Body Activity Recognition using Wearable Sensors

Long Cheng, IEEE/ACM Member
Research and Development Department
Kiwii Power Technology Corporation
Troy, NY,
USA
dearlongcheng@gmail.com

Chenyu You, Yani Guan
Department of Electrical, Computer and Systems Engineering
Rensselaer Polytechnic Institute
Troy, NY, USA
{youc2, guany3}@rpi.edu

Yiyi Yu
Department of Applied and Computational Math
Johns Hopkins University
Baltimore, MD, USA
yyu48@jhu.edu

***Abstract*—Efficient recognition of human body activities is of great significance in many fields. With the development of wearable sensor technology, wearable sensors are playing a more and more important role in recognizing human body activities. How to accurately and quickly recognize human body activities while saving computing resource is a major challenge to be addressed. Utilizing data sampled from wearable sensors, this paper solves the human body activity recognition problem from perspectives of both machine learning and compressed sensing. Specifically, three different machine learning algorithms, artificial neural network, support vector machine and hidden Markov model, and one compressed sensing involved algorithm, sparse representation classification method based on random projections, are used to recognize human body activities, respectively. Meanwhile, various numerical experiments based on a real-world dataset collected using wearable sensors are conducted to validate the effectiveness of these algorithms. Numerical results demonstrate that all of the four algorithms achieve satisfactory recognition performance and the sparse representation classification method based on random projections outperforms the other three machine learning algorithms.**

Keywords—wearable sensor; support vector machine; hidden Markov model; neural network; sparse representation; random projections

I. INTRODUCTION

Recognizing human body activities plays a critical role in many applications, such as elderly care, health monitoring and security scanning. With the fast development of wearable technology, the wearable sensor, which can accurately measure accelerations of body movements in real time, is becoming more and more popular in many applications. For instance, the wearable sensor can be integrated into many assistive products, making the elderly live a healthy life [1].

With the help of wearable sensors, many scientists have proposed different body activity recognition algorithms, including supervised classification algorithms such as Support Vector Machine (SVM), k-Nearest Neighbor and Random Forest, and unsupervised classification algorithms such as Hidden Markov Model (HMM), k-Means and Gaussian Mixture Models. For instance, Cheng *et al*. applied SVM, HMM and Artificial Neural Network (ANN) to recognize human activities using wearable sensors [2]. Yang *et al.* implemented a neural network classifier for off-line human activity recognition [3]. However, applying such a complicated algorithm in [3] in the embedded system is limited by the computing resource. Gao *et al*. proposed some decision tree based algorithms to identify human activities [4], but its recognition rate, which is just near eighty percent, is not satisfactory. So some more accurate algorithms are required to be proposed to address the human activity recognition problem.

During recent years, compressed sensing develops very fast and has solved many problems in more efficient ways. For instance, it was successfully used in the classification of power system faults [5], the identification of power line outages [6] and the robust face recognition [7]. More recently, some researchers started to recognize human body activities from the perspective of compressed sensing. For example, the authors in [8] represented each human activity test signal as a linear sparse combination of all the training signals and used the L1 minimization algorithm to determine the activity class of each test signal. Sparse representation and random projection were used in [9] to recognize human body activities. Orthogonal Matching Pursuit was used in [10] to recognize human activities. But its recognition rate in [10] could be further improved by using some other convex optimization methods. Authors in [11] solved human action recognition problem using sparse representation and L1 minimization to avoid the sensitivity of parameter selection in the nearest neighbor algorithm. But the computation of motion context descriptors involved in this method was inconvenient and computationally expensive. Cheng *et al.* proposed a sparse representation classification method with the consideration of random projections (SRC-RP) in order to accurately and quickly recognize human activities while saving computing resource [12]. In almost all of these compressed sensing based human activity recognition methods, sparsely representing the activity test signal in some form and efficiently solving the corresponding optimization problem are very important to discriminate different human activities.

In this paper, three different machine learning algorithms, SVM, ANN and HMM, and one compressed sensing algorithm, SRC-RP are proposed to recognize human body activities. And numerical experiments are conducted on a real-world human activity dataset collected by wearable sensors to verify the effectiveness of these algorithms.

The rest of the paper is organized as follows. Section II describes the data sampled using wearable sensors and the numerical experiments in details. Three different machine learning algorithms and one compressed sensing based algorithm for the recognition of human body activities are presented in Section III. The results of the numerical

978-1-5090-5443-5/17/$31.00 ©2017 IEEE

experiments are summarized and discussed in Section IV. And Section V concludes the paper and introduces the future work.

II. DATA AND EXPERIMENT DESCRIPTION

The recognition of human body activities depends on the features extracted from body motions since each activity can be represented using some kinds of motion features. And meanwhile, motion features can be abstractly represented using data sampled from wearable sensors. Therefore, with corresponding features extracted from wearable sensor data, different activities can be probably distinguished. The feasibility and the correctness of this idea are further verified by performing numerical experiments on a real-world dataset, which was contributed by W. Ugulino's team. In order to recognize five common human activities, sitting, sitting down, standing up, standing, and walking, this team put four wearable sensors on four person subjects at four different body locations, the right arm, the left thigh, the right ankle and the abdomen[13]. Wearing the four wearable sensors, each subject acted the five classes of activities for about 2 hours. Each of the four sensors measured the accelerations in the x, y, z directions. So 12 values were recorded at each sampling point for each subject. To improve the accuracy of the sensor measurement, all sensors were calibrated before the experiments. The total amount of sample data collected by the wearable sensors is 165,633. At the end of the experiment, the collected sample data were further discretized into one second time window, with one hundred and fifty milliseconds overlapping.

Two different human activity recognition experiments were conducted. The first one was One-against-Own, which indicates that each subject was tested against itself. In this case, every sixth data point for a subject was used as the test data and every other data for that subject was used for training. The other one was One-against-All, which indicates that training on all but one subject, and subsequently testing on that left subject.

First, text labels in the dataset were changed into integers for the convenience: 1 is used to stand for sitting; 2 indicates sitting down; 3 indicates standing; 4 indicates standing up; and 5 indicates walking.

