ACTIVITY RECOGNITION FROM ACCELERATION DATA USING AR MODEL REPRESENTATION AND SVM

ZHEN-YU HE, LIAN-WEN JIN

School of Electronic and Information Engineering，
South China University of Technology, Guangzhou, China. 510640
E-MAIL: eelwjin@scut.edu.cn

Abstract:

   In this paper, the autoregressive (AR) model of time-series is presented to recognize human activity from a tri-axial accelerometer data. Four orders of autoregressive model for accelerometer data is built and the AR coefficients are extracted as features for activity recognition. Classification of the human activities is performed with Support Vector Machine (SVM). The average recognition results for four activities (running, still, jumping and walking) using the proposed AR-based features are 92.25%, which are better than using traditional frequently used time domains features (mean, standard deviation, energy and correlation of acceleration data) and FFT features. The results show that AR coefficients obvious discriminate different human activities and it can be extract as an effective feature for the recognition of accelerometer date.

Keywords:

   Tri-axial accelerometer data; Activity recognition; Autoregressive model; Feature extraction; SVM

1. Introduction

   Context awareness is a central issue in ubiquitous and wearable computing [1]. Accurate recognition and tracking of human activities is an important goal of ubiquitous computing. Activity recognition is also one technology frequently embedded in wearable systems [3~8]. For example, several activities such as ambulation, typing, talking were distinguished in [6] with five small bi-axial accelerometers. In [7] and [8], daily activities of standing, walking, climbing up/down stairs and brushing teeth, were analyzed based on the data collected from accelerometers.

   Although a great number of progress have been made in the field of wearable computing, the users of most current wearable systems don't always feel comfortable when wearing sensors. Many of the existed activity recognition systems require the users to bring multiple sensors which have to be fixed to specific parts of his/her bodies to achieve a high degree of accuracy. In order to resolve these

issues, we attempt to recognize activities using a single tri-axial accelerometer. Rather than fixing the accelerometer to his body, the user can put the accelerometer in his/her trousers pocket. Therefore it is a more natural and friendly way to recognize human's activities.

   As activity recognition can be formulated as a typical classification problem and just like many pattern recognition problem, features extraction plays an crucial role during the recognition process. Although in the literature there are already many study on exploring the extraction of features from acceleration data, few works which make quantitative comparison of their quality are reported. In general, most of the attempts to extract features from acceleration date can be classified into two categories, say, time domains features and frequency domains features. Traditional widely used time domains features are mean [2,6,8], variance or standard deviation [2,6], energy [2,6,8], entropy [6], correlation between axes [2,6,8] and so on. The most popular frequency domains features are FFT coefficients[1]. However, FFT require much high components to discriminate different activities. Hence it will increase computation and do not suitable for real time application. As the time-domains features can be easily extracted in real time, they are more popular in many practical acceleration activity recognition systems.

   Although activity recognition using time-domains features were successful to some limit, the recognition results using these features have not had a high success rate because such methods assume that activity acceleration signals are deterministic. However, in fact, the activity acceleration signals are random signals in their nature. Thus we should establish a better mathematical model using stochastic time series analysis to describe these data. In this paper, we use the pure autoregressive (AR) model to represent the acceleration activity signals. Although the AR model is just a simple linear and second-moment stationary model, it has been successfully used in wide range of applications such as speech analysis, spectrum estimation,

and surface electromyography [9, 10]. However, to our knowledge there is no previous research works using AR model for activity recognition.

In this paper, a new AR model based features for activity recognition from a tri-axial accelerometer data has been proposed. Four orders of autoregressive model accelerometer data is built and the AR coefficients are chosen as the input features of the SVM classifier. The classification of four activities shows encouraging results.

The rest of this paper is organized as follows: In section 2, we introduce the AR model. Section 3 presents the detailed information about the proposed recognition strategy, including data collection, feature extraction based on AR modeling of acceleration signals and classification. Section 4 gives the experiment design for validating the effectiveness of the proposed features. Finally, conclusions are given in section 5.

