

Research Article

Daily Human Physical Activity Recognition Based on Kernel Discriminant Analysis and Extreme Learning Machine

Wendong Xiao and Yingjie Lu

School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, China

Correspondence should be addressed to Wendong Xiao; wdxiao@ustb.edu.cn

Received 10 September 2014; Accepted 26 November 2014

Academic Editor: Amaury Lendasse

Copyright © 2015 W. Xiao and Y. Lu. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Wearable sensor based human physical activity recognition has extensive applications in many fields such as physical training and health care. This paper will be focused on the development of highly efficient approach for daily human activity recognition by a triaxial accelerometer. In the proposed approach, a number of features, including the tilt angle, the signal magnitude area (SMA), and the wavelet energy, are extracted from the raw measurement signal via the time domain, the frequency domain, and the time-frequency domain analysis. A nonlinear kernel discriminant analysis (KDA) scheme is introduced to enhance the discrimination between different activities. Extreme learning machine (ELM) is proposed as a novel activity recognition algorithm. Experimental results show that the proposed KDA based ELM classifier can achieve superior recognition performance with higher accuracy and faster learning speed than the back-propagation (BP) and the support vector machine (SVM) algorithms.

1. Introduction

Recognition of physical activity plays an important role in many fields such as physical training and health care. In particular nowadays, faced with the aging problem, a growing number of old people live alone and urgently demand advanced solutions for their health monitoring, including their physical activity recognition.

A number of key research issues are related to physical activity recognition, including how to improve the data collection mechanisms, how to select more effective features, and how to design high-performance classification algorithms. Different solutions have been proposed to address these issues, usually based on the video analysis and the wearable sensor signal analysis. The video analysis approach needs to obtain the position and attitude information from a series of body image sequences. Due to the complexity and the variability of the human physical activities, it often suffers from low accuracy and low efficiency in many practical scenarios. It is also weak for privacy of the monitored person [1, 2]. With the development of microelectromechanical system (MEMS) and wearable sensor networks, activity recognition approaches based on the wearable sensors, especially

accelerometers, have received increasing interests, due to their advantages that can be implemented easily anywhere anytime [3–8].

In previous studies, many researchers use multiple sensors for physical activity recognition to improve the accuracy. But the hardware would obstruct the movements of the human and is not practical for long-term wearing. Also the cost of system will increase with the number of sensors [3–5]. Therefore recently more researchers are seeking activity recognition approaches by using only one accelerometer sensor for collecting the signal [6–8].

Feature extraction is crucial for activity recognition. It is required to extract useful features from the raw measurement data to reduce the processing time and improve the recognition accuracy. Based on the accelerometer data, a number of features and feature extraction methods have been proposed, covering from the time domain analysis and the frequency domain analysis to the time-frequency domain analysis [4, 9]. The time domain method extracts features, such as the mean, the standard deviation, and the correlation coefficient, from the collected signal directly. The frequency domain method extracts features from the frequency-domain parameters, such as FFT coefficients. More recently, wavelet analysis,

which can incorporate time and frequency information, has been used to perform the time-frequency analysis.

Classifier is the key for activity recognition. Besides the high accuracy and short training time, the classifier is often expected to meet the real-time and the generalization requirements. A large number of classification methods have been investigated, including the artificial neural network (ANN) [7] and the support vector machine (SVM) [10], which have been widely used in machine learning and data analysis. But these popular learning techniques often face some challenging issues such as intensive human intervene, slow learning speed, and poor learning scalability [11–13]. Therefore, the recent extreme learning machine (ELM) [11] classifier will be proposed in this paper as a highly efficient and accurate activity recognition approach.

ELM is developed as a single-hidden layer feed-forward network (SLFN) that can randomly assign the weights between the input nodes and the hidden nodes and analytically determine the output weights between the hidden nodes and output nodes. It learns much faster than traditional gradient-based approaches, such as back-propagation (BP) algorithm. ELM tends to reach the small norm of the network output weights and achieve better generalization performance according to Bartlett's theory states [12]. In theory, ELM with the same kernels outperforms SVM in both regression and classification applications [13, 14]. ELM provides efficient unified solutions to generalized feed-forward networks including but not limited to neural networks, radial basis function (RBF) networks, and kernel learning [15]. Many improvements of ELM are under study in recent years, such as fast online learning [16], large scale ELM [17], and voting based ELM [18]. Also various applications can be found based on ELM and its variants, including wireless indoor localization [19], multicollinear problem [20, 21], and protein sequence classification [22].