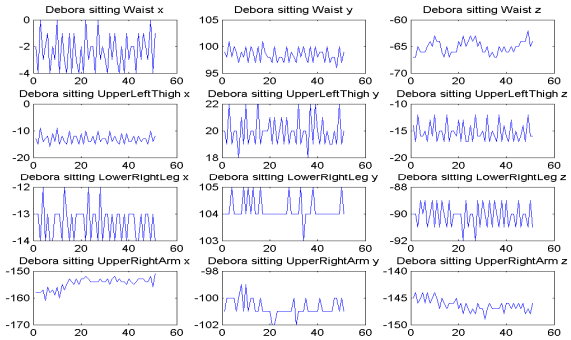

Fig. 1. Data visualization for Debora's sitting

TABLE I. DATA SUMMARY TABLE

Index	Name of the person	Label	Number of samples
2-15616	Debora	1	15615
15617-29896	Katia	1	14280
29897-44889	Wallace	1	14993
44890-50632	Jose	1	5743
50633-54179	Debora	2	3547
54180-58196	Katia	2	4017
58197-61682	Wallace	2	3485
61683-62459	Jose	2	777
62460-77399	Debora	3	14940
77400-91633	Katia	3	14234
91634-106100	Wallace	3	14468
106101-109829	Jose	3	3729
109830-113682	Debora	4	3852
113683-117392	Katia	4	3710
117393-121507	Wallace	4	4115
121508-122244	Jose	4	737
122245-135866	Debora	5	13622
135866-149422	Wallace	5	13557
149423-163459	Jose	5	14037
163460-165634	Debora	5	2175

Meanwhile, the index of the data was used to separate the dataset into the training and the testing parts. The data is summarized in TABLE I.

Data visualization examples for Debora's body activities are illustrated in Fig. 1-5, respectively.

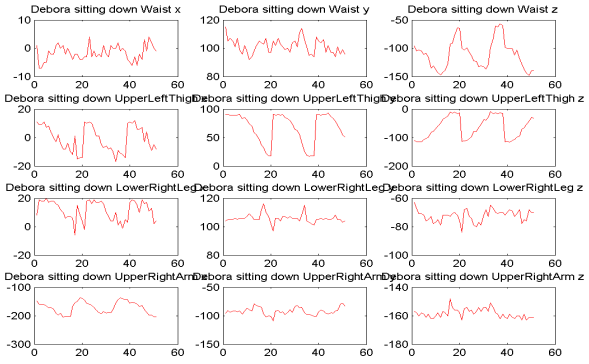

Fig. 2. Data visualization for Debora's sitting down

978-1-5090-5443-5/17/$31.00 ©2017 IEEE

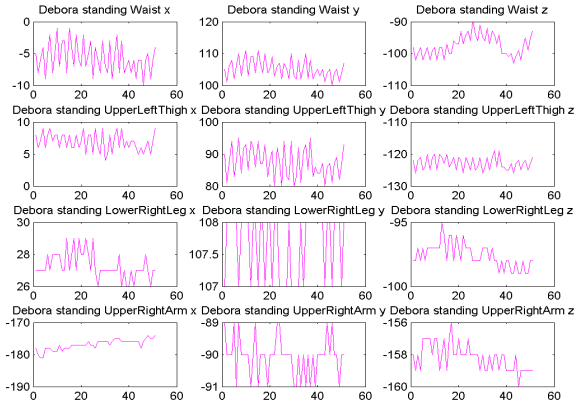

Fig. 3. Data visualization for Debora's standing

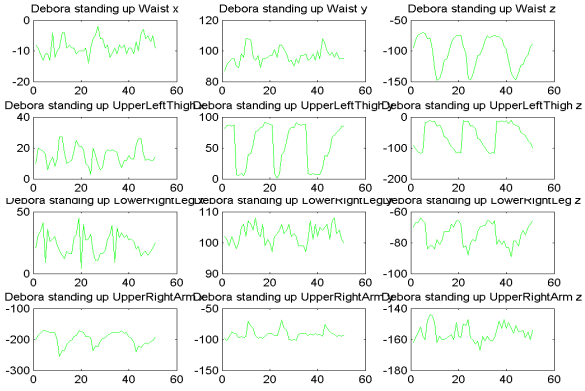

Fig. 4. Data visualization for Debora's standing up

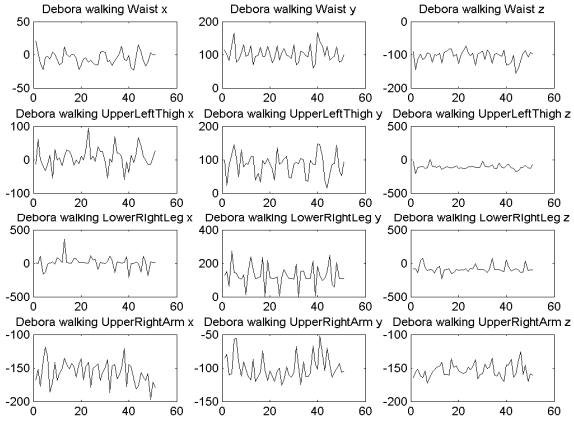

Fig. 5. Data visualization for Debora's walking

III. Optimization Algorithms

A. SVM based Algorithm

SVM is originally a two-class linear classification algorithm, which can also be extended to be a multi-class classifier or a nonlinear classifier by mapping the data from the input space to a new desired space. SVM seeks solutions for the following optimization problem:

$$\min_{\omega,b,\xi} \frac{1}{2}\omega^{\mathrm{T}}\omega + \mathrm{C}\sum_{i=1}^{l}\xi_i \tag{1}$$

subject to

$$\mathrm{y}_i(\omega^T\phi(\mathrm{x}_i)+\mathrm{b}) \geq 1-\xi_i \tag{2}$$

$$\xi_i \geq 0 \tag{3}$$

where C is essentially a regularization parameter that controls the trade-off between achieving a low error on the training data and minimizing the norm of the weights. Tuning C correctly is a vital step in best practice in the use of the SVM, Specifically, for large values of C, the optimization will choose a smaller-margin hyperplane if that hyperplane does a better job of getting all the training points classified correctly. Conversely, a very small value of C will cause the optimizer to look for a larger-margin separating hyperplane, even if that hyperplane misclassifies more points [14].

The Radius Basis Function (RBF) kernel, instead of polynomial kernel function, was used in this paper.

$$K(\mathrm{x}_i, \mathrm{x}_j) = \exp(-\frac{\| x_i - x_j \|^2}{2\sigma^2}) \tag{4}$$

where $\frac{1}{2\sigma^2} = \gamma$. The parameter γ sets the width of bell-shaped curve of the kernel. The larger the value of γ , the narrower will be the bell. Smaller values of γ result in wide bells.