2. AR Model Representation

It is well known that linear prediction filters can be used to model a second-order stationary random signal. In short time interval, such signal can be equivalent to the output of linear system which is power by the zero-mean white noise process. Hefftner [10] observed that any model can be approximated by an AR model of finite order n, with a specified degree of closeness. The order of an autoregressive model represents the amount of information necessary to predict an estimated signal that can be represented by the AR coefficients. The short duration acceleration data can indeed be a kind of stationary random signal. Therefore, it is very reasonable to use AR model to describe acceleration signals. Then we can use the AR model coefficients as features to perform the recognition of activity.

Auto-Regression (AR) model is expressed as the following difference equation [11] :

$$x(n) = -\sum_{k=1}^{p} a(k) * x(n-k) + u(n) \qquad (1)$$

Where *x(n)* is the random signal to be processed, the input $u(n)$ is assumed to be a white noise signal and p is the model order. The selection of the model order in AR model is a critical problem, too low an order is a smoothed estimation, while too large an order causes spurious peaks and general statistic instability. We use the methods proposed by Akaike [12] to determine the model order p, which aims to minimize the following Akaike information criterion (AIC) function:

$$AIC(k) = N \ln \hat{\rho}_k + 2k \qquad (2)$$

Where $\hat{\rho}_k$ is the estimate of the white noise variance for the *k* th order AR model.

Once the order of AR model has been determined, the AR coefficients $a_1 \cdots a_n$ can be estimated by an auto-correlation method, known as the Yule-Walker equations:

$$\begin{bmatrix} r_x(0) & r_x(1) & \cdots & r_x(p) \\ r_x(1) & r_x(0) & \cdots & r_x(p-1) \\ r_x(2) & r_x(1) & \cdots & r_x(p-2) \\ \cdots & \cdots & \cdots & \cdots \\ r_x(p) & r_x(p-1) \cdots & r_x(0) \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ a_2 \\ \cdots \\ a_p \end{bmatrix} = \begin{bmatrix} \sigma^2 \\ 0 \\ 0 \\ \cdots \\ 0 \end{bmatrix} \qquad (3)$$

It can be proved that matrix in (3) is hermitian and toeplitz. And the Levinson-Durbin algorithm [11] can be used to solve the equations. From the minimum-phase theorem, the resulting estimated poles are always within the unit circle by, therefore to guarantee the AR model to be stable.

3. Activity Recognition

3.1. Data Collection

The diagram of our experimental setup is shown in Fig.1. We used a tri-axial accelerometer ADXL330 manufactured by Analog Devices, which is capable of sensing accelerations from –3.0g to +3.0g with tolerances within 10%. The output signal of the accelerometer is sampled at 100 Hz. The data generated by the accelerometer was transmitted to a personal computer wirelessly over Bluetooth. Data from the accelerometer has the following attributes: time, acceleration along x axis, acceleration along y axis and acceleration along z axis. We collected four common domestic activities: running, still, jumping and walking. In order to achieve robustness with regard to sensor position, subjects put the accelerometer in their trousers pocket. As we don't fix the sensor with the body, it may move randomly in the pocket (such as rotation) and therefore produces more variations among different collectors. Eleven subjects (night male and two female) were asked to perform each activity about one minute. The activities were performed in two rounds over different days. The subjects keep standing still five seconds while every activity start and stop. Figure 2 shows the example of raw data and the corresponding colors (X-red, Y-green, Z-blue)
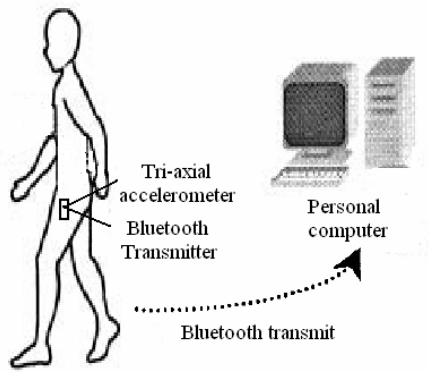
Figure 1. Diagram of experimental setup