The patterns of the physical activities vary from simple activities (such as standing, sitting, and walking) to more complex action strings (such as eating, drinking, and cycling). With the addition of new activities, the features of the existing systems may be ineffective, and the recognition accuracy may decrease significantly. To find new features will increase the workload of the researchers, and classifier design will become difficult with the increase of the number of the features. Principal component analysis (PCA) [23] and linear discriminant analysis (LDA) [7, 24] are applied to select most discriminative features for activity recognition. Kernel discriminant analysis (KDA) [10, 25] is an extension of LDA to obtain nonlinear discriminating features by the kernel technique for mapping the data to the feature space. In this paper, we will introduce KDA to extract more meaningful features of the activities and integrate it with ELM classifier to achieve improved classification performance.

The paper is organized as follows. The proposed approach is detailed in Section 2, including its general description, system design, data acquisition and preprocessing, feature extraction method, KDA, and the ELM algorithm. Experimental results are reported in Section 3. Finally, conclusions and the future work are given in Section 4.

2. Proposed Approach

The proposed activity recognition approach uses a triaxial accelerometer for data collection. As shown in Figure 1, the overall approach consists of the following steps: data acquisition, preprocessing, feature extraction, KDA, ELM training and classification.

2.1. Wearable Component. The wearable sensor in this paper employs a MPU-6000 sensor, which contains a triaxial accelerometer and a triaxial gyroscope. The triaxial accelerometer is used to collect the raw acceleration measurement signal. The sampling frequency is set at 50 Hz with the output ranging in $[-4g, +4g]$. The sensor transmits the data to a mobile phone via Bluetooth module called HC-05 and the data are stored in the SD card of the phone.

The wearable sensor is small and convenient to carry and is worn on the subject's thigh of right leg, as shown in Figure 2. When the body wears the accelerometer at the state of standing, the x -axis represents the acceleration in the lateral direction, the y -axis represents the acceleration in the longitudinal direction, and the z -axis represents the acceleration in the vertical direction. In order to facilitate observations, the measurement of each axis is divided by the gravitational acceleration (taken as 9.8 m/s^2); therefore, the data are multiples of the gravitational acceleration.

2.2. Data Acquisition and Preprocessing. In the experiments, 10 subjects' data are collected, 5 females and 5 males, wearing the wearable sensor each day for a period to collect data for six different activities of daily life, including sitting, standing, walking, running, going upstairs, and going downstairs. Figure 3 shows the example acceleration signals of sitting and walking for each axis of the triaxial accelerometer.

As the sensor will be influenced by the gravity, the data are filtered by a high-pass filter with a cut-off frequency of 0.5 Hz to eliminate the influence of the gravity. As the frequency of human daily activities is not too high, a low-pass filter with a cut-off frequency of 20 Hz is used to filter high frequency noise. In addition, the data are smoothed by median filter to eliminate independent noise.

2.3. Feature Extraction. Features are extracted from the raw accelerometer data over the sliding window. In this paper, features were computed from 128 sampling points (2.56 s) with 64 samples overlapping between the consecutive sliding windows. Feature extraction on sliding windows with 50% overlap has been shown to be effective in previous studies on activity classification [4].

A number of features, including the mean, the standard deviation, the median, the correlation, the tilt angle (TA), the signal magnitude area (SMA), the frequency energy, the frequency entropy, and the wavelet energy, were extracted from the signals for activity recognition. A brief description of each feature is given as follows.

(1) *Tilt Angle (TA).* No matter what the state of the device is, there will be vertical gravity. When tilting the device, there will be a gravitational component in the z -axis. The tilt angle

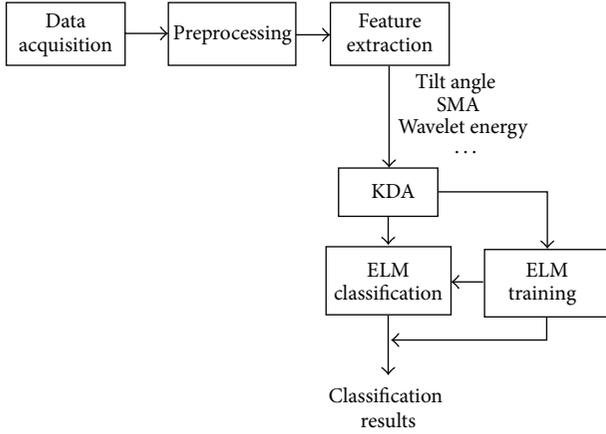


FIGURE 1: The architecture of the proposed activity recognition approach.

