be represented as a linear span of the over-complete dictionary matrix

$$\boldsymbol{y}=\boldsymbol{A\alpha} \tag{5}$$

where $\boldsymbol{\alpha}\in\mathbb{R}^{M\times 1}$ is the sparse coefficient. Using this formulation, the class membership of $\boldsymbol{y}$, which is encoded as the sparsest solution of the underdetermined system given in Eq. (1), can be evaluated by resolving Eq. (3).

The above approach is used in existing human activity recognition methods based on compressed sensing theory. As compared with the existing studies, ours differs in the following two aspects:

1) To set up an over-complete dictionary matrix, instead of calculating synthetic accelerations and extracting features, we use raw tri-axis acceleration data sampled by the mobile phone accelerometer. This eliminates the time and energy consumption required for calculating synthetic acceleration data and extracting features, resulting in a better performance, which is valuable, in particular for mobile applications that run under limited central processing unit (CPU) resources and energy supply conditions.

2) We argue that the phone placement information can be useful for activity recognition. Include the phone placement information in the over-complete dictionary matrix can achieve better recognition accuracy. Therefore, our method requires no additional operations to eliminate the effects of different users' diverse phone placements as compared to existing recognition methods. This also helps to reduce the calculation time and energy consumption.

3 Framework

In this section, we present our proposed framework for human activity recognition by means of raw tri-axis acceleration data and phone placement information. As Fig. 1 shows, the framework consists of three components: construction of the over-complete dictionary matrix, resolution of the sparse coefficient via L1 minimization, and residual value calculation to recognize the activity.

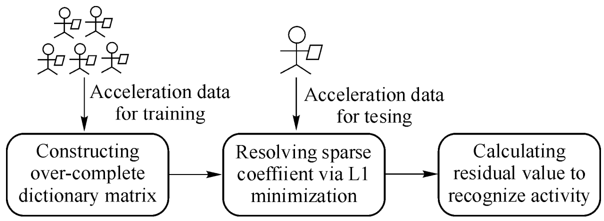

Fig. 1 Activity recognition framework

3.1 Construction of an over-complete dictionary matrix and resolution of sparse coefficients

It has been proved that human activity signals captured by a mobile phone accelerometer are sparse in certain transformation bases [14], and therefore, we can use sparse data representation to describe human activity data based on compressed sensing theory. Suppose there are K different human activities that need to be recognized. For each activity, tri-axis acceleration data a_x, a_y, and a_z are sampled at one moment. We define vector $\boldsymbol{a}$ as

$$\boldsymbol{a}=[a_x,a_y,a_z];\quad \boldsymbol{a}\in\mathbb{R}^3 \tag{6}$$

which is one observed vector. During a time period, n observed vectors are sampled. We define

$$\boldsymbol{V}=[\boldsymbol{a}_1,\boldsymbol{a}_2,...,\boldsymbol{a}_n];\quad \boldsymbol{V}\in\mathbb{R}^N, N=3n \tag{7}$$

as one data sample. For each human activity, we collect training samples for L different user mobile phone placements, such as in the hand, trouser pocket, and handbag. Suppose M_{i_j} $(i=1,2,...,K,\ j=1,2,...,L)$ training samples are collected for each activity with each placement. We construct matrix A_i $(i=1,2,...,K)$ for each activity in the format

$$\begin{aligned}A_i = [A_{i_1}, A_{i_2}, ..., A_{i_j}, ..., A_{i_L}] = [[V_{i_{11}}, V_{i_{12}}, ..., V_{i_1 M_{i_1}}], \\ [V_{i_{21}}, V_{i_{22}}, ..., V_{i_2 M_{i_2}}], ..., [V_{i_{j1}}, V_{i_{j2}}, ..., V_{i_j M_{i_j}}], ..., \\ [V_{i_{L1}}, V_{i_{L2}}, ..., V_{i_L M_{i_L}}]]\end{aligned} \tag{8}$$

where $i=1,2,...,K$, $j=1,2,...,L$, and $A_{i_j} \in \mathbb{R}^{N \times M_{i_j}}$ is a sub-array combining all the sample data for activity *i* with placement *j*. Based on compressed sensing theory, matrix $A_i \in \mathbb{R}^{N \times M_i}$ is the over-complete dictionary matrix for activity *i*.

