PCA & HMM Based Arm Gesture Recognition Using Inertial Measurement Unit

Yinlong Zhang
Key Laboratory of Networked Control Systems, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, 110016, China
University of Chinese Academy of Sciences, Beijing, 100049, China
zhangyinlong@sia.cn

Wei Liang
Key Laboratory of Networked Control Systems, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, 110016, China
weiliang@sia.cn

Jindong Tan
Department of Mechanical Aerospace and Biomedical Engineering
The University of Tennessee
Knoxville, Tennessee, U.S.
tan@utk.edu

Yang Li
Key Laboratory of Networked Control Systems, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang, 110016, China
University of Chinese Academy of Sciences, Beijing, 100049, China
yangli@sia.cn

Ziming Zeng
Information and Control Engineering Faculty,
Shenyang Jianzhu University
Shenyang, 110168, China
Zengziming1983@gmail.com

ABSTRACT

This paper presents a novel arm gesture recognition approach that is capable of recognizing seven commonly used sequential arm gestures based upon the outputs from Inertial Measurement Unit (IMU) integrated with 3-D accelerometer and 3-D gyroscope. Unlike the traditional gesture recognition methods where the states in the gesture sequence are irrelevant, our proposed recognition system is intentionally designed to recognize the meaningful gesture sequence where each gesture state relates to the contiguous states which is applicable in the specific occasions such as the police directing the traffic and the arm-injured patients performing a set of arm gestures for effective rehabilitation. In the proposed arm gesture recognition system, the waveforms of the inertial outputs, i.e., 3-D accelerations and 3-D angular rates are automatically segmented for each arm gesture trace at first. Then we employ the Principal Component Analysis (PCA) - a computationally efficient feature selection method characteristic of compressing the inertial data and minimizing the influences of gesture variations. These selected features from PCA are compared with those standard features stored in pattern templates to acquire the gesture observation sequence that satisfy the Markov property. Finally, the Hidden Markov Model is applied in deducing the most likely arm gesture sequence. The experimental results show that our arm gesture classifier achieves up to 93% accuracy. By comparing with the other published recognition methods, our approach verifies the robustness and feasibility in arm gesture recognition using wearable MEMS sensors.

Keywords

Principal Component Analysis; Hidden Markov Model; Arm Gesture Recognition; Inertial Measurement Unit

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
BODYNETS 2013, September 30-October 02, Boston, United States
Copyright © 2013 ICST 978-1-936968-89-3
DOI 10.4108/icst.bodynets.2013.253667

1. INTRODUCTION

Gesture recognition is an important application and pertains to recognizing the meaningful expressions of motion by a human, involving the hands, arm, face and body [1]. As the proliferation of low-cost, light-weight and compact MEMS has motivated the researches on wearable computing, multiple gesture recognition techniques using wearable devices like Apple Iphone, Wiimote controller, IMU, contribute to its pervasive use in our daily life.

Among the human gesture recognition categories, the arm gesture recognition plays an indispensable role. The arm gestures (especially the semantic arm gestures) inherently include an underlying characteristic that the contiguous gesture states are largely relevant, i.e., the current gesture state is dependent on the previous state proportionally which satisfies the Markov property. Yet, quite a few researchers on gesture recognition fail to take that aspect into consideration. Apart from that, arm gesture recognition could be applied to a variety of occasions. For example, it would be helpful for traffic police mounting the wearable MEMS sensors for recognizing the meaningful arm gestures in directing the traffic [2] as well as for the wheelchair-propulsion upper-limb kinetic researches [3]. Another example of arm gesture recognition is for arm-injured patients in rehabilitation, which will obviously shorten their recovery time. However, the researches on arm gesture recognition are still limited, and there is still potential for exploration.

There are two major challenges in developing an arm gesture recognition system. The fact that the collected IMU traces from the same gesture are discrepant and it causes the ambiguity in gesture trace comparison with the stored gesture template, poses the first major challenge. The second challenge in arm gesture recognition, however, is the effectiveness and robustness of the recognition algorithm. To be specific, how to develop a universally applicable algorithm which is computationally efficient and suitable for most users is still imperfectly resolved.

In gesture segmentation, Xu Zhang et al. [4] combine the accelerometer signals and EMG signals to automatically distinguish the valid gestures. But the EMG signals are easily affected by the soft tissue artifacts which contribute to the false gesture segmentation. Thomas et al. [5] use Wiimote buttons to determine the gesture trace. This method is efficient, yet requires the user to manually press the controller buttons, which is not suitable for daily activity recognition. In the stage of feature extraction, traditional data analysis methods such as FFT, wavelet transform [6] in frequency domain and the familiar statistical parameters such as max, mean, variance are usually taken as the inherent features. However, these feature selection methods are either improper for the large sum of data like the inertial output waveforms or unable to sufficiently represent the collected data features. By comparison, (Principal Component Analysis) PCA and (Linear Discriminant Analysis) LDA [7] are well known for their elegant dimensionality reduction and computationally efficient property in feature extraction. In classification, k-means and k nearest neighbor (KNN) [8] are usually applied due to their simplicity in practice. However, these approaches fail to take the intrinsic contiguous states existed in data waveforms into account. Conversely, HMM and Dynamic time warping [9] are useful and robust in solving time-varying co-related signals.

With respect to the above problems, we propose an arm gesture recognition system that characterizes the computationally efficient feature selection and inherently co-related gesture identification. To be specific, "scale invariant" features corresponding to each gesture trace are selected using PCA. And we apply the HMM to deduce the arm gesture sequence as the arm movements satisfy the Markov property that the current gesture state is merely associated with that of the previous gesture and observation outputs are dependent on arm gesture states with a probability distribution.

2. Methodology

The proposed method contains three steps. In the first step, we set three threshold pairs to segment the gesture traces. In the second step, a computationally efficient method: Principal Component Analysis is applied to reduce the dataset dimensionality and extract the inertial features inherently existed in IMU waveforms. Then the arm gesture observation sequence is derived by comparing the gesture features with the standard template pattern features. Lastly, we take advantage of HMM to deduce the hidden arm gesture state sequence. Figure1. shows the overview of our arm gesture recognition approach.

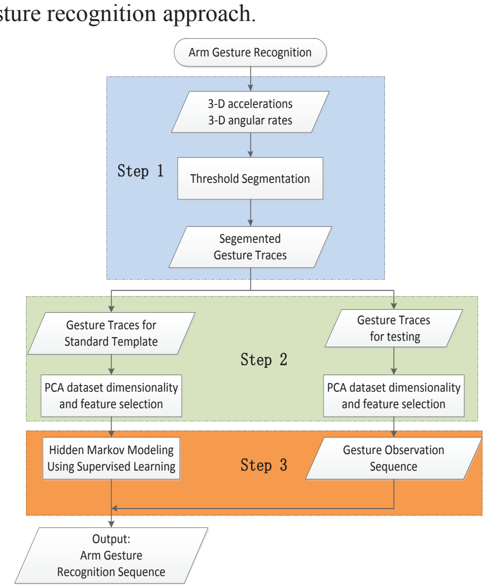

Figure1.

General overview of the arm gesture recognition approach

2.1 Inertial data segmentation

The outputs from the 3-D accelerometer fixed on forearm include the instantaneous meaningful gestures. However, the acceleration waveforms corresponding to the same movement conducted by the specific subject can differ drastically with time, not to mention the diverse accelerometer outputs related to the same movement conducted by different subjects. Thus, it's challenging to select a general window size in segmenting the 3-D accelerometer waveforms in time domain. It would be helpful in designing a general system that enables the automatic segmentation for the acceleration waveforms.