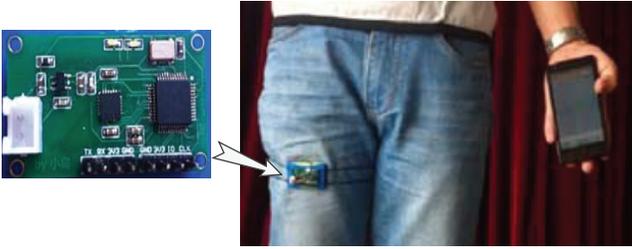


FIGURE 2: Wearable sensor being worn by a subject.

refers to the relative tilt of the body in space and is calculated from the acceleration in the z -axis according to

$$\theta = \arccos \frac{a_z}{g}, \quad (1)$$

where a_z is the gravity component in the z -axis and g is the gravity coefficient.

(2) *Signal Magnitude Area (SMA)*. We adopt the SMA to extract a feature quantity according to

$$SMA = \sum_{i=1}^N (|x(i)| + |y(i)| + |z(i)|), \quad (2)$$

where $x(i)$, $y(i)$, and $z(i)$ indicate the values of x -axis, y -axis, and z -axis acceleration signals at the i th sampling point after preprocessing and N is the length of the sliding window. SMA can indicate the fluctuation degree of the acceleration signal; the higher its value is, the more violent the fluctuation is.

(3) *Wavelet Energy (WE)*. This paper uses db5 as the mother wavelet. By decomposing the vertical component to 5 layers, WE is calculated as the sum of the squared detail coefficients at levels 4 and 5, according to

$$WE = \sum_{j=4}^5 CD_j^2, \quad (3)$$

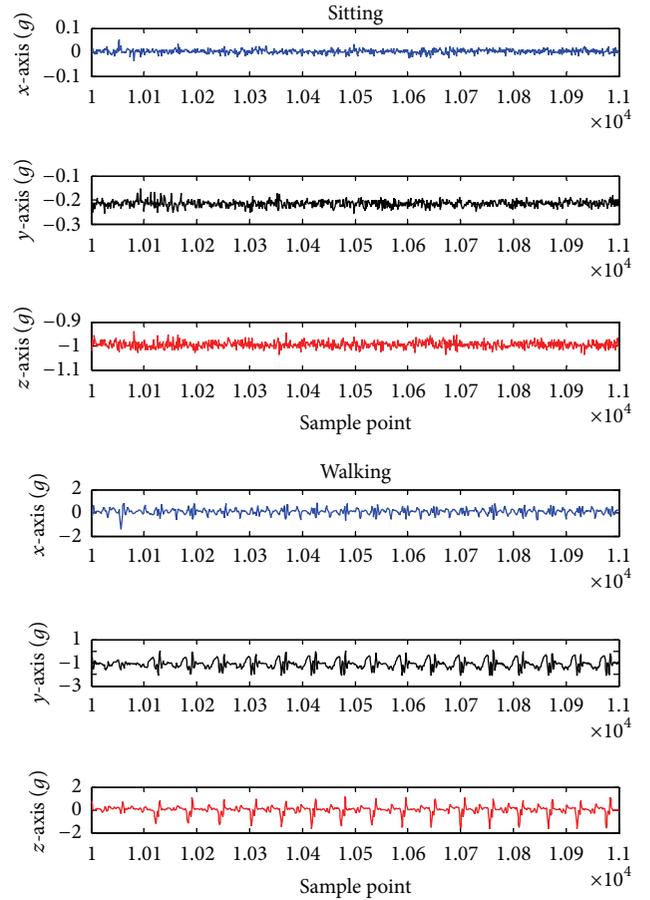


FIGURE 3: The raw acceleration signals of sitting and walking.

where CD_j are the detail coefficients of y -axis acceleration signal.

(4) *Other Features*. Mean, standard deviation, and median are the average, standard deviation, and median value of the signal over the sliding window, respectively. Correlation coefficient is calculated between the y -axis and z -axis of accelerometer signal. The frequency energy feature is calculated as the sum of the squared magnitudes of the discrete FFT components of the signal. Entropy is calculated as the normalized information entropy magnitudes of the discrete FFT components of the signal.

Figure 4 shows some of these features for different activities. We can find that it is easy to distinguish some activities, such as sitting and standing, but the differences among walking, going upstairs, and going downstairs are not significant, likely to cause misrecognition.

2.4. *Kernel Discriminant Analysis (KDA)*. In order to improve the classification accuracy, further operations on the derived features are performed.