Finally, we define a matrix A with N rows and M columns:

$$A = [A_1, A_2, ..., A_i, ..., A_K] \in \mathbb{R}^{N \times M} \tag{9}$$

where $M = M_1 + M_2 + ... + M_K$. Here, A is the over-complete dictionary matrix for all the activities with all placements. According to the description in the previous section, for any observed data y, we have

$$y = A\alpha \tag{10}$$

where $\alpha = [0, 0, ..., 0, \alpha_{i1}, \alpha_{i2}, ..., \alpha_{iM_i}, 0, ..., 0]^{\mathrm{T}} \in \mathbb{R}^{M \times 1}$ is the sparse coefficient, which can be evaluated by resolving Eq. (3). Here, random observation projection should be performed for both A and y before the resolution process.

As compared with existing human activity recognition methods that do not using phone placement information, our method uses data for different phone placements to construct the over-complete dictionary matrix, which results in a more sufficient over-complete dictionary matrix with more information of activities. This can facilitate a better solution of the sparse coefficient to improve the activity recognition rate based on the compressed sensing theory. Second, we use raw tri-axis acceleration data sampled by the mobile phone accelerometer to construct the over-complete dictionary matrix, instead of calculating synthetic acceleration data and extracting features for this purpose as is done in the existing recognition methods. This reduces the calculation time and energy consumption, which results in a better recognition performance and is valuable for mobile device applications.

3.2 Residual value calculation to recognize activity

As described above, we can obtain sparse coefficient α of Eq. (3) by resolving L1 minimization as a convex optimization problem. Ideally, α is a sparse vector in which all data are zeros, except the data located in the position representing the activity that is to be recognized, which are non-zero. Thus, when we resolve sparse coefficient α, we can recognize the activity by the distribution of the non-zero data. Unfortunately, because of the effects of sampling noise and calculating deviation, the sparse coefficient α may be a little different from the ideal one for practical applications. The non-zero data may be located in places that do not match the tested activity. Fig. 2 shows the distribution for a resolved sparse coefficient α against the sparse coefficient distribution for the standing activity. This example includes a total of five activities, standing, walking, running, walking upstairs, and walking downstairs. We can see that the great majority of the data are zeros and most of the non-zero data are located at the position for the standing activity. However, a few non-zero data are located at the positions for other activities.

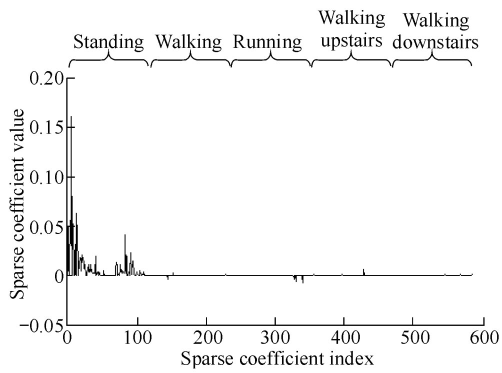

Fig. 2 Distribution of sparse coefficient values

To handle this problem, we can calculate the residual value *r* between y and the training sample vector for each activity:

$$r_i = \| y - A_i \delta_i(\hat{\alpha}) \|_2; \quad i = 1, 2, ..., K \tag{11}$$

where A_i is the sub-matrix of dictionary matrix A that represents the activity class *i* and $\delta_i(\hat{\alpha})$ denotes the sub-vector of vector $\hat{\alpha}$, which is in the same position as A_i in A. Suppose there are K activities, then, we can obtain K residual values, the minimum of which is selected to indicate the class of activity to be recognized. The reason is that the data within sparse coefficient α should all be zeros, which results in a considerably larger residual value. Table 1 shows the calculated residual values for different activities based on the sparse coefficient values shown in Fig. 1. It clearly shows that the residual value for the standing activity is the minimum, and thus, indicates that the activity in the case of this sample should be recognized as activity standing.

The activity recognition process of our solution can be summarized as the followings steps:

Step 1 Construction of over-complete dictionary matrix $\boldsymbol{A}$ by sufficient training samples of different activities and different phone placements.