To select proper values for parameters C and γ , the five-fold cross-validation and the grid search were used in the paper. The grid was constructed by $C = [2^{-3}, 2^{-2}, \ldots, 2^{3}]$ and $\gamma = [2^{-6}, 2^{-5}, \ldots, 2^{3}]$.

To classify multiple classes, the "One-against-All" approach was used here. If the number of classes is k , this approach first constructed k classifiers. To train the k classifiers, the *svmtrain* function in the LIBSVM toolbox [15] was used for the training procedure. After training, *svmtrain* returned a trained model for the k *th* class against all other classes, and the test data can be classified with the *svmpredict* function in the LIBSVM toolbox.

B. HMM based Algorithm

The reason that recognizes human body activities using HMM is that the sequence of human body activities can be extracted as first-order Markov chains. Given the labeled observation sequences, which can be seen as feature vectors extracted from wearable sensors' measurements, each human activity class can be modeled using a single HMM. And then each test observation sequence can be put into each HMM model and hence the likelihood of the observation can also be calculated. The class with the highest likelihood for the HMM can be finally determined to be the class that the test observation belongs to.

978-1-5090-5443-5/17/$31.00 ©2017 IEEE

To represent the HMM, λ = (A; B; π) is used. S is the state alphabet set, and V is the observation alphabet set:

$$\mathrm{S}=(\mathrm{s}_1,\mathrm{s}_2,...,\mathrm{s}_N) \tag{5}$$

$$\mathrm{V}=(\mathrm{v}_1,\mathrm{v}_2,...,\mathrm{v}_\mathrm{M}) \tag{6}$$

Define Q as a fixed state sequence with length T, and O as the corresponding observations:

$$\mathrm{Q}=q_1,q_2,...,q_T \tag{7}$$

$$O=o_1,o_2,...,o_T \tag{8}$$

Present the probability of state j following state i with transition matrix $\mathbf{A}$.

$$\mathbf{A}=\left\{a_{ij}=P\left[q_t=s_j|q_t=s_i\right]\right\},\ 1\le i,j\le M \tag{9}$$

where $\sum_{j=1}^{N} a_{ij}=1$.

Present the probability of observation k produced from the state j using observation matrix $\mathbf{B}$.

$$\mathbf{B}=\left\{b_{jk}=P\left[o_t=v_k|q_t=s_j\right]\right\} \tag{10}$$

where $\sum_{k=1}^{M} b_{jk}=1$.

Define π as the initial probability array:

$$\pi=\left[\pi_\mathrm{i}\right],\ \pi=P(\mathrm{q}_1=\mathrm{s}_i) \tag{11}$$

To train the HMM parameters, the Baum-Welch algorithm [16] was used in this paper. Specifically, Sitting and Standing were modeled as three states HMM. Sitting-down, Standing-up and Walking were modeled as six states HMM. All are left-to-right. The Gaussian mixture with three Gaussian distributions was used to describe the probability density of the observations for each state.

With observations $O=o_1,o_2,...,o_T$ and states $Q=q_1,q_2,...,q_M$, for each class λ_i, a corresponding HMM was trained using Baum-Welch algorithm as follows:

$$\max_{\lambda}\left\{P(O|\lambda)\right\} \tag{12}$$

$$\begin{aligned} P(\mathbf{O}\mid\lambda)&=\sum_{Q} P(\mathrm{O},\mathrm{Q}\mid\lambda)=\sum_{Q} P(\mathrm{O}\mid\mathrm{Q},\lambda)\,\mathrm{P}(\mathrm{Q}\mid\lambda)\\ &=\sum_{Q} \pi_{q_1} b_{q_1 O_1} a_{q_1 q_2} b_{q_2 O_2}\ldots a_{q_{T-1} q_T} b_{q_T O_T} \end{aligned} \tag{13}$$

where

subject to

$$\sum_{j=1}^{N} a_{ij}=1 \tag{14}$$

$$\sum_{k=1}^{M} b_{jk}=1 \tag{15}$$

$$\sum_{i=1}^{N} \pi_i=1 \tag{16}$$

The HMM classification procedure for a new test sample was performed using the forward algorithm. When a new test sequence, Y, was given, the probabilities for the five human activity classes were calculated respectively as follows:

$$P_j=\log(\mathrm{P}(\mathrm{Y}\mid\lambda_j)),\quad \mathrm{j}=1,\ldots,5 \tag{17}$$

The activity class whose HMM had the highest log likelihood was finally determined to be the class that the Y associated with, which was shown as:

$$\lambda^{*}=\max_{j}\{\mathrm{P}(\mathrm{Y}\mid\lambda_j)\} \tag{18}$$

C. ANN based Algorithm

ANN can process one human activity training sample at one time or a batch of training samples at the same time, and continuously improve itself by comparing their classification results with their actual classes [17]. Errors from the previous training result are fed back into the network, and reused to modify the parameters of the network again. Repeat this procedure for several times until the output errors are reduced to some certain values. After the neural network is well trained, when a new test sample is input into this neural network, the output nodes will output some values, which are used to assign the test sample to a class whose output node has the largest value.

In sum, the optimization goal of ANN was simplified as follows:

$$\min_{w}\{\mathrm{J}(\omega)\} \tag{19}$$

$$J(\omega)=\frac{1}{2}\|\mathrm{t}-\mathrm{z}\|^2 \tag{20}$$

where t is the label of the training/testing data, z is the result of a feedforward process in the network, and ω stands for the weight between connected layers.

Some basic rules for the design of the ANN are as follows: (1) With the increase of the complex relationship between the input sample data and the desired output data, the number of elements in the hidden layer(s) usually should also be increased. (2) If the classification process model is separated into several stages, some additional hidden layer(s) might be needed in the network. (3) The redundant hidden layers might make the trained network overfit. (4) The number of training data can be used to set an upper bound for the amount of elements in the hidden layer(s).

D. SRC-RP Algorithm

Suppose there are k different human activity classes and L wearable sensors put on each human subject. And each activity signal is *m*-dimensional. Construct a training sample set A with

978-1-5090-5443-5/17/$31.00 ©2017 IEEE

k subsets, which are denoted as A_i, i=1, 2, ..., k, meaning k different activities are considered in the training sample set. So A = [A_1, A_2, …, A_i, …, A_{k-1}, A_k]. Suppose there are n_i samples in subset A_i and $n_1+n_2+\cdots+n_k = n$. So A_i can be represented as

$$A_i=[a_{i1}, a_{i2}, \ldots, a_{ij}, \ldots, a_{in_i}]\in R^{m\times n_i} \quad (21)$$

where $a_{ij}\in R^{m\times 1}$ denotes the j*th* training activity signal in the training sample subset A_i.

