Activity Recognition Using One Triaxial Accelerometer: A Neuro-fuzzy Classifier with Feature Reduction

Jhun-Ying Yang[1], Yen-Ping Chen[1], Gwo-Yun Lee[2], Shun-Nan Liou[2], and Jeen-Shing Wang[1]

[1] Department of Electrical Engineering,
National Cheng Kung University,
Tainan 701, Taiwan
[2] Micro Systems Technology Research Laboratories,
Industrial Technology Research Institute,
Tainan 709, Taiwan
`jeenshin@mail.ncku.edu.tw`

Abstract. This paper presents a neuro-fuzzy classifer for activity recognition using one triaxial accelerometer and feature reduction approaches. We use a triaxial accelerometer to acquire subjects' acceleration data and train the neuro-fuzzy classifier to distinguish different activities/movements. To construct the neuro-fuzzy classifier, a modified mapping-constrained agglomerative clustering algorithm is devised to reveal a compact data configuration from the acceleration data. In addition, we investigate two different feature reduction methods, a feature subset selection and linear discriminate analysis. These two methods are used to determine the significant feature subsets and retain the characteristics of the data distribution in the feature space for training the neuro-fuzzy classifier. Experimental results have successfully validated the effectiveness of the proposed classifier.

Keywords: Acceleration, activity recognition, feature extraction, linear discriminate analysis, neuro-fuzzy system, triaxial accelerometer.

1 Introduction

The development of wearable systems using advanced miniature sensors and wireless technologies has allowed people to issue a command for controlling electric appliances by gestures/activities. Activity recognition is one of the technologies frequently embedded in wearable systems for recognizing human activities or gestures. Nowadays, many researchers have focused on diversifying application domains of human activity recognition in biomedical engineering, medical nursing, and interactive entertainment. Signals for recognition can be obtained from different kinds of detectors. These include electromyography (EMG), audio sensors, image sensors, and accelerometers. Among the aforementioned sensors, accelerometers that can record acceleration data caused by movements and gravity have drawn much attention in the field of activity/gesture recognition [1].
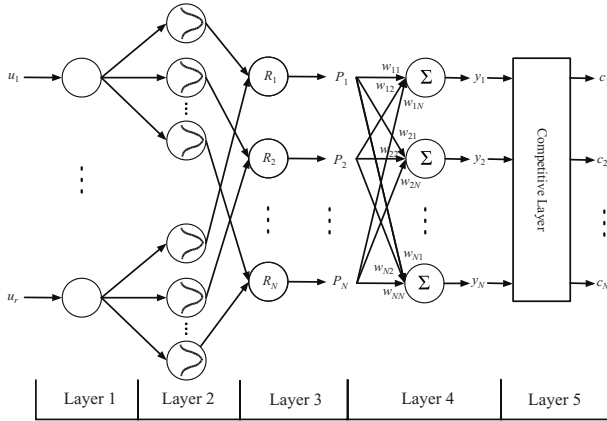
Fig. 1. Topology of the neuro-fuzzy classifier

Recently, research studies have proposed a variety of classifiers for activity recognition from acceleration data. These methodologies used to build classifiers include support vector machine (SVM) [4], [6], K-nearest neighbors [4], decision tree, [3], naïve Bayes classifiers [1], [4], hidden Markov models [8], and neural networks [6]. In this paper, we adopt a neuro-fuzzy classifier to recognize human activities using one triaxial accelerometer to acquire a subject's acceleration data. In addition, we compare two different dimensionality reduction methods in our recognition scheme: one is the feature selection method proposed in [2] and the other is linear discriminate analysis (LDA) [5].

The rest of this paper is organized as follows. In section 2, we introduce the structure of a neuro-fuzzy classifier and its construction algorithm. Section 3 presents the detailed information about the proposed recognition strategy, including feature extraction and dimensionality reduction. Section 4 provides the experiment design for validating the effectiveness of the proposed classifier. Finally, conclusions are given the last section.