Linear discriminant analysis (LDA) is a supervised dimensionality reduction technique used for data analysis and pattern recognition. LDA tries to maximize the separation between different classes and minimize the separation

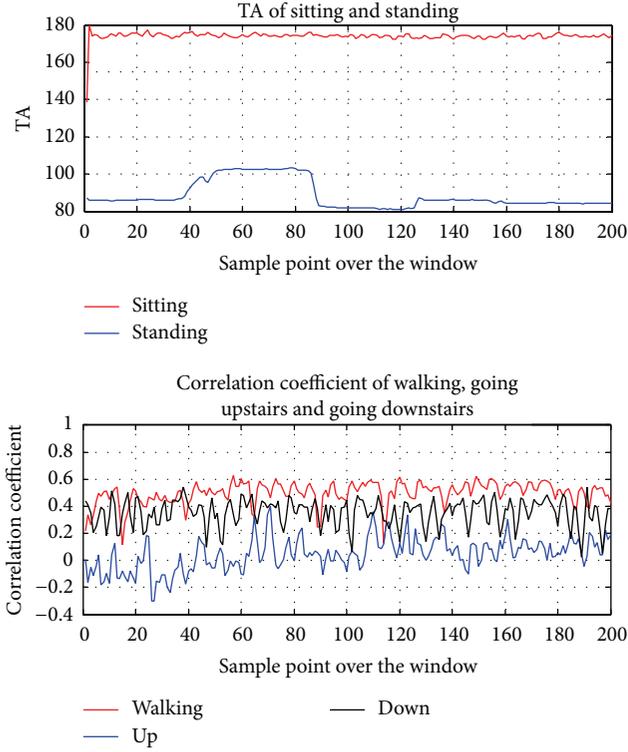


FIGURE 4: Difference in feature values for discriminating different activities.

within the same class simultaneously. KDA is a nonlinear extension of LDA to obtain nonlinear discriminating features by the kernel technique [26]. In KDA, input data are mapped to the high dimensional feature space F by nonlinear feature mapping $\phi : R^n \rightarrow F$. We select Gaussian radial basis function (RBF) for ϕ . In KDA, the cross class and intraclass scatter matrices are computed as

$$S_B^\phi = \sum_{i=1}^c N_i (\mu_i^\phi - \mu^\phi) (\mu_i^\phi - \mu^\phi)^T, \quad (4)$$

$$S_W^\phi = \sum_{i=1}^c \sum_{x \in X_i} (\phi(x) - \mu_i^\phi) (\phi(x) - \mu_i^\phi)^T,$$

where N_i is the number of samples from the i th class, c is the number of classes, μ_i^ϕ is the centroid of the i th class and μ^ϕ is the global centroid, x is a vector for a specific class, and X_i is the set of samples of the i th class. S_W^ϕ represents the degree of scattering within classes of activities and is calculated as the summation of covariance matrices of each class, whereas S_B^ϕ represents the degree of scattering between classes of activities and is calculated as the summation of the covariance matrix of the means of each class. The optimal discrimination transformation in the projection space is obtained by solving the following optimization problem:

$$\omega_{\text{opt}} = \arg \max_{\omega} \frac{\omega^T S_B^\phi \omega}{\omega^T S_W^\phi \omega}. \quad (5)$$

The optimization solution ω_{opt} , corresponding to the largest eigenvalues λ , can be explained by the generalized eigenvalue problem:

$$S_B^\phi \omega_i = \lambda_i S_W^\phi \omega_i. \quad (6)$$

It is proved that the above equation can be equivalently represented as

$$A_{\text{opt}} = \arg \max_A \frac{A^T K W K A}{A^T K K A}. \quad (7)$$

The corresponding eigenvalue problem is represented as

$$K W K A = \lambda K K A, \quad (8)$$

where K is the kernel matrix with its element defined as

$$K_{ij} = k(x_i, x_j), \quad (9)$$

where $k(\cdot, \cdot)$ is a positive semidefinite kernel function and W is a matrix with its element defined as

$$W_{ij} = \begin{cases} \frac{1}{n_c}, & \text{if } x_i \text{ and } x_j \text{ belong to the } c\text{th class} \\ 0, & \text{otherwise.} \end{cases} \quad (10)$$

More details on KDA can be found in [26].

2.5. ELM Algorithm. Extreme learning machine (ELM) was introduced by Huang et al. [11], originally proposed for standard single-hidden layer feed-forward neural networks (SLFNs), with random hidden nodes, and has recently been extended to kernel learning as well [27]. Compared with traditional training methods, ELM has the advantages of high accuracy, fast learning speed, and good generalization property. It can provide a unified learning platform with widespread type of feature mappings and can be applied in regression and multiclass classification applications directly.

