Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

Activity Recognition from acceleration data Based on Discrete Consine Transform and SVM

Zhenyu He, Lianwen Jin*
School of Electronic and Information Engineering,
South China University of Technology,
Guangzhou, China.510640
*E-mail: lianwen.jin@gmail.com

***Abstract*—This paper developed a high-accuracy human activity recognition system based on single tri-axis accelerometer for use in a naturalistic environment. This system exploits the discrete cosine transform (DCT), the Principal Component Analysis (PCA) and Support Vector Machine (SVM) for classification human different activity. First, the effective features are extracted from accelerometer data using DCT. Next, feature dimension is reduced by PCA in DCT domain. After implementing the PCA, the most invariant and discriminating information for recognition is maintained. As a consequence, Multi-class Support Vector Machines is adopted to distinguish different human activities. Experiment results show that the proposed system achieves the best accuracy is 97.51%, which is better than other approaches.**

***Keywords*—tri-axial accelerometer; activity recognition; Discrete Cosine Transform; Principal Component Analysis; SVM**

I. INTRODUCTION

Context awareness is a central issue in ubiquitous and wearable computing. 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 [1~6]. For example, several activities such as ambulation, typing, talking were distinguished in [2] with five small bi-axial accelerometers. In [3] and [4], 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 a crucial role during the recognition process. The choice of good features is a fundamental step in statistical pattern recognition and a highly problem dependent task. Generally speaking, most of the attempts to extract features from acceleration date can be classified into three categories, say, time domains features， frequency domains features and time-frequency analysis. Traditional widely used time domains features are mean [1~3], variance or standard deviation [1,2], energy [1,2,4], entropy [2], correlation between axes [1,2,4] and so on. The most popular frequency domains features are FFT coefficient [6]. Mäntyjärvi et al. [3] decompose the signal using wavelet transform and computer power of selected scales wavelet coefficients as features. Besides, the autoregressive model of time-series is presented to recognize human activity from a tri-axial accelerometer data [5].

In this paper, an efficient method for high-accuracy activity recognition based on the discrete cosine transform (DCT), the Principal Component Analysis (PCA) and Support Vector Machine (SVM) is presented. Discrete cosine transform [7] is an important orthogonal transform and its performance has been found to be asymptotically equivalent to the optimal Karhunent-Loeve transform for signal decorrelation. Recently, DCT has been successfully used in wide range of applications such as digital signal processing community, image compression and high-speed face recognition [9,10].In order to reduce the dimension of DCT features, the PCA is also employed in the DCT domain to extract the most discriminating features for recognition. With the combination of DCT and PCA, more DCT coefficients can be kept and the most discriminating features can be extracted. After that we adopt multi-class support vector machines to distinguish the different activities of human. The classification of four human daily activities shows encouraging results. The block diagram of our proposed activity recognition system is shown in Fig. 1.

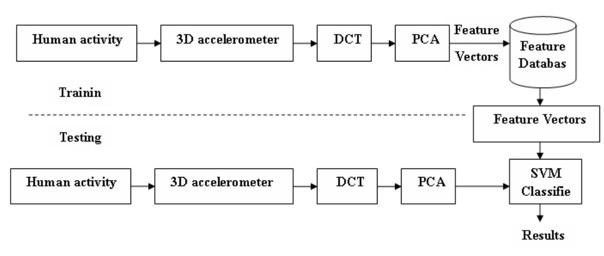


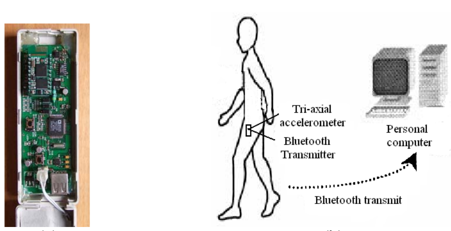
Figure 1. Overview of our system

978-1-4244-2794-9/09/$25.00 ©2009 IEEE

The rest of this paper is organized as follows: Data collected platform based on tri-axial accelerometer introduce in Section II. Section III presents the DCT-based feature extraction method and the PCA method to reduce features dimension. Section III describes the multi-class Support Vector Machine we used for classification. Experimental results and discussion are presented in Section IV. Finally, conclusions are drawn in Section V.

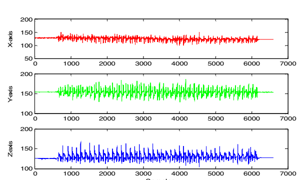
II. DATA COLLECTION

The diagram of our experimental setup is shown in Fig.2. 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 daily 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 3 shows the example of raw data and the corresponding colors (X-red, Y-green, Z-blue).