2   Neuro-fuzzy Classifier

In this paper, we adopt a neuro-fuzzy classifier to recognize daily activities from acceleration data. The structure of the neuro-fuzzy classifier is shown in Fig. 1. It contains five layers and the description of the layers is given as follows. **Layer 1**: The nodes in this layer only transmit input values to the nodes of the next layer directly. **u** = $[u_1, u_2, ..., u_r]^T$ is the input vector, where r represents the number of elements in the input feature set. **Layer 2**: The nodes represent Gaussian membership functions:

$\mu_{M_i^j} = \exp(-(\frac{u_i - m_i^j}{\sigma_i^j})^2)$, where M_i^j is the input fuzzy term set, m_i^j and σ_i^j are the center and width of the Gaussian membership function of the jth term of the ith input variable u_i. **Layer 3**: The nodes in this layer constitute the antecedents of the fuzzy rule base.

The output of the node is $p_j = \prod_{i=1}^{n} \mu_{M_i^j}$, where p_j is the output function of the jth rule node. **Layer 4**: This is the output layer. The output nodes integrate all the inferred information from Layer 3 with the corresponding singleton values w_{hj} and act as a defuzzifier. $y_h = \sum_{j=1}^{N} w_{hj} p_j$, where N is the number of classes and $\mathbf{y} = [y_1, y_2, ..., y_N]^T$ are the output vector. We set the numbers of the fuzzy rules and output neurons equal to the number of activity classes. The fuzzy rules can be expressed as

$$\text{Rule } j\text{: } \mathbf{IF} \ u_1 \text{ is } M_1^j \text{ and } \cdots \text{ and } u_r \text{ is } M_r^j,$$
$$\mathbf{THEN} \ y_1(k) = w_{1j} p_j \text{ and } \cdots \text{ and } y_N = w_{Nj} p_j. \tag{1}$$

Layer 5: This is the competitive layer. Each element in the output vector competes with the others. We set the winner's output be 1 and zeros for the rest outputs.

In general, a clustering technique is appropriate to determine the parameters of the neuro-fuzzy structure. To effectively construct the network structure, we modify the mapping-constrained agglomerative (MCA) clustering algorithm [7] that takes the number of fuzzy rules as the number of clusters in the input and output spaces. The parameters of these clusters including the input centers ($\mathbf{m}$), variances ($\boldsymbol{\sigma}$), and output centers ($\mathbf{W}$), can be directly used to construct the neuro-fuzzy classifier. For each cluster, say cluster α, the components of $\mathbf{m}_{IO\alpha} = [\mathbf{m}_{I\alpha}, \mathbf{m}_{O\alpha}]^T$ ($\boldsymbol{\sigma}_{IO\alpha} = [\boldsymbol{\sigma}_{I\alpha}, \boldsymbol{\sigma}_{O\alpha}]^T$) denote the centers (variances) in the input space (I) and output space (O), respectively, and let z_α be the counter representing the number of patterns in cluster α. Given a set of training patterns $\mathbf{V} = [\mathbf{u}, \mathbf{y}]^T$, we first assign the initial seed clusters by randomly selecting the data points in the training patterns and then feed the data points as the training patterns. The seed cluster closest to the current data point is defined as the winner cluster. We perform $winner = \arg\min \left\{ \|\mathbf{V} - \mathbf{m}_{IO\alpha}\| \right\}, \alpha = 1, ..., N$, and update the parameters of the winner cluster: Set the counter of data points $z_{winner} = z_{winner} + 1$. For the input space, set

$$temp\_\mathbf{m}_{Iwinner} = \mathbf{m}_{Iwinner} + \frac{\mathbf{u} - \mathbf{m}_{Iwinner}}{z_{winner}};$$
$$\boldsymbol{\sigma}^2_{Iwinner} = \frac{(z_{winner} - 1)(\boldsymbol{\sigma}^2_{Iwinner} + \mathbf{m}^2_{Iwinner}) + \mathbf{u}^2}{z_{winner}} - temp\_\mathbf{m}^2_{Iwinner}; \tag{2}$$
$$\mathbf{m}_{Iwinner} = temp\_\mathbf{m}_{Iwinner}.$$

Similarly, for the output space, set $\mathbf{m}_{Owinner} = \mathbf{m}_{Owinner} + \frac{\mathbf{y} - \mathbf{m}_{Owinner}}{z_{winner}}$. Repeat the above steps until there remains no data to be clustered. The above modified MCA algorithm determines the parameters of $\mathbf{m}$, $\boldsymbol{\sigma}$ and $\mathbf{W}$ by setting $\mathbf{m} = \mathbf{m}_I$, $\boldsymbol{\sigma} = \boldsymbol{\sigma}_I$ and $\mathbf{W} = \mathbf{m}_O$.