Figure 5 is the structure of ELM. The network consists of an input layer, a hidden layer, and an output layer. For N arbitrary distinct samples (x_i, t_i) , denote the input signal vector $x_i = [x_{i1}, x_{i2}, \dots, x_{im}]^T \in R^m$ and the output signal vector $t_i = [t_{i1}, t_{i2}, \dots, t_{in}]^T \in R^n$, and $w_j = [w_{j1}, w_{j2}, \dots, w_{jm}]^T$ is the weight vector connecting the j th hidden node and the input nodes, $\beta_j = [\beta_{j1}, \beta_{j2}, \dots, \beta_{jn}]^T$ is the weight vector connecting the j th hidden node and the output nodes, and b_j is the threshold of the j th hidden node.

The activation function of hidden layer neurons is $g(x)$, and the hidden layer has L neurons. There exist w_j , β_j , and b_j such that

$$\sum_{j=1}^L \beta_j g(w_j x_i + b_j) = t_i, \quad i = 1, \dots, N. \quad (11)$$

Denote the output of the network as

$$T = \begin{bmatrix} t_1^T \\ \vdots \\ t_N^T \end{bmatrix}_{N \times n}. \quad (12)$$

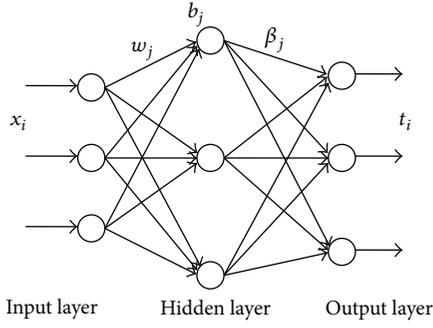


FIGURE 5: The structure of ELM.

Then (11) can be rewritten as

$$H\beta = T, \quad (13)$$

where H is the hidden layer output matrix of the ELM:

$$H(w_1, w_2, \dots, w_L, b_1, b_2, \dots, b_L, x_1, x_2, \dots, x_N) = \begin{bmatrix} g(w_1 x_1 + b_1) & g(w_2 x_1 + b_2) & \dots & g(w_L x_1 + b_L) \\ g(w_1 x_2 + b_1) & g(w_2 x_2 + b_2) & \dots & g(w_L x_2 + b_L) \\ \vdots & \vdots & \ddots & \vdots \\ g(w_1 x_N + b_1) & g(w_2 x_N + b_2) & \dots & g(w_L x_N + b_L) \end{bmatrix}_{N \times L},$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_L^T \end{bmatrix}_{L \times n}.$$

According to the theorem proven in [11], when the activation function is infinitely differentiable, and $L \leq N$, the parameters of SLFNs do not need all to be adjusted. Parameters w and b can be selected randomly before training and remain constant during the training process. To train SLFNs is simply equivalent to finding a least-square solution $\hat{\beta}$ of the linear system $H\beta = T$; that is,

$$\|H\hat{\beta} - T\| = \min_{\beta} \|H\beta - T\|. \quad (15)$$

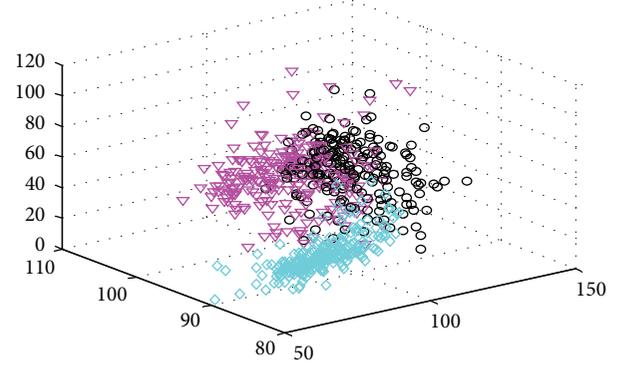
Then the optimal output weights can be calculated as

$$\beta = H^\dagger T, \quad (16)$$

where H^\dagger is the Moore-Penrose generalized inverse of the matrix H .