Step 2 Random observation projection for over-complete dictionary matrix $\boldsymbol{A}$ and any testing sample $\boldsymbol{y}$ that needs to be recognized, and then, evaluation of sparse coefficient $\boldsymbol{\alpha}$ by resolving L1 minimization.

Step 3 Calculation of the residual value between $\boldsymbol{y}$ and the training sample vector for each activity, and selection of the activity that is represented by the minimum residual value as the activity for the testing sample.

Table 1 Residual value for different activities

Activity	Residual value
Standing	0.094 4
Walking	3.663 7
Running	3.713 8
Walking upstairs	3.667 3
Walking downstairs	3.686 7

4 Experiments and evaluation

4.1 Dataset and experimental results

We developed a program for Android mobile phones to collect tri-axis acceleration data for our experiment. Data of five types of activity (standing, walking, running, walking upstairs, and walking downstairs) were collected from 12 subjects (6 males and 6 females) whose ages ranged from 22 years to 53 years. The mobile phones used to collect acceleration data were normal mobile phones that the subjects used in their daily life, and therefore, the brand and model of the mobile phones varied. The subjects were asked to carry the mobiles phone in three places (in the hand, trouser pocket, and handbag) and we collected acceleration data for each activity. Therefore, a total of 15 combinations of activity and phone placement were used. For each activity with each placement, we collected continuous acceleration data for 10 s, and repeated the procedure 10 times to reduce accidental error. The sampling rate was 50 Hz. We adopted a validation strategy consisting of dividing the entire dataset into two parts: 80% of the data were used to construct the over-complete dictionary matrix and the remaining 20% were used for testing. We conducted two experiments in a Matlab simulation environment.

First, we calculated the recognition rate of different activities based on our proposed framework when four widely used random observation matrices, Gaussian random matrix, Bernoulli random matrix, sparse random matrix, and Hadamard random matrix, were applied. The experiment was conducted 10 times with 10 different compression ratios in order to evaluate the effects of the compression ratio on the activity recognition accuracy. L1 minimization was calculated using the L_1-MAGIC toolkit provided by Stanford University [17]. Table 2 and Fig. 3 show the experimental results.

Table 2 Recognition rate of different random observations under different compression ratios

Compression ratio/(%)	Recognition rate/(%)			
	Gaussian random matrix	Bernoulli random matrix	Sparse random matrix	Hadamard random matrix
100	89.86	86.84	86.61	83.99
90	86.32	80.51	78.35	78.35
80	82.56	75.73	76.92	70.66
70	76.07	69.86	69.34	66.55
60	71.51	66.95	67.98	66.55
50	69.97	65.24	67.24	65.81
40	69.23	64.44	65.19	64.44
30	67.81	64.44	62.39	63.76
20	66.72	62.39	61.65	60.68
10	64.56	59.60	60.97	60.28

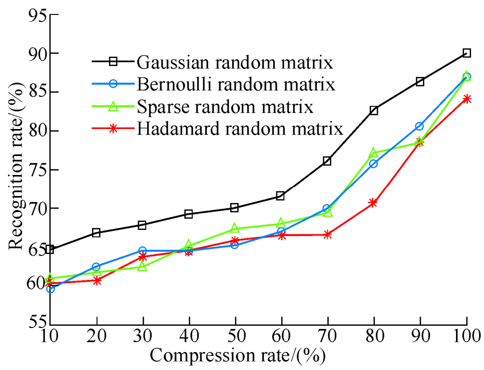

Fig. 3 Recognition rate of different random observations under different compression ratios

Second, we calculated the recognition rate of human activities achieved by a recognition method that does not use mobile phone placement information. The results were compared with those of our proposed method to evaluate the benefit of using mobile phone placement information. For this experiment, we mixed the acceleration data of each activity with different placements together to construct the over-complete dictionary matrix. Fig. 4 shows the recognition rate with and without phone placement information using the Gaussian random matrix and the Bernoulli random matrix as the random observation matrices.