3.2. Feature extraction

Features were extracted from the raw accelerometer data using a window size of 512 with 256 samples overlapping between consecutive windows. Feature extraction on windows with 50% overlap has demonstrated success in previous work [6]. At a sampling frequency of 100Hz, each window represents 5.12 seconds. According to our experience, a much shorter window can't seize the activity pattern properly. However, if it is too long, there will be a delayed response and not suitable for real time application. For each window, we computed the n order AR coefficients as feature.
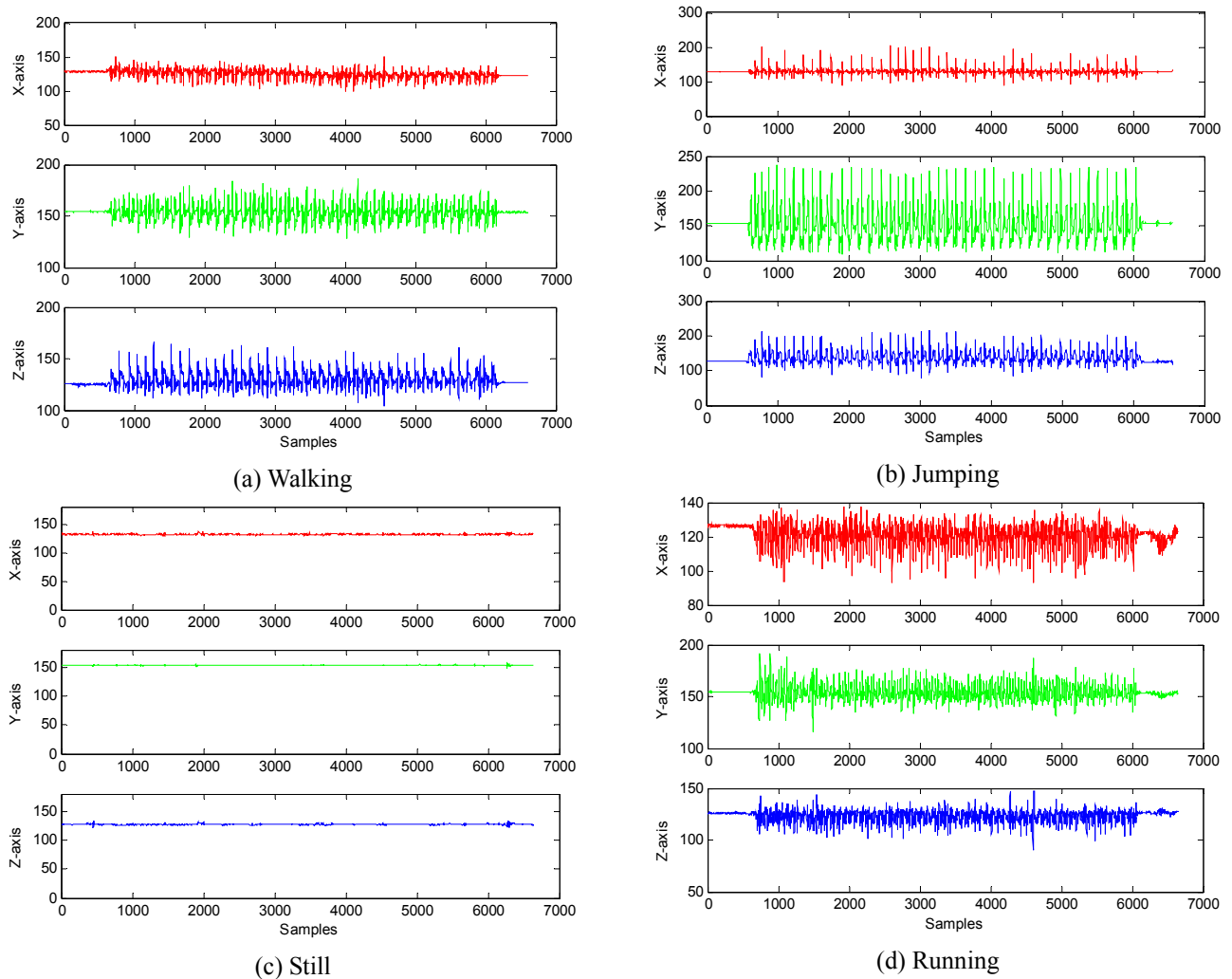