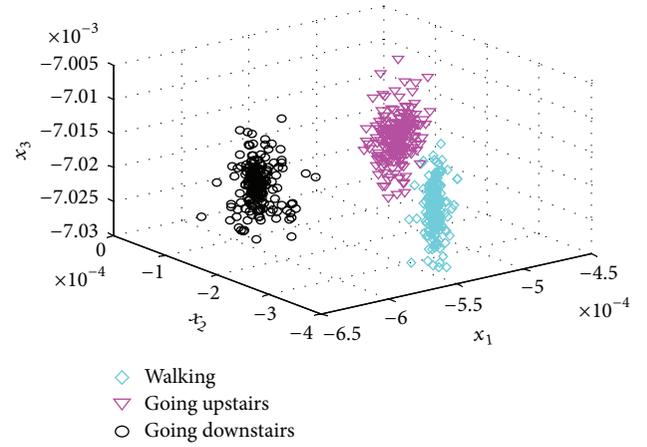
3. Approach Implementation and Experimental Results

In this section, we will implement the approach for extracting features of the data, using KDA to map the features to the high dimensional feature space, applying ELM algorithm to classify the samples and comparing ELM with other classic classifiers.



- ◆ Walking
- ▼ Going upstairs
- Going downstairs

(a) 3D feature space representation on original features



(b) 3D feature space representation for KDA implementation on original features

FIGURE 6: Features without and with KDA operations.

3.1. Recognition Using KDA on Original Features. The difference of the features among the three activities (walking, going upstairs, and going downstairs) is not so significant. Therefore, KDA is implemented to deal with the problem. The 3D plot in Figure 6 shows the low intraclass variance and high cross class variance for the three activities. Figure 6(a) shows the three features, SMA, TA, and WE, extracted from the collected samples. It is difficult to classify them. Figure 6(b) shows significantly improved discrimination among the three activities after the KDA operation.

3.2. Classification Results. In this subsection, we will compare the classification performance among ELM, BP, and SVM classifiers. All the experiments for the algorithms are carried out in MATLAB 8.0 environment running in an Inter i5, 2.6 GHz CPU. There are many variants of BP and SVM algorithms. Levenberg Marquardt BP (LM-BP) and least-square SVM (LS-SVM) [28] are used in this paper.

The data are randomly divided into two sets, namely, the training set and the testing set. The classification process

consists of two parts, generating the model by the training set and testing the performance of the model by the testing set. In our experiments, all the inputs (attributes) have been normalized into the range $[0, 1]$. Table 1 shows the confusion matrices between the 6 classes of activities (including sitting, standing, walking, running, going upstairs, and going downstairs) for the algorithms. The row represents the actual class and the column represents the recognized class by the algorithms, and the element at the i th row and j th column of the confusion matrix represents the probability of the actual class i is recognized as class j by an algorithm. The probability was calculated by 1000 samples over the window of six activities. From the matrices, we can find that ELM can distinguish the 6 classes of activities better than LM-BP and LS-SVM in general, and the KDA implementation on original features can improve the classification performance.

Furthermore, 50 trials have been conducted for a thorough comparison study and the average results are outputted as the classification results as shown in Table 2. The classifier of KDA based ELM takes the shortest testing time (0.0012 s) and achieves the highest testing accuracy (99.81%). Although KDA based ELM takes slight more training time than the original ELM, it has better classification result than the original ELM. That is because KDA can achieve the low intraclass variance and high cross class variance for activities and improve the classification accuracy effectively.

As shown in Table 2, KDA based LM-BP can classify samples faster than the original LM-BP and has higher classification accuracy. Also KDA based LS-SVM can improve the accuracy. We can find that the KDA strategy is useful in improving the recognition performance.

Now we give the comparison among ELM, BP, and SVM classifiers. It can be found from Table 2 that the classification accuracy of ELM, LM-BP, and LS-SVM based on the original features is not significantly different, but ELM has much faster learning speed (up to hundreds times) than LM-BP and LS-SVM. Based on the KDA implementation, ELM and LS-SVM can achieve better classification performance than LM-BP, but LS-SVM needs longer time for training than ELM. Also, the parameters of LS-SVM and LM-BP need to be determined before training, while ELM selects parameters randomly before training. Overall, the experiments demonstrate that ELM can achieve superior recognition performance compared with the LM-BP and LS-SVM classifiers.

4. Conclusion

This paper develops a highly efficient approach for human activity recognition based on ELM and using only one triaxial accelerometer. A number of features, respectively, in the time domain, the frequency domain, and the time-frequency domain, are defined and extracted from the raw measurement signals. KDA is performed on the original features to achieve the low intraclass variance and high cross class variance for activities. ELM classifier is proposed to classify the activities. Experimental results show that KDA based classifier can improve the classification accuracy effectively and ELM can achieve superior recognition performance compared with SVM and BP classifiers.

TABLE 1: The confusion matrices between the 6 classes of activities. R indicates running, ST indicates sitting, SD indicates standing, W indicates walking, U indicates going upstairs, and D indicates going downstairs.