When the amount of training samples in A_i is sufficient, for any new test sample signal $y\in\mathbb{R}^m$ belonging to the i*th* class, it can be approximately represented using a linear superposition of the training samples in subset A_i [5]

$$y = w_{i1}\bullet a_{i1}+\cdots+w_{ij}\bullet a_{ij}+\cdots+w_{in_i}\bullet a_{in_i} \quad (22)$$

where $w_{ij}\in\mathbb{R}$, j=1, 2, …, n_i.

In general, each test sample y has to be represented using all the training sample signals in the training sample set, since the subset that the new test sample y belongs to could not be predicted in advance,

$$\begin{aligned} y &= \mathbf{A}x_0 = [A_1, A_2, \cdots, A_k]x_0 \\ &= [a_{11}, \cdots, a_{1n_1}, a_{21}, \cdots, a_{2n_2}, \cdots, a_{k1}, \cdots, a_{kn_k}]x_0 \end{aligned} \quad (23)$$

where $A\in\mathbb{R}^{m\times n}$, $n=n_1+n_2+\ldots+n_k$; x_0 = [0, … , 0, w_{i1}, w_{i2}, ... , w_{in_i}, 0, …, 0]$^T\in\mathbb{R}^n$. x_0 is a sparse coefficient vector whose entries, ideally, are zeros except for those entries correspondingly associated with the class that y belongs to.

To save computing resources, random projection is used on each wearable sensor node to reduce the dimensionality of the activity signal before the signal is finally sent to the central node. Intuitively, dimensionality reduction can be interpreted as to find an $m-$ dimensional signal to represent an $n-$ dimensional signal $\mathbf{x}$, with $m\ll n$, while maintaining enough useful information [18]. Suppose a human activity signal collected by a wearable sensor is denoted as $x\in\mathbb{R}^n$. A straightforward way to reduce the dimensionality of signal $\mathbf{x}$ is to apply a mapping matrix $W\in\mathbb{R}^{m\times n}$ such that $\hat{\mathbf{x}}=W\mathbf{x}\in\mathbb{R}^m$, which takes $o(nm)$ computations [19]. If a sparse mapping matrix W with only r nonzero entries in each column is applied, the computational complexity for the dimensionality reduction can be reduced to $o(rm)$ [12]. Therefore, a careful design of the mapping matrix W is very helpful to improve the computational efficiency.

Random projection makes use of randomized constructions for the mapping matrix W, for instance, sparse random binary matrix, random Gaussian matrix, and random Bernoulli matrix, to ensure the mapping matrix can meet the Restricted Isometry Property (RIP) condition with a high probability. In addition, random projection is very convenient to use since it is neither dependent of the training sample data nor in need of extra storage space to record the intermediate variables [5].

In this paper, specifically, random projection W_j is applied on each wearable sensor node to reduce the dimensionality of the human activity signal y_j' measured by the wearable sensor j (j= 1, 2, …, L) during a certain time period.

$$\hat{y}_j' = W_j y_j' \in R^{d_j} \quad (24)$$

An integral human activity signal $\hat{y}$ can be constructed using signals collected by all the L wearable sensors.

$$\hat{y}=[y_1'^T, y_2'^T, \cdots, y_L'^T]^T\in R^d \quad (25)$$

which can also be denoted as

$$\hat{y}=\begin{bmatrix} W_1 & 0 & \cdots & 0 \\ 0 & W_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & W_L \end{bmatrix}\begin{bmatrix} y_1' \\ y_2' \\ \vdots \\ y_L' \end{bmatrix}=Wy' \quad (26)$$

The same random projection matrix W is applied on both the training sample set and test sample set. Thus the SRC-RP based human body activity recognition model can be denoted as

$$Wy'=WAx+\varepsilon \quad (27)$$

where ε is the noise factor to account for the measurement noise and $\|\varepsilon\|_2\le\beta$. β is the noise tolerance and $\beta>0$.

Considering the amount of the samples is larger than the dimensionality of the sample signals, i.e., $m<n$, equation (27) is an underdetermined equation., which can be solved by L0 minimization. If x is a sparse vector, L1 minimization algorithm can be used to solve for the sparse representation coefficients to avoid the NP problem involved in the L0 minimization.

$$\hat{x}=\arg\min\|x\|_1 \quad s.t. \quad \|Wy'-WAx\|_2\le\beta \quad (28)$$

To decide which human activity class that the test signal y belongs to, first, equation (28) can be used to solve for the sparse representation coefficient vector. Then a classification strategy of comparing how well the various components of the coefficient vector $\hat{x}$ associated with different activity classes can approximately reproduce the test signal y. Intuitively, the nonzero coefficients in $\hat{x}$ ideally should be mostly associated with those corresponding training samples which belong to the same activity class as y. In fact, due to the measurement noise and incomplete training sample set, a few small nonzero coefficients of $\hat{x}$ might also be associated with some other activity classes which y does not belong to. Define a characteristic function $\theta_j(\hat{x})\in\mathbb{R}^n$, which only reserves the

978-1-5090-5443-5/17/$31.00 ©2017 IEEE

coefficients in $\hat{x}$ associated with class j and set all the other coefficients to be zeros. So the reproduced approximate value of y corresponding to the j*th* class is denoted as:

$$\hat{y}_j = A\theta_j(\hat{x}) \tag{29}$$

At last, the test human activity signal y is set to be the activity class resulting in the minimum residual value.

$$\min_j r_j(y) = \left\| y - A\theta_j(\hat{x}_1) \right\|_2 \tag{30}$$

In addition, to measure the classification confidence, the classification result is also evaluated using matric sparsity concentration index (*SCI*), ranging from 0 to 1, proposed in [7]. The definition of SCI for a signal x is described as

$$SCI(x) = \frac{k \cdot \max_j \| \theta_j(x) \|_1 / \| x \|_1 - 1}{k - 1} \tag{31}$$

The complete SRC-RP algorithm is stated in Algorithm 1.