3 Activity Recognition Strategy

We have developed an effective activity recognition strategy to increase the recognition accuracy and reduce the computational burden of the neuro-fuzzy classifier. First, we cut the raw acceleration sequences into many overlapping windows of the same length and extract features from each window into a feature set.

In order to reduce the dimension of the feature set, we investigate two dimensionality reduction methods in Section 3.2.

3.1 Feature Extraction

We extract the features in time domain and frequency domain from each window of the triaxial acceleration data. The time domain features are extracted from the raw data and include mean, correlation between axes, interquartile range, mean absolute deviation, root mean square, standard deviation, and variance. The frequency domain feature is energy calculated as the sum of the squared discrete FFT component magnitudes of the signal in a window [6]. Since we collect signals from a triaxial accelerometer, a total of 24 features (3 axes × 8 features) are calculated from a window of the acceleration data for x-axis, y-axis, and z-axis.

3.2 Dimensionality Reduction

In general, the following two approaches are usually used in dimensionality reduction: 1) Selecting the significant features and eliminating irrelative ones to preserve as much of the original information as possible for recognition; and 2) Transforming original feature sets into a lower dimensional feature space with class separability. In this study, we investigate two approaches, a feature subset selection (FSS) [2] and linear discriminate analysis (LDA) [5].

Feature subset selection. The FSS is based on common principal components (CPCs) generalized from the principal component analysis (PCA). First, the FSS performs PCA on each class to obtain the CPC loadings [2]. Each row vector in the CPC loadings represents the projection of the corresponding feature of the classes to a lower dimensional common space. Then, a clustering technique is utilized to group the row vectors of the CPC loadings which have similar contribution in the data distribution. We select the points closest to the centers of clusters as the selected features. That is, each data point represents a feature and the number of the selected features equals to the number of clusters.

Linear discriminate analysis. The basic concept of the LDA is to seek the most efficient projective direction which minimizes the data distribution in the same class and separates the data distribution in the different classes for discrimination. For this purpose, two scatter matrices are defined: the between-class covariance matrix $\mathbf{S}_B$ and the within-class covariance matrix $\mathbf{S}_W$. $\mathbf{S}_B$ shows the scatter of the expected vectors around the mixture mean:

$$\mathbf{S}_B = \sum_{\alpha=1}^{N} n_\alpha (\mathbf{m}^{(\alpha)} - \mathbf{m})(\mathbf{m}^{(\alpha)} - \mathbf{m})^T, \qquad (3)$$

and $\mathbf{S}_W$ represents the scatter of samples around their respective class expected vectors:

$$\mathbf{S}_W = \sum_{\alpha=1}^{N} n_\alpha \sum_{i=1}^{n_\alpha} (\mathbf{x}_i^{(\alpha)} - \mathbf{m}^{(\alpha)})(\mathbf{x}_i^{(\alpha)} - \mathbf{m}^{(\alpha)})^T, \qquad (4)$$

where n_α is the number of samples in the class α, $x_i^{(\alpha)}$ is the ith sample of αth class, $\mathbf{m}^{(\alpha)}$ is the mean vector of the samples in the class α, and $\mathbf{m}$ is the mean vector of all the data points. The LDA preserves class separability in a lower dimensional space by finding a unit projective vector $\mathbf{w}$ which maximizes the covariance between classes and minimizes the covariance within class by maximizing the following criterion [5]:

$$J(\mathbf{w}) = \frac{\mathbf{w}^T \mathbf{S}_B \mathbf{w}}{\mathbf{w}^T \mathbf{S}_W \mathbf{w}}. \tag{5}$$

4   Experimental Results

The acceleration data was collected using the MMA7260Q triaxial accelerometer on a wearable board. The accelerometer's sensitivity is set from –4.0g to +4.0g and the output signal of the acceleometer is sampled at 100 Hz by a 10-bit ADC.

The classification tasks include eight common domestic activities: walking, running, scrubbing, standing, working at a computer, vacuuming, brushing teeth and sitting. We gathered acceleration data from a single triaxial accelerometer module mounted on the dominant wrist of each subject. All the subjects (seven normal and healthy subjects) were asked to perform each activity for two minutes. We took the window size of 512 with 256 samples overlapped with consecutive windows. A total of 45 windows were obtained from the acceleration data for each activity of one subject.