(a) ELM based on original features						
	R	ST	SD	W	U	D
R	1	0	0	0	0	0
ST	0	1	0	0	0	0
SD	0	0	0.999	0	0	0.001
W	0	0	0	0.985	0.005	0.01
U	0	0	0	0.015	0.968	0.017
D	0	0	0	0.021	0.009	0.97
(b) ELM based on KDA features						
	R	ST	SD	W	U	D
R	1	0	0	0	0	0
ST	0	1	0	0	0	0
SD	0	0	1	0	0	0
W	0	0	0	1	0	0
U	0	0	0	0	0.999	0.001
D	0	0	0	0.001	0	0.999
(c) LM-BP based on original features						
	R	ST	SD	W	U	D
R	0.998	0	0	0.002	0	0
ST	0	0.998	0	0	0.002	0
SD	0	0	0.996	0.003	0	0.001
W	0	0	0	0.98	0.006	0.014
U	0	0	0	0.018	0.962	0.02
D	0	0	0.006	0.024	0.012	0.958
(d) LM-BP based on KDA features						
	R	ST	SD	W	U	D
R	1	0	0	0	0	0
ST	0	1	0	0	0	0
SD	0	0	1	0	0	0
W	0	0	0	0.99	0.006	0.004
U	0	0	0	0.01	0.988	0.002
D	0	0	0	0.009	0.006	0.985
(e) LS-SVM based on original features						
	R	ST	SD	W	U	D
R	0.993	0	0	0.006	0	0.001
ST	0	0.996	0	0	0.004	0
SD	0	0	0.997	0.001	0	0.002
W	0	0	0	0.981	0.006	0.013
U	0	0	0	0.01	0.976	0.014
D	0	0	0	0.02	0.012	0.968
(f) LS-SVM based on KDA features						
	R	ST	SD	W	U	D
R	1	0	0	0	0	0
ST	0	1	0	0	0	0
SD	0	0	1	0	0	0
W	0	0	0	0.998	0	0.002
U	0	0	0	0.003	0.993	0.004
D	0	0	0	0.002	0.002	0.996

TABLE 2: Classification results of ELM, LM-BP, and LS-SVM classifiers.

Features	Classifier	Time (s)		Accuracy rate (%)	
		Training	Testing	Training	Testing
Original	ELM	0.0022	0.0031	96.44	96.07
KDA based	ELM	0.0031	0.0012	99.96	99.81
Original	LM-BP	1.8985	0.0175	95.59	95.24
KDA based	LM-BP	0.8861	0.0246	98.48	98.02
Original	LS-SVM	0.7775	0.1819	97.36	95.97
KDA based	LS-SVM	0.7678	0.2760	99.94	99.05

Further research is required to determine the most appropriate feature set for more specific subject groups, such as the elderly or the neurologically impaired. In the future, we would also like to examine more complex physical activities and apply the ELM algorithm to solve more complex classification problems.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors thank Professor Jiankang Wu for his valuable comments, Jian Cui, Peiyuan Chen, and Yuejie Lu for their work related to this research issue, and Dr. Lianying Ji for his insightful suggestions.