Algorithm 1: SRC-RP

Input Parameters: Initial training sample signal set $A = [A_1, A_2, \ldots, A_i, \ldots, A_{k-1}, A_k] \in \mathbb{R}^{m \times n}$; Random projection mapping matrix W; Test sample signal $y \in \mathbb{R}^m$; Noise tolerance $\beta > 0$; SCI threshold value $\tau \in [0,1]$.

i. Construct the test sample set Wy and the training sample set WA.

ii. Normalize all the columns of Wy and WA.

iii. Solve for $\hat{x}$ by L1 minimization algorithm:

$$\hat{x} = \arg\min \|x\|_1 \quad s.t. \quad \|Wy - WAx\|_2 \le \beta .$$

iv. Calculate the residuals:

$$r_j(y) = \left\| Wy - WA\theta_j(\hat{x}) \right\|_2 , \text{ j=1, 2, …, k.}$$

v. Determine $c(y) = \arg\min_j r_j(y)$.

vi. Compute $SCI(\hat{x})$

If $SCI(\hat{x}) \ge \tau$ Return: Class (y)= $\arg\min_j r_j(y)$.

IV. Numerical Experiments and Results

A. SVM based Classification

There were two kinds experiments conducted in this section. First, the One-against-Own experiment used the first 400 signal samples for each human subject. Specifically, the first 4/5 portion of the samples was reserved for training while the rest 1/5 part for test. Second, the One-against-All experiment used 600 training signal samples in total, which were equally contributed by Debora, Katia and Wallace, and used another 100 samples from Jose for test.

1) One-against-Own: This kind of experiment was conducted on the dataset of Debora, Katia and Wallace. The reason why this experiment was not conducted on the dataset of Jose is because the amount of data samples collected from Jose is not sufficient enough. The parameters of the SVMs in this experiment are shown as follows:

SVM for Debora: C = 0.25 and γ = 0.0156;

SVM for Katia: C = 0.125 and γ = 0.0156;

SVM for Wallace: C = 1 and γ = 0.0156;

The results of the recognition accuracy are illustrated in Fig.6-8 for Debora, Katia and Wallace respectively. The final recognition accuracy is 92.5% for experiment conducted on Debora, 99.5% for Katia, and 93.5% for Wallace.

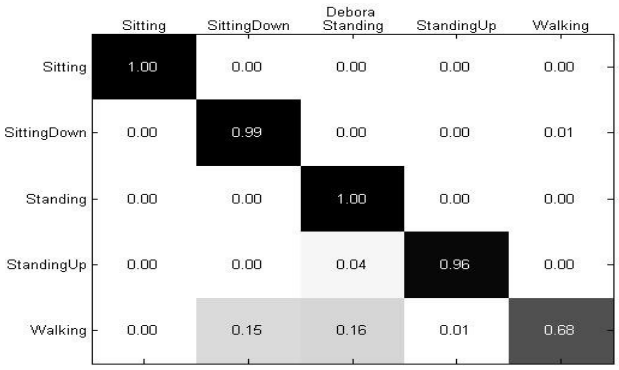

Fig. 6. Recognition result using SVM for Debora vs herself

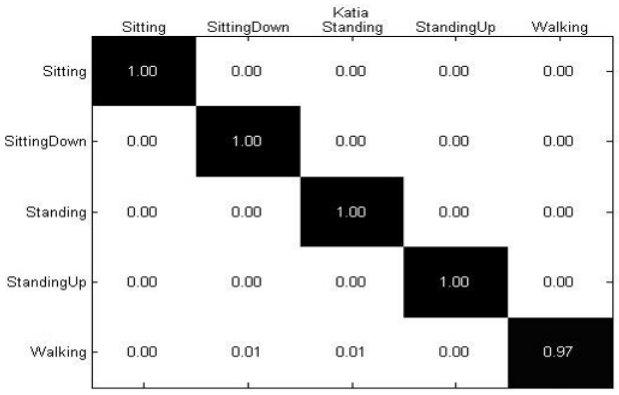

Fig. 7. Recognition result using SVM for Katia vs herself

978-1-5090-5443-5/17/$31.00 ©2017 IEEE

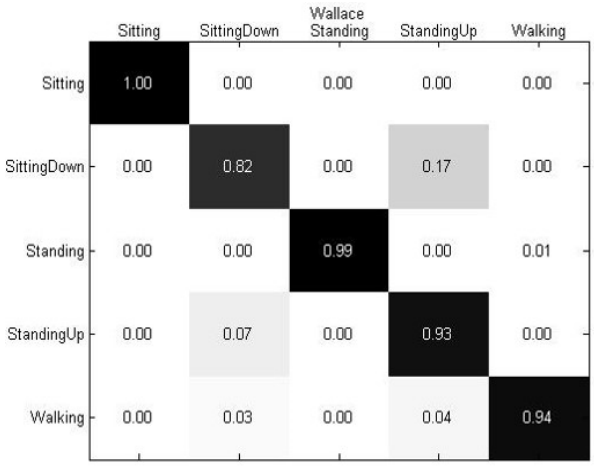

Fig. 8. Recognition result using SVM for Wallace vs himself

2) One-against-All: The kind of experiment was first conducted on Debora, Katia and Wallace against Jose, with the SVM parameters were set to be C = 1 and $\gamma = 0.0156$. The final recognition result is shown in Fig. 9, which indicates a low recognition accuracy, only 41.2%. And then, this kind of experiment was conducted again on Katia, Wallace and Jose against Debora, with the SM parameters setting at $C = 0.25$ and $\gamma = 0.0156$. The final result is illustrated in Fig.10, which indicates that the recognition accuracy is improved to be 60%.

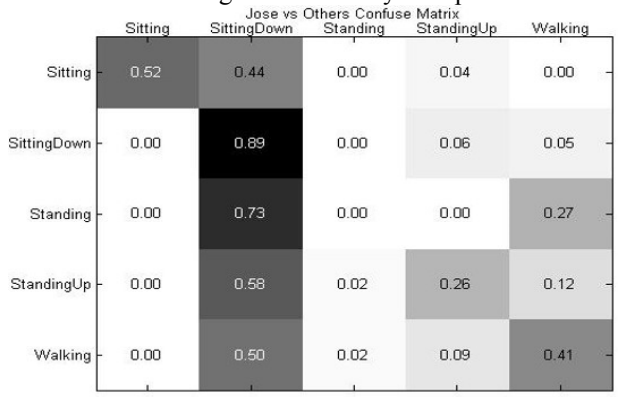

Fig. 9. Recognition result using SVM for Jose vs others

There are several possible reasons for this result of low recognition accuracy. First, different subjects could move their bodies in different patterns, especially considering that Jose is elder than the other three human subjects. Second, the number of subjects is too small and the training sample set is not large enough to train this SVM model. The reason why the recognition accuracy is improved in the second experiment is probably related with the physical features of the subject Debora, who has more similarities in age, health status with some subjects in the training sample set. Anyway, it is better to take this kind of experiment on a larger dataset with more subjects to improve the recognition accuracy.