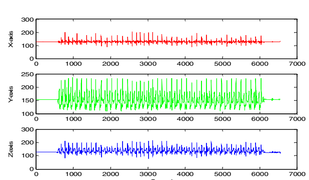


(a) (b)

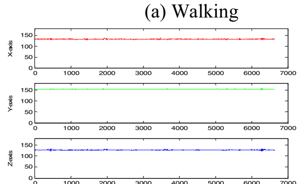
Figure 2. Data collection apparatus (a), Diagram of experimental setup (b)



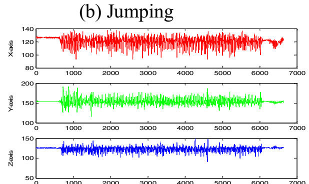
(a) Walking



(b) Jumping



(c) Still



(d) Running

Figure 3. Examples of raw signals for different activities

III. FEATURE EXTRACTION

A. *DISCRETE COSINE TRANSFORM*

Ahmed, Natarajan, and Rao [7] first introduced the discrete cosine transform (DCT) in the early seventies. Ever since, the DCT has grown in popularity, and several variants have been proposed. In particular, the DCT was categorized by Wang [8] into four slightly different transformations named DCT-I, DCT-II, DCT-III, and DCT-IV. Of the four classes Wang defined, DCT-II was the one first suggested by Ahmed et al., and it is the one of concern in this paper. The DCT of a data sequence $x(n), n = 0,1,\cdots,(N-1)$ is defined as:

$$X_c(0) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x(n) \tag{1}$$

$$X_c(k) = \sqrt{\frac{2}{N}} \sum_{n=0}^{N-1} x(n) \cos\frac{(2n+1)k\pi}{2N}, k = 1,2,\cdots,(N-1) \tag{2}$$

Where $X_c(k)$ is the kth DCT coefficient. All N DCT coefficients can be computed using a 2N-point fast Fourier transform. It can be shown that $X_c(k)$ is a bandpass filter with a center frequency at $(2k+1)/2N$ when the sampling frequency is normalized to 1. Hence, the magnitude of the output of $X_c(k)$ for small k is generally larger. In other words, the DCT can be concentrated in the low indices of the DCT if the remaining DCT coefficients can be set to zero without a significant impact on the energy of the signal. DCT is widely used in image compression also because of its excellent energy compaction property. As shown in Fig. 4, lots of frequency component of our activity acceleration data are centralized at the low-frequency. Most of the visually significant information is concentrated in just a few DCT coefficients. Therefore, we discard the high-frequency DCT coefficients, and select the low-frequency DCT coefficients as activity features. In this paper, we extract the first 48 magnitude of DCT coefficients from each axis acceleration data for features.

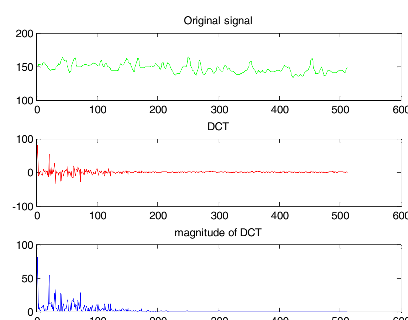


Figure 4. Y-axis data and its DCT

B. Principal Component Analysis

Feature reduction techniques attempt to reduce dimensionality by discarding some of the original features, whereas feature transform methods attempt to map the original features into a lower dimensional subspace. In this work, the principal component analysis (PCA) [3] is applied to reduce dimension of features. It was expected that they would reveal essential information in acceleration signals that describe human activity.

Let $\boldsymbol{X} = [x_1, x_2, \cdots x_n]^T$ be an *n*-dimensional random vector having zero mean. The task is to find an orthonormal matrix $\boldsymbol{V}$ of size $n \times k$, $k \leq n$, so that the reduced *k*-dimensional projection $X' = VX$ retains as much of the variance of X as possible. The matrix V defines the principal direction of the projection. In practice, the principal directions and components can be calculated using the eigendecomposition $C = E\Lambda E^T$ of the sample covariance matrix $C = E(XX^T)$ The eigenvalues $\Lambda = diag(\lambda_1, \lambda_2, \cdots, \lambda_n)$ determine the variance that each PC captures. The *n* PCs $X' = [x_1', x_2', \cdots x'_n]$ are computed by projecting the original data to the principal directions $X' = E^T X$.

IV. Classification

The classification algorithm we used is Support Vector Machine (SVM) [11]. 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. 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.

Since the training dataset is relative limited here, generalization capability of the classifier is more important for recognition. We used the leave-one-subject-out validation test [1] 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. As we known, the characteristics of the gait signals is unique for every person, therefore generalization capability based on leave-one-subject-out test is better than other test method

V. Experiment Results and Discussion

This section describes experiments with the developed activity recognition system. 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 [2]. 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 extract the first N magnitude of DCT coefficients from each axis acceleration data for features. Because our data included a large DC offset that represents the gravitational acceleration component of signal we threw away the DC component, leaving a $N-1$ dimensional each axis and giving a total of $(N-1)\times 3$ dimension features.