References

- [1] R.-F. Li, L.-L. Wang, and K. Wang, "A survey of human body action recognition," *Pattern Recognition and Artificial Intelligence*, vol. 27, no. 1, pp. 35–48, 2014.
- [2] Q. R. Sun, W. M. Wang, and H. Liu, "Study of human action representation in video sequences," *CAAI Transaction on Intelligent Systems*, vol. 8, pp. 189–198, 2013.
- [3] N. Kern, B. Schiele, and A. Schmidt, "Multi-sensor activity context detection for wearable computing," in *Ambient Intelligence*, vol. 2875 of *Lecture Notes in Computer Science*, pp. 220–232, Springer, Berlin, Germany, 2003.
- [4] L. Bao and S. S. Intille, "Activity recognition from user-annotated acceleration data," in *Pervasive Computing*, vol. 3001 of *Lecture Notes in Computer Science*, pp. 1–17, Springer, Berlin, Germany, 2004.
- [5] U. Maurer, A. Smailagic, D. P. Siewiorek, and M. Deisher, "Activity recognition and monitoring using multiple sensors on different body positions," in *Proceedings of the International Workshop on Wearable and Implantable Body Sensor Networks (BSN '06)*, pp. 113–116, April 2006.
- [6] M. J. Mathie, A. C. F. Coster, N. H. Lovell, and B. G. Celler, "Detection of daily physical activities using a triaxial accelerometer," *Medical & Biological Engineering & Computing*, vol. 41, no. 3, pp. 296–301, 2003.
- [7] A. M. Khan, Y.-K. Lee, S. Y. Lee, and T.-S. Kim, "A triaxial accelerometer-based physical-activity recognition via augmented-signal features and a hierarchical recognizer," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 5, pp. 1166–1172, 2010.
- [8] M. Li, V. Rozgić, G. Thatté et al., "Multimodal physical activity recognition by fusing temporal and cepstral information," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 18, no. 4, pp. 369–380, 2010.
- [9] S. J. Preece, J. Y. Goulermas, L. P. J. Kenney, and D. Howard, "A comparison of feature extraction methods for the classification of dynamic activities from accelerometer data," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 3, pp. 871–879, 2009.
- [10] T. Ishii and S. Abe, "Feature selection based on kernel discriminant analysis for multi-class problems," in *Proceedings of the International Joint Conference on Neural Networks (IJCNN '08)*, pp. 2455–2460, June 2008.
- [11] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: a new learning scheme of feedforward neural networks," in *Proceedings of the IEEE International Joint Conference on Neural Networks*, vol. 2, pp. 985–990, July 2004.
- [12] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: theory and applications," *Neurocomputing*, vol. 70, no. 1–3, pp. 489–501, 2006.
- [13] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 42, no. 2, pp. 513–529, 2012.
- [14] G. B. Huang, "An insight into extreme learning machines: random neurons, random features and kernels," *Cognitive Computation*, vol. 6, no. 3, pp. 376–390, 2014.
- [15] E. Cambria, G. B. Huang, L. L. C. Kasun et al., "Extreme learning machines [trends & controversies]," *IEEE Intelligent Systems*, vol. 28, no. 6, pp. 30–59, 2013.
- [16] J. W. Cao, T. Chen, and J. Fan, "Fast online learning algorithm for landmark recognition based on BoW framework," in *Proceedings of the IEEE 9th Conference on Industrial Electronics and Applications (ICIEA '14)*, pp. 1163–1168, Hangzhou, China, June 2014.
- [17] W. Xiao, P. Liu, W.-S. Soh, and G.-B. Huang, "Large scale wireless indoor localization by clustering and extreme learning machine," in *Proceedings of the 15th International Conference on Information Fusion (FUSION '12)*, pp. 1609–1614, September 2012.
- [18] J. Cao, Z. Lin, G.-B. Huang, and N. Liu, "Voting based extreme learning machine," *Information Sciences*, vol. 185, no. 1, pp. 66–77, 2012.

- [19] W. Xiao, P. Liu, W.-S. Soh, and Y. Jin, "Extreme learning machine for wireless indoor localization," in *Proceedings of the 11th ACM/IEEE Conference on Information Processing in Sensing Networks (IPSN '12)*, pp. 101–102, April 2012.
- [20] H. G. Zhang, S. Zhang, and Y. X. Yin, "An improved ELM algorithm based on EM-ELM and ridge regression," in *Proceedings of the International Conference on Intelligence Science and Big Data Engineering*, Lecture Notes in Computer Science, pp. 756–763, Springer, 2013.
- [21] H. G. Zhang, S. Zhang, and Y. X. Yin, "A novel improved ELM algorithm for a real industrial application," *Mathematical Problems in Engineering*, vol. 2014, Article ID 824765, 7 pages, 2014.
- [22] J. Cao and L. Xiong, "Protein sequence classification with improved extreme learning machine algorithms," *BioMed Research International*, vol. 2014, Article ID 103054, 12 pages, 2014.
- [23] D. Wang, D. Li, and Y. Lin, "A new method of face recognition with data field and PCA," in *Proceedings of the IEEE International Conference on Granular Computing (GrC '13)*, pp. 320–325, December 2013.
- [24] A. Iosifidis, A. Tefas, and I. Pitas, "Activity-based person identification using fuzzy representation and discriminant learning," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 2, pp. 530–542, 2012.
- [25] Z. A. Khan and W. Sohn, "Hierarchical human activity recognition system based on R-transform and nonlinear kernel discriminant features," *Electronics Letters*, vol. 48, no. 18, pp. 1119–1120, 2012.
- [26] B. Schölkopf and A. J. Smola, *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*, MIT press, 2002.
- [27] G.-B. Huang and C.-K. Siew, "Extreme learning machine: RBF network case," in *Proceedings of the 8th International Conference on Control, Automation, Robotics and Vision*, vol. 2, pp. 1029–1036, December 2004.
- [28] J. A. K. Suykens and J. Vandewalle, "Least squares support vector machine classifiers," *Neural Processing Letters*, vol. 9, no. 3, pp. 293–300, 1999.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