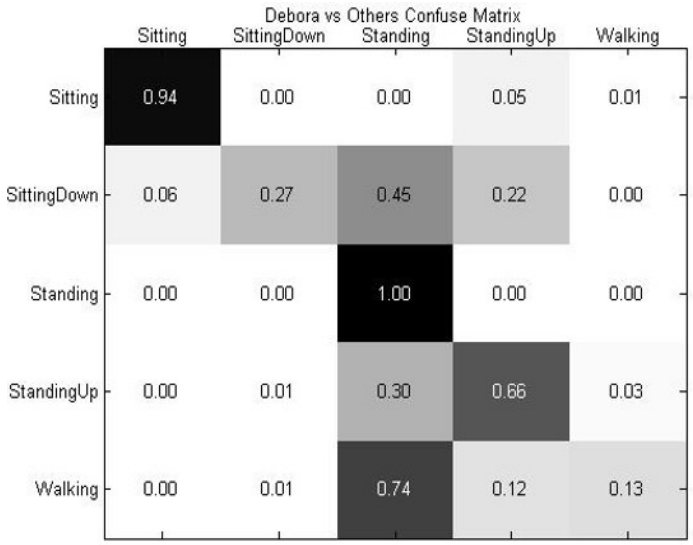

Fig. 10. Recognition result using SVM for Debora vs others

B. *HMM based Classification*

In the HMM experiment, the data samples were first grouped into sequences of 10-sample date in order to properly apply the HMM classifier. The One-against-Own experiment was conducted on the dataset of Debora, Katia and Wallace. The number of Gaussian mix in the HMM was set to be three. The numbers of states for classes of sitting, walking, standing, standing-up and sitting-down were set at three, six, three, six and six respectively, which were based on the fact that sitting and standing are relatively simpler activities than the other three. The number of Gaussian mix and the number of states were selected to keep the balance between the classification accuracy and the computational complexity. The MATLAB HMM toolbox based on [16] was used in this paper to train and test the different HMMs.

The result is illustrated in Fig. 11-13 for the One-against-Own experiment using the confuse matrix. In sum, the recognition accuracy is 90.3% for Debora, 99.7% for Katia, and 97.3% for Wallace, all achieving high recognition accuracy.

In the One-against-All experiment, Debora's, Katia's and Wallace's data were used for training and Jose's data for test. The result is illustrated in Fig. 14. It can be seen that the recognition accuracy is only 36.8%.

The reason why the recognition accuracy is high for the One-against-Own experiment while it is low for the One-against-All is that since each subject keeps certain and similar patterns in his/her own activities, one subject can easily and accurately learn these patterns from his/her training data making the recognition rate very high. But different subjects have different patterns in their activities considering their age, gender, height, and so on. In order to achieve a high recognition accuracy in the One-against-All experiment, more training data with more subjects are needed.