A. Number of DCT Coefficients

In order to determine how many DCT coefficients should be chosen, we test the recognition performance with different number of DCT coefficients. Experimental results are summarized in Table I. We observe that there is a slight decrease in recognition accuracy as we go to higher numbers of coefficients. In other words, more DCT coefficients do not necessarily mean better recognition performance because high-frequency components are related to unstable activity features such as noise. There will be more variable information for recognition when the DCT coefficients increase. According to Table I, the best performance of our system is obtained when *N*=48.

TABLE I. RECOGNITION PERFORMANCE VERSUS NUMBER OF DCT COEFFICIENTS

N	16	24	32	48	64
Accuracy	91.55	93.47	96.09	**97.16**	96.87

B. Number of PCA Component

As discuss above, the optimal number of the DCT coefficients is 48 and the total dimension of features is 141. In order to reduce the feature dimension of feature and obtain the most salient feature of human activity, the PCA is applied in the truncated DCT domain. The recognition performances versus different principal components are show in Table II. It can be seen that only the first 20 components are enough to obtained high recognition accuracy. The best recognition rate for our new algorithm is 97. 51 %. It should be highlighted that more components do not imply better recognition results, because by adding them, they may be representing more irrelevant information which is bad for recognition.

TABLE II. RECOGNITION PERFORMANCE VERSUS NUMBER OF PCA COMPOENT

NO.	10	15	20	25	30
Accuracy	96.09	97.37	**97.51**	97.51	97.44

C. Comparisons With Other Approaches

In order to compare the performance of our new features against other traditional features, we carry out several experiments under same experimental conditions. For example, the following kinds of traditional time-domains features were extracted from each axes of accelerometer: mean, standard deviation, energy and correlation between axes [4]. In [5]，Four orders of autoregressive model accelerometer data is built and the AR coefficients are chosen as features. Besides, Mäntyjärvi et al. [3] transform the signals using db8 and extract the power of wavelet coefficients from level 5 to 8. In [6], the FFT is calculated for signals from forward-axis and up-axis acceleration signals. The first 40 FFT coefficients per channel are concatenated as a feature vector. Comparative results of different approaches are shown in Table III. It is evident from the table III that accuracy using we proposed DCT+PCA feature is higher than other 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.3).

TABLE III. RECOGNITION PERFORMANCE COMPARISON OF DIFFERENT APPROACHES

Method	run	still	jump	walk	average
[4]	58.23	100	75.85	100	83.52
[5]	69.03	100	81.53	96.87	86.86
[3]	80.11	100	79.82	100	89.98
[6]	90.06	86.36	89.77	92.61	89.70
DCT	94.32	100	96.30	99.43	**97.51**

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 IV shows the aggregate confusion matrix for the SVM classifier 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 IV. CONFUSION MATRIX FOR THE PROPOSED METHOD

	run	still	jump	walk
run	332	0	**7**	13
still	0	352	0	0
jump	**13**	0	339	0
walk	2	0	0	350

VI. CONCLUSION

A high-accuracy human activity recognition system based on a tri-axial acceleration signals have been proposed in this paper. An efficient feature extraction method for high-accuracy activity recognition based on the discrete cosine transform and Principal Component Analysis is presented. To recognize different human activities, we adopt multi-class support vector machines in our system. Our experimental result demonstrates 97.51% accuracy for four different daily activities. We believe that the promising results will potentially contribute to the user's awareness of his daily activity level and promote a more active lifestyle.

ACKNOWLEDGMENT

This paper is partially supported by the following research funding: Motorola Foundation (B06-D8070740), NSFC (No. U0735004,60772216). The authors would like to thank the date collecting volunteers from SCUT.

REFERENCES

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

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

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

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

[5] He, Zhen-Yu; Jin, Lian-Wen. Activity recognition from acceleration data using AR model representation and SVM, IEEE International Conference on Machine Learning and Cybernetics,vol.4,pp. 2245-2250, July,2008

[6] J. Mantyjarvi, M. Lindholm, E. Vildjiounaite, S.-M.,Makela, H.A. Ailisto, "Identifying users of portable devices from gait pattern with accelerometers", IEEE International Conference on coustics, Speech, and Signal Processing, vol.2,PP: 973-976,2005.

[7] N. Ahmed, T. Natarajan, and K. R. Rao. "Discrete Cosine Transform", IEEE Trans. on Computers, vol. 23, pp.90-93,Jan 1974.

[8] Z. Wang, "Fast algorithms for the discrete W transform and for the discrete fourier transform". IEEE Trans. Acoust., Speech, and Signal Proc., 32:803–816,1984

[9] Meng Joo Er, W. Chen, and Wu. Shiqian, "High-Speed Face Recognition Based on Discrete Cosine Transform and RBF Neural Networks", IEEE Trans. on Neural Networks, vol.16, No.3, pp:679-691,may,2005.

[10] Z. M. Hafed and M. D. Levine "Face Recognition Using the Discrete Cosine Transform.International", Journal of Computer Vision, 43(3),pp.167-188,2001

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

.