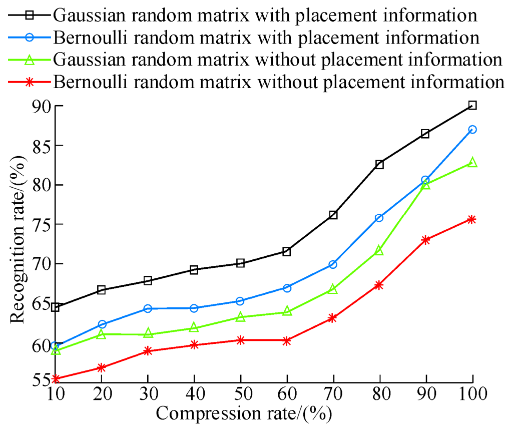

Fig. 4 Recognition rate with and without phone placement information

4.2 Recognition performance evaluation

In this section, we discuss the evaluation of the recognition performance of our framework. The results of the first experiment (Table 2 and Fig. 3) show that for our proposed method, regardless of the type of random observation matrix used, the trends of the recognition rate curves are similar. When the compression ratio is 100%, the recognition rate is the highest. When the compression ratio is decreased, the recognition rate is reduced. The reason is that when the compression ratio is reduced, the amount of data contained in the over-complete dictionary matrix is reduced, which affects the accuracy of the data reconstruction. Among the four random observation matrices we tested, the Gaussian random matrix shows the best recognition performance, regardless of the compression ratio used. The recognition performances for the remaining three random observation matrices are similar. When the compression ratio is set as 70% and 80%, the Hadamard random matrix shows the lowest recognition performance. This proves that the Gaussian random matrix is the best choice for our compressed sensing method to recognize human activity.

In Fig. 3, we can also see that, when the compression ratio is reduced from 100% to 60%, the recognition rate is decreased faster, whereas when the compression ratio is reduced from 60% to 10%, the decrease in the recognition rate is slower. When the compression ratio is 10%, the recognition rate remains around 60%. These results prove the advantage of using a compressed sensing method that allows a signal to be reconstructed with limited or incomplete samples. It also indicates that it is possible to recognize activity using only a few compressed acceleration data instead of all the acceleration data, if the recognition rate is acceptable under that condition. This is reasonable for some mobile applications that need acceleration data transferred from the mobile phone to support the servers' activity recognition. Transferal of compressed data reduces the network flows significantly in this case.

The results of the second experiment (Fig. 4) shows that the recognition rate is higher when mobile phone placement information is used than when it is not, at any compression ratio for both the Gaussian random matrix and the Bernoulli random matrix. As mentioned above, we mixed the acceleration data of each activity with different placements together to construct the over-complete dictionary matrix for the second experiment. Thus, the over-complete dictionary matrix contained more information of activities than did that used in the first experiment, which was constructed by separating the different phone placements. This results in a better recognition performance when a compressed sensing method is used to resolve L1 minimization and finally recognize the target activity. These experimental results prove the effectiveness of our recognition framework.

We used a confusion table (Table 3) to examine the experimental results in more detail. The table lists the specific recognition results of the five types of activity with the three placements when using the Gaussian random matrix with a compression ratio of 100%. In Table 3, P_1, P_2 and P_3 represent the three placements, in the hand, trouser pocket, and handbag, respectively. It can be seen in Table 3 that the recognition rate of the standing activity type is the highest, while the recognition rate of the remaining types is lower. The reason is that the curve of the acceleration data for the standing activity differs significantly from those of the acceleration data for the remaining four activities. However, the curves of the acceleration data for the activities walking, running, walking upstairs, and walking downstairs have some similarities of varying degrees. The traditional activity recognition methods extract features from raw acceleration data and use these features to recognize activity. Thus, they can avoid the problem of similarity in the raw acceleration data. Our compressed sensing method uses raw acceleration data directly to calculate the sparse coefficients to recognize the activity. Similar data may interfere with this calculation, resulting in incorrect results. On the basis of the theory of compressed sensing, this weak point can be improved from two aspects. On the one hand, the dimension of training sample M for the over-complete dictionary matrix can be increased so that the dimension N of the unknown samples is considerably smaller than M. On the other hand, the algorithm that resolves L1 minimization can be improved in order to improve the accuracy of the solved sparse coefficients, and thereby, improve the recognition rate.