978-1-5090-5443-5/17/$31.00 ©2017 IEEE

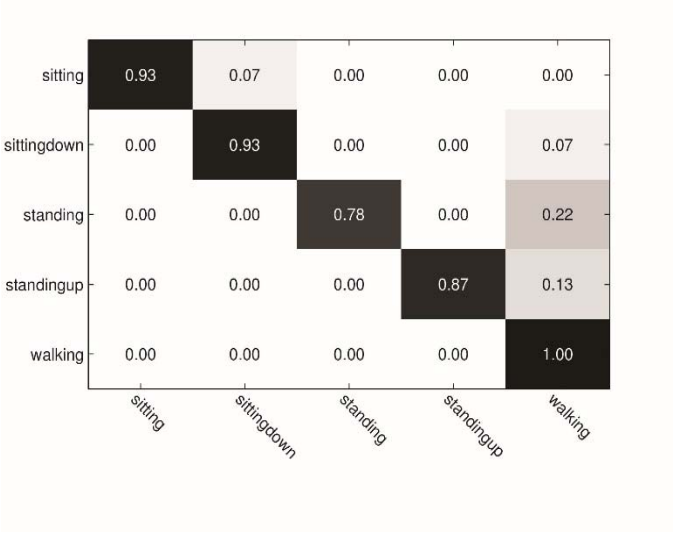

Fig. 11. Recognition result using HMM for Debora vs herself

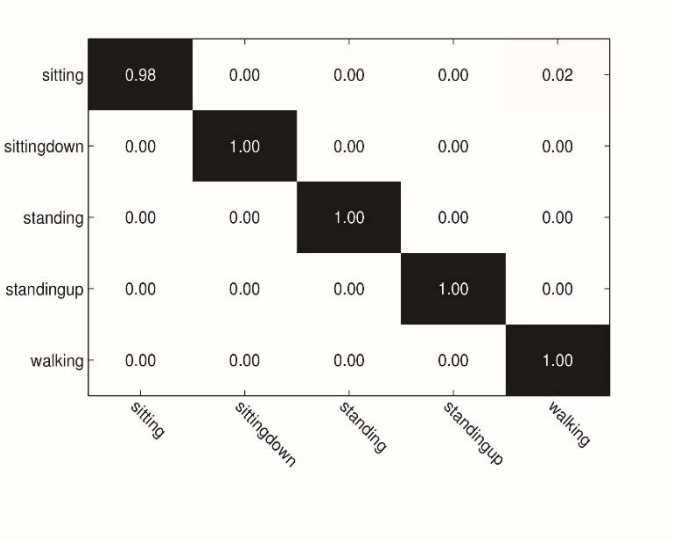

Fig. 12. Recognition result using HMM for Katia vs herself

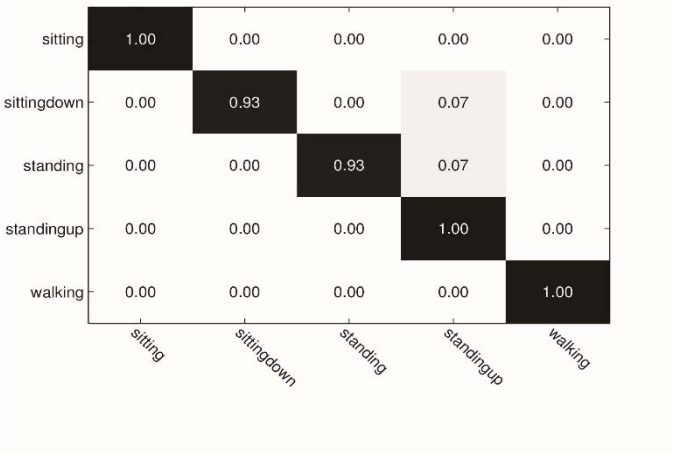

Fig. 13. Recognition result using HMM for Wallace vs himself

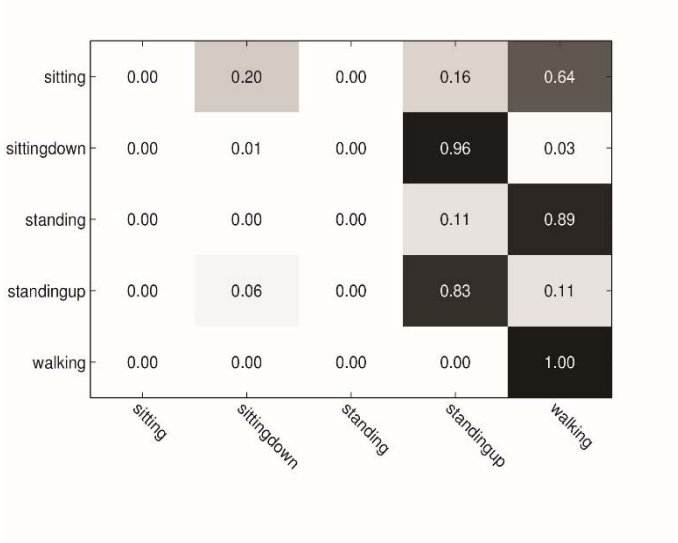

Fig. 14. Recognition result using HMM for One-against-all experiment

C. ANN based Classification

In the One-vs-Own experiment, four hundred data samples are taken from each subject, and the first 4/5 portion of the data samples was kept for training while the rest 1/5 portion was reserved for test. This kind of experiment was performed on Debora, Katia and Wallace, excluding Jose, since not enough data samples were collected from Jose. The recognition accuracy achieves 93.4% for Debora, 99.6% for Katia and 95.6% for Wallace. The network model used is a 12-10-10-5 multi-layer neural network.

In the One-vs-All experiment, six hundred data samples were taken from one subject for training, and another one hundred data samples were taken from another subject for test. A 12-10-10-5 multilayer neural network was designed for this experiment. For example, when we used Debora's data for training and Katia's data for test, the recognition accuracy is about 61.9%.

The deep learning MATLAB toolbox developed by [20] was used for both experiments in this paper.

The advantages of ANN are reflected in the following aspects. First, it is a data driven and self-adaptive method, which means it can learn the patterns from the data without explicitly formulating any function in advance. Furthermore, it is a relatively easy and flexible way to form nonlinear models as well as approximately representing any function with a pre-set accuracy. In addition, it can also be used to estimate the posterior probabilities, offering the basis for establishing recognition rules and performing further statistical analysis.

The human activity recognition accuracy of the three machine learning methods is summarized in Table II.

From the One-against-Own experiments, the activity recognition accuracies for all of the three machine learning classifiers, SVM, HMM and ANN, are above 90%. It can be concluded that a high activity recognition accuracy is generally easy to be achieved using the data collected from the wearable

978-1-5090-5443-5/17/$31.00 ©2017 IEEE

sensors which are installed on the subject. This is because the activity pattern of one subject maintains consistently which can be well learned even using a small training dataset. However, the activity recognition accuracy for the One-against-All experiments is not good enough. The reason behind this is that there are many discrepancies among different subjects, such as age, gender, weight and so on. In order to improve the recognition accuracy of the body activity classifier, a larger training dataset is required and some better features and more appropriate parameters need to be designed.

TABLE II. RECOGNITION ACCURACY SUMMARY

Experiment	**SVM**	**HMM**	**NN**
Debora vs Own	92.5%	90.3%	93.4%
Katia vs Own	99.5%	99.7%	99.6%
Wallace vs Own	93.5%	97.3%	95.6%
One vs All	41.2%	36.8%	61.9%

D. SRC-RP based Classification

In this experiment, 40 original sample signals were concatenated together to create a new sample signal, changing the dimensionality of the sample signal to be 480. The confusion table summarizes the activity recognition results for the five human subjects with a compression ratio of 50% for SRC-RP, which is illustrated in Table III. Compression ratio is the ratio of dimensionalities of the activity sample signals after and before applying the random projection. For convenience, in Table III, 1 stands for Sitting; 2 for Sitting Down; 3 for Standing; 4 for Standing up; 5 for Walking.

TABLE III. CONFUSION TABLE FOR RECOGNIZING FIVE HUMAN ACTIVITIES USING SRC-RP

	1	2	3	4	5	RATE
1	31002	190	237	361	279	96.67%
2	14	10586	135	230	119	95.60%
3	281	279	41656	421	493	95.61%
4	246	132	153	9002	198	92.51%
5	221	247	443	472	30259	95.63%

It can been seen from Table III that the activity Sitting has the best recognition performance while the activity Standing Up has the worst recognition performance.

To compare the performance of SRC-RP with the performance of the other three machine learning methods, the human activity recognition rates under different compression ratios are illustrated in Fig. 15. And the compression ratio is set at 10%, 30%, 50%, 70% and 90%, respectively. Meanwhile, different random projection matrixes including sparse random binary matrix, random Gaussian matrix and random Bernoulli matrix are also compared for the SRC-RP algorithm.

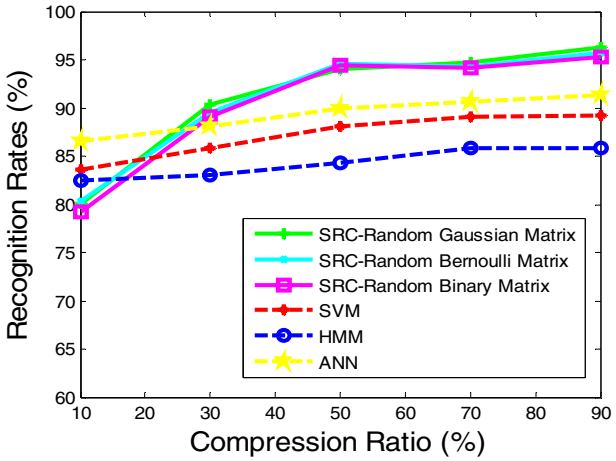

Fig. 15. Human activity recognition rates for different classification methods with the consideration of different compression ratios

It can be seen from Fig. 15 that when the compression ratio goes to 50%, the human activity recognition performance based on SRC-RP is apparently better than the performance based on SVM, HMM and ANN. When the compression ratio is less than 50%, increasing the compression ratio can apparently increase the human activity recognition accuracy for SRC-RP. When the compression ratio is higher than 50%, the human activity recognition accuracy for SRC-RP appears to be stable. Compared to the SRC-RP based method, the human activity recognition rates of SVM, HMM and ANN are less likely to be affected by the different compression ratios.