(a) Walking

(b) Jumping

(c) Still

(d) Running

Figure 2. Examples of raw signals for different activities

The optimal order *n* of the AR model is obtained according to equation (2). Fig. 3 shows the Akaike information criterion of acceleration data. It can be seen that the optimal order for accelerometer data is 3. From experiments we also found that the recognition result is hardly improved for orders above 3 (see Fig.4). However, in order to make our results comparable with traditional 12 dimension time-domains features and 12 FFT features, we chose 4 as AR model order here to produces the same number of features. In other word, the 4-order AR coefficients were extracted from each of the three axes of the accelerometer data, giving a total of 12 features.
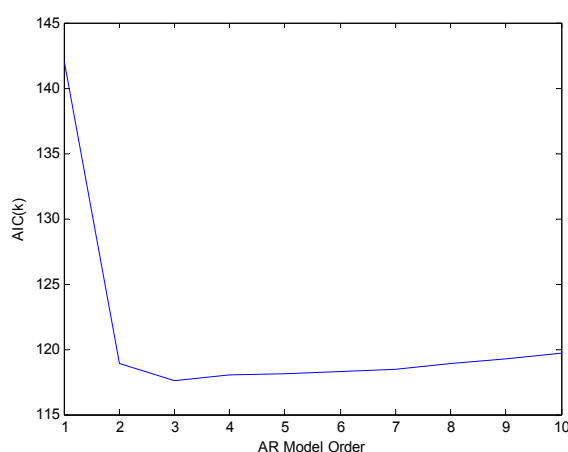

Figure 3. AIC in model order selection

In order to compare the performance of our new features against traditional features (TF), the following four kinds of traditional time-domains features (TF) were extracted from each axes of accelerometer (also result in total 12 features): mean (DC feature), standard deviation, energy and correlation between axes. The effectiveness of these features has been demonstrated in many prior work ［e.g. 6，8］. The DC feature is the mean acceleration value of the sliding window. Standard deviation was used to characterize the stability of the signal. The energy feature, which can capture the data periodicity, was calculated as the sum of the squared discrete FFT component magnitudes of the signal ［2］.The correlation feature aims to find out the relationship between axes in three directions which calculates between each pair of axes as the ratio of the covariance and the product of the standard deviations. Besides, for comparing with traditional frequency-domains features, we extracted the first 4 FFT coefficients from each axes of accelerometer data (also a total of 12 features).

3.3. Classification

The classification algorithm we used is Support Vector Machine (SVM) [13]. SVM has become one of the most popular classification methods in Machine Learning field in recent years. As SVM was originally designed for binary classification, it cannot deal with multi-class classification directly. The multi-class classification problem is usually solved by decomposition of the problem into several two-class problems. In this paper, we used One-versus-One Strategy (OVO), where a set of binary classifiers are constructed using corresponding data from two classes. While testing, we used the voting strategy of "Max-Wins" to produce the output.

Five-fold cross-validation was used for classifier assessment. The data was randomly divided into 5 groups with the same number of samples for different classes. The classifier was built ten times. Each time one group in turn was excluded from the training and used solely as a test set. The cross-validated classification result is the average of the ten testing results.

Since the training dataset is relative limited here, generalization capability of the classifier is more important for recognition. We also use the leave-one-subject-out validation test [2] to evaluate the classifiers' ability to recognize unacquainted actions. Classifiers were trained on activity data for all subjects except one. The classifiers were then tested on the data for only the subject left out of the training data set. This process was repeated for all subjects. In other words, the recognized are subject-independent.