Table 3 Confusion table for using Gaussian random matrix with compression ratio 100%

Activity	Placement	Standing			Walking			Running			Walking upstairs			Walking downstairs		
		P_1	P_2	P_3	P_1	P_2	P_3	P_1	P_2	P_3	P_1	P_2	P_3	P_1	P_2	P_3
Standing	P_1	116	0	1	0	0	0	0	0	0	0	0	0	0	0	0
	P_2	2	115	0	0	0	0	0	0	0	0	0	0	0	0	0
	P_3	1	0	116	0	0	0	0	0	0	0	0	0	0	0	0
Walking	P_1	3	0	0	108	0	0	0	0	0	1	0	5	0	0	0
	P_2	0	0	0	4	100	0	0	0	0	0	0	3	0	10	0
	P_3	0	0	0	0	0	107	0	0	3	0	0	5	0	2	0
Running	P_1	0	0	0	0	0	6	101	0	6	0	0	0	0	0	4
	P_2	0	0	0	0	7	0	2	105	0	0	0	0	0	3	0
	P_3	0	0	0	0	0	0	9	0	103	2	0	0	0	0	3
Walking upstairs	P_1	0	0	0	3	0	5	0	0	0	100	0	2	0	3	4
	P_2	0	0	0	0	0	0	0	0	0	1	98	0	9	9	0
	P_3	0	0	0	0	0	1	0	0	0	0	0	107	0	0	9
Walking downstairs	P_1	0	0	0	0	0	0	0	0	2	0	8	0	99	0	8
	P_2	0	0	0	5	0	0	0	0	0	3	7	0	0	102	0
	P_3	0	0	0	0	0	6	0	2	0	0	0	4	5	0	100

The accuracy of our recognition method in the real world may be affected by one factor. We considered the three most frequently used placements of mobile phones (in the hand, trouser pocket, and handbag) in our experiment for building the framework, but users may carry their phones in additional different places that we have not considered. There are two means of resolving this problem. One is to train the framework with a larger-scale dataset that includes more placement labels and the second is to use transfer learning methodology to handle new placements. In our framework, the application of transfer learning is enabled by re-constructing the over-complete dictionary matrix using newly collected data. We are considering conducting further experiments with the large-scale dataset and applying transfer learning methodology to evaluate more activities and phone placements in future work.

5 Conclusions

In this paper, we proposed a compressed sensing method for recognizing human activity in which both acceleration data and phone placement information are utilized. The addition of phone placements information into the over-complete dictionary matrix results in a better over-complete dictionary matrix with more information of activities, which can help to improve the activity recognition rate. The experimental results showed that the proposed method can achieve a recognition accuracy up to 89.86% for five human activities with three placements. This recognition accuracy is higher than that of a method that does not take into account phone placement information. Moreover, we also evaluated the effects of using different random observation matrixes. The experimental results showed that the Gaussian random matrix is the best choice for the proposed method.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (61373116).

References

1. Shoaib M, Bosch S, Durmaz Incel O D, et al. A survey of online activity recognition using mobile phones. Sensors, 2015, 15(1): 2059–2085
2. Su X, Tong H H, Ji P. Activity recognition with smartphone sensors. Tsinghua Science and Technology, 2014, 19(3): 235–249
3. Incel O D, Kose M, Ersoy C. A review and taxonomy of activity recognition on mobile phones. BioNanoScience, 2013,3(2): 145–171
4. Parviainen J, Bojja J, Collin J, et al. Adaptive activity and environment recognition for mobile phones. Sensors, 2014, 14(11): 20753–20778
5. Lee Y S, Cho S B. Activity recognition with android phone using mixture-of-experts co-trained with labeled and unlabeled data. Neurocomputing, 2014, 126: 106–115
6. Deng W Y, Zheng Q H, Wang Z M. Cross-person activity recognition using reduced kernel extreme learning machine. Neural Networks, 2014, 53: 1–7
7. Büber E, Guvensan A M. Discriminative time-domain features for activity recognition on a mobile phone. Proceedings of the IEEE 9th International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP'14), Apr 21–24, 2014, Singapore. Piscataway, NJ, USA: IEEE, 2014: 6p

To p. 71