In addition, Fig. 15 also indicates that the three different random projection matrices almost have rare different effects on the human activity recognition performance. One of the biggest advantages of using sparse random binary matrix is that there are less nonzero entries in the sparse random projection matrix, saving more computing and memory resources than the random Gaussian matrix or the random Bernoulli matrix.

V. OVERALL CONCLUSION AND DIRECTION FOR FUTURE WORK

This paper presents three different machine learning algorithms, SVM, HMM and ANN, and one compressed sensing based algorithm, SRC-RP, to recognize human body activities. And numerical experiments are also conducted on a real-world human activity dataset collected using wearable sensors to validate the effectiveness of these algorithms.

To improve the One-against-All recognition performance, other features, such as age, gender, height and weight of the subjects can be also used in the experiment.

For the SVM based recognition algorithm, future work would be to propose a more efficient method to determine the optimal model parameters rather than only relying on cross-validation. For example, a coarse cross-validation grid can be first generated to determine the range of the optimal parameters and then a fine grid can further generated within that range,

978-1-5090-5443-5/17/$31.00 ©2017 IEEE

which would decrease the overall computational complexity. In addition, other kernels could also be used to see whether they can achieve better recognition performance than RBF.

For the HMM based recognition algorithm, future work would be to find the optimal number of statues and the mixture of Gaussian using more efficient approaches instead of relying on intuition or trial and error.

Future work on SRC-RP would be to select the training sample set in a more efficient way to reduce the overall runtime and improve the recognition accuracy.

Some other recognition methods can also be tried to solve the human activities recognition problem. For example, the clustering analysis and regression models proposed in [21] and [22] could probably also provide an easy way to recognize various human activities.

REFERENCES

[1] Y.J. Hong, K. Ig-Jae, C. A. Sang, and K. Hyoung-Gon, "Activity recognition using wearable sensors for elder care," *Second International Conference on Future Generation Communication and Networking*, vol. 2, pp. 302-305, 2008.

[2] L. Cheng, Y. Guan, K. Zhu, and Y. Li, "Recognition of body activities using machine learning methods with wearable sensors," *The 7th IEEE Annual Computing and Communication Workshop and Conference*, pp. 1-7, 2017.

[3] J. Y. Yang, J. S. Wang, and Y. P. Chen, "Using acceleration measurements for activity recognition: An effective learning algorithm for constructing neural classifiers," *Pattern Recognition Letters*, vol. 29, pp. 2213-2220, 2008.

[4] L. Gao, A. K. Bourke, and J. Nelson, "A system for activity recognition using multi-sensor fusion," *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 7869-7872, 2011.

[5] L. Cheng, L. Wang, and F. Gao, "Power system fault classification method based on random dimensionality reduction projection and sparse representation," *IEEE Power & Energy Society General Meeting*, pp. 1-5, 2015.

[6] L. Cheng, C. You, and L. Chen, "Identification of power line outages based on PMU measurements and sparse overcomplete representation," *IEEE 17th International Conference on Information Reuse and Integration,* pp. 343-349, 2016.

[7] J Wright, AY Yang, A Ganesh, et al, "Robust face recognition via sparse representation," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 31, no. 2, pp. 210-227, 2009.

[8] M. Zhang, and A. A. Sawchuk, "Human daily activity recognition with sparse representation using wearable sensors," *IEEE Journal of Biomedical and Health Informatics*, vol. 17, no. 3, pp. 553-560, 2013.

[9] L. Xiao, R. Li, and J. Luo, "Recognition of human activity based on compressed sensing in body sensor networks," *Journal of Electronics & Information Technology*, vol. 35, no. 1, pp. 119-125, 2013.

[10] T. Guha, and R. K. Ward, "Learning sparse representations for human action recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 8, pp. 1576-1588, 2012.

[11] C. Liu, Y. Yang, and Y. Chen, "Human action recognition using sparse representation," *IEEE International Conference on Intelligent Computing and Intelligent Systems*, vol. 4, pp. 184-188, 2009.

[12] L. Cheng, Y. Li, and Y. Guan, "Human activity recognition based on compressed sensing," *The 7th IEEE Annual Computing and Communication Workshop and Conference*, pp. 1-7, 2017.

[13] W. Ugulino, D. Cardador, K. Vega, E. Velloso, R. Milidiú, and H. Fuks, "Wearable computing: Accelerometers' data classification of body postures and movements," *Advances in Artificial Intelligence-SBIA 2012*, pp. 52-61. Springer Berlin Heidelberg, 2012.

[14] C.W. Hsu, C.C. Chang and C.J. Lin, "A Practical Guide to Support Vector Classification", http://www.csie.ntu.edu.tw/ cjlin, 2010.

[15] C.C. Chang and C.J. Lin, "LIBSVM: a library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, pp. 27:1-27:27, 2011.

[16] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proceedings of the IEEE 77*, no. 2, pp. 257-286, 1989.

[17] S. A. Kustrin and R. Beresford, "Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research," *Journal of Pharmaceutical and Biomedical Analysis 22*, no. 5, pp. 717-727, 2000.

[18] L. Cheng, and C. You, "Hybrid non-linear dimensionality reduction method framework based on random projections," *IEEE International Conference on Cloud Computing and Big Data Analysis*, pp. 43-48, 2016.

[19] L. Cheng, C. You, and Y. Guan, "Random projections for non-linear dimensionality reduction," *International Journal of Machine Learning and Computing*, vol. 6, no. 4, pp. 220-225, 2016.

[20] R. B. Palm, "Prediction as a candidate for learning deep hierarchical models of data," *Technical University of Denmark 5*, 2012.

[21] L. Cheng and C. You, "Analysis of rising tuition rates in the United States based on clustering analysis and regression models," *Computer Science Conference Proceedings in Computer Science & Information Technology*, vol. 6, no. 6, pp. 127-144, 2016.

[22] L. Cheng and C. You, "Analysis of tuition growth rates based on clustering and regression models," *International Journal of Data Mining & Knowledge Management Process*, vol. 6, no. 4, pp. 1-17, 2016.

978-1-5090-5443-5/17/$31.00 ©2017 IEEE