4. Experiment Results and Discussion

For the first experiment, we carried out five-cross-validation procedure to validate the effectiveness of the proposed activity recognition strategy. The recognition result is given in Table 1.

Table 1. The accuracy of the five-cross-validated test

	run	still	jump	walk	average
TF	64.48	100	93.19	100	89.42
FFT	93.74	92.30	89.21	87.47	90.69
AR	80.69	100	90.05	98.29	*92.25*

It can be seen from Table 1 that accuracy using the proposed AR features is higher than using traditional time-domains features and FFT features. The recognition rate for activities running is lowest among four activities. This result is reasonable, because the raw signals of running are similar to the jumping. (see fig.2).

Secondly, we utilized a leave-one-subject-out

cross-validation test. The results are shown in Table 2. From table 2, we can also see that in the case of the leave-one-subject-out test, classification accuracy of the AR features is still better than four traditional time-domains features and FFT features. We noticed that the average accuracy of the leave-one-subject-out test is lower than that of the five-cross-validated test, because the leave-one-subject-out test is subject-independent. As we known, the characteristics of the gait signals is unique for every person, therefore generalization capability based on leave-one-subject-out training is better than cross-validated training. The FFT features particularly produce bad result, this may due to the reason that the number of FFT coefficients we used as features is too small (only 12), therefore many high frequency characteristic of the signal that may important to distinguish different subjects' activities are ignored.

Table 2. The accuracy of the leave-one-subject-out test

	run	still	jump	walk	average
TF	58.23	100	75.85	100	83.52
FFT	61.93	31.81	59.94	29.54	45.80
AR	69.03	100	81.53	96.87	86.86

In order to find out which activities are relatively harder to be recognized, we analyzed the confusion matrices, which give information about the actual and predicted classification results given by the classifiers. Table 3 shows the aggregate confusion matrix for the AR features based on all eleven trials of leave-one-subject-out validation. It can be seen that running is often confused with jumping and is in general hard to recognize.

Table 3. Confusion matrix for leave-one-subject-out test

	run	still	jump	walk
run	243	0	*58*	51
still	0	352	0	0
jump	*54*	0	287	11
walk	10	0	1	341

In order to validate the effects of model order from the aspect of discrimination capability, a series of experiments was carried out. Fig.4 shows the accuracy based on leave-one-subject-out test from 1-order to 10-order AR model. The curse shows that activity discrimination accuracy is hardly at all improved for orders above 3.
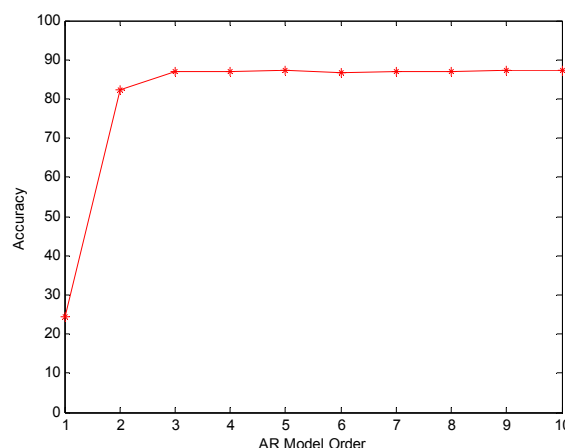

Figure 4. Recognition results vs. AR model order

Table 4. The typical AR coefficients corresponding to four activities

	run	still	jump	walk
a_{x1}	-1.0659	-0.9980	-1.0536	-1.0492
a_{x2}	0.0483	-0.0009	0.1834	0.0413
a_{x3}	0.0045	-0.0002	-0.0956	0.0179
a_{x4}	0.0167	0.0010	-0.0315	-0.0077
a_{y1}	-1.2806	-0.9990	-1.2502	-1.0421
a_{y2}	0.3308	0.0033	0.2766	0.0339
a_{y3}	0.0402	-0.0065	0.1410	0.0219
a_{y4}	-0.0857	0.0042	-0.1634	-0.0114
a_{z1}	-1.2184	-1.0009	-1.0907	-1.0741
a_{z2}	0.4709	0.0017	0.2699	0.0647
a_{z3}	-0.1840	0.0003	-0.0415	0.0357
a_{z4}	-0.0629	0.0009	-0.1343	-0.0240

Furthermore, the AR coefficients corresponding to four activities were obtained with the purpose of validating their discrimination. From the Table 4, we can find the phenomena that differences among AR coefficients corresponding to different activity patterns are obvious. This observation also confirms us that it is reasonable to extract AR coefficients as acceleration signal's features.

5. Conclusion

A new AR model based features for activity recognition from a tri-axial acceleration signals have been proposed in this paper. Four orders of autoregressive model for accelerometer signals is built and the AR coefficients are chosen as the input features of the SVM classifier. The average recognition results for four activities (running, still, jumping and walking) using the proposed AR-based features are 92.25%, which are better than using traditional widely used time domains features (mean, standard deviation, energy and correlation of acceleration data) and FFT features. The experiment results show that AR coefficients corresponding to different activities patterns are obvious discriminative. It is found from our study that the AR coefficients can provide sufficient discrimination between different types of human activity and provides new choice of features for activity recognition.

Acknowledgements

References

[1] Huynh. T. Schiele. B. Analyzing Features for Activity Recognition. Proceedings of the 2005 joint conference on Smart objects, 2005

[2] Wang, S., Yang, J., Chen, N., Chen, X., Zhang, Q., "Human Activity Recognition with User-Free Accelerometers in the Sensor Networks", IEEE Int. Conf. Neural Networks and Brain, vol. 2, pp. 1212–1217, 2005.

[3] G.D. Abowd, A.K. Dey, R. Orr, J.A. Brotherton, "Context-awareness in wearable and ubiquitous computing". Virtual Reality, vol.3, pp.200-211, 1998.

[4] Sekine, M., Tamura, T., Akay, M., ect al Fujimoto, T., Togawa, T., Fukui, Y., "Discrimination of Walking Patterns using Wavelet-Based Fractal Analysis", IEEE Transactions on Neural Systems and Rehabilitation Engineering，vol.10, PP.188–196, 2002.

[5] Randell C., Muller H., "Context awareness by analyzing accelerometer data". Fourth International Symposium on Wearable Computers, pp. 175-176, 2000.

[6] Bao L., Intille S.S., "Activity recognition from user-annotated acceleration data", PERVASIVE 2004. LNCS, vol. 300, PP.1-17, Apr. 2004.

[7] Mantyla J., Himberg J., Seppanen T., "Recognizing Human Motion with Multiple Acceleration Sensors", IEEE International Conference on Systems, Man and Cybernetics, Tucson, USA, vol. 3494, pp. 747-752, 2001.

[8] Ravi N., Dandekar N., Mysore P., Littman M.L., "Activity Recognition from Accelerometer Data", Proceedings of the Seventeenth Innovative Applications of Artificial Intelligence Conference, pp. 1541–1546, 2005.

[9] Zhizeng Luo, Fei Wang, Wenjie Ma, "Pattern Classification of Surface Electromyography Based on AR Model and High-order Neural Network", Proceedings of the 2nd IEEE/ASME International Conference on Mechatronic and Embedded Systems and Applications, PP. 1-6, Aug. 2006.

[10] Hefftner G., Zucchini W., Jaros,G G., "The Eletromyogram (EMG) as a Control Signal for Functional Neuromuscular Stimulation – Part I: Autoregressive Modeling as a Means of EMG Signature Discrimination", IEEE Transactions on Biomedical Engineering, vol. 35, NO. 34, pp. 230-237, Apr. 1988.

[11] Steven M Kay. Modern Spectral Estimation: Theory and Application, Prentice Hall, 1988.

[12] Akaike, A New Look at the Statistical Model Identification, IEEE Trans. Autom. Control, Vol. AC19, pp.716-723, 1994

[13] Vapnik V. The nature of statistical learning theory, Springer Press, New York, 1999.