We utilized a leave-one-subject-out cross-validation procedure to validate the effectiveness of the proposed activity recognition strategy. Six subjects were trained in the recognition scheme and then tested on the subject left out of the training data set. We repeated the same procedure for all the subjects.

The LDA and FSS methods were applied to reduce the dimensionality of the feature space, and then we utilized the lower dimensional feature sets to train and test the proposed neuro-fuzzy classifier. The original feature space was transformed into a new seven-dimensional feature space after performing the LDA. The average cross-validation recognition accuracy achieves 92.86±5.91%. Table 1 shows the confusion matrix that records the number of recognition errors for all the subjects. To compare the performance of the LDA and FSS methods, the FSS method selected seven features from the original 24 features. The average recognition accuracy is

Table 1. Confusion matrix for all subjects

Classified as	Walking	Running	Scrubbing	Standing	Working at a PC	Vacuuming	Brushing teeth	Sitting
Walking	304	0	2	0	0	10	0	0
Running	0	265	0	0	0	1	0	0
Scrubbing	0	10	279	0	0	2	16	0
Standing	0	0	0	306	0	0	0	0
Working at a PC	0	0	0	0	315	0	0	45
Vacuuming	11	31	5	9	0	302	0	0
Brushing teeth	0	6	29	0	0	0	299	0
Sitting	0	3	0	0	0	0	0	270

83.41±5.93%. The LDA method outperforms the FSS method. The recognition results show that the proposed activity recognition strategy can provide satisfactory accuracy by the proposed neuro-fuzzy classifier.

5   Conclusions

In this study, we have developed a human activity recognition scheme based on a neuro-fuzzy classifier using acceleration data acquired by a single accelerometer. We employed the modified MCA and two dimensionality reduction methods to construct the neuro-fuzzy classifier. The performances of the proposed classifiers with the FSS and LDA achieve 83% and 93%, respectively. If accuracy is the major concern for building the classifier, the construction algorithm can be associated with optimization techniques to further fine-tune the parameters of the neuro-fuzzy classifiers in improve the overall performance of the classifier.

References

1. Bao, L., Intille, S.S.: Activity Recognition from User-Annotated Acceleration Data. In: Ferscha, A., Mattern, F. (eds.) PERVASIVE 2004. LNCS, vol. 3001, pp. 1–17. Springer, Heidelberg (2004)
2. Chen, Y.-P., Yang, J.-Y., Liou, S.-N., Lee, G.-Y., Wang, J.-S.: Neural Classifiers for Activity Recognition Using Acceleration Measurements. In: Applied Mathematics and Computation (to appear)
3. Mathie, M.J., Celler, B.G., Lovell, N.H., Coster, A.C.F.: Classification of Basic Daily Movements Using a Triaxial Accelerometer. Medical and Biological Engineering and Computing 42(5), 679–687 (2004)
4. Ravi, N., Dandekar, N., Mysore, P., Littman, M.L.: Activity Recognition from Accelerometer Data. In: Proceedings of the Seventeenth Innovative Applications of Artificial Intelligence Conference, pp. 1541–1546 (2005)
5. Tsymbal, A., Puuronen, S., Pechenizkiy, M., Baumgarten, M., Patterson, D.: Eigenvector-Based Feature Extraction for Classification. In: Proceedings of the Fifteenth International Florida Artificial Intelligence Research Society Conference, pp. 354–358.
6. Wang, S., Yang, J., Chen, N., Chen, X., Zhang, Q.: Human Activity Recognition with User-Free Accelerometers in the Sensor Networks. In: IEEE Int. Conf. Neural Networks and Brain, vol. 2, pp. 1212–1217 (2005)
7. Wang, J.-S., Lee, C.S.G.: Self-Adaptive Neuro-Fuzzy Inference Systems for Classification Applications. IEEE Trans. Fuzzy Systems 10(6), 790–802 (2002)
8. Ward, J.A., Lukowicz, P., Troster, G., Starner, T.E.: Activity Recognition of Assembly Task Using Body-Worn Microphones and Accelerometers. IEEE Trans. Pattern Analysis and Machine Intelligence 28(10), 1553–1567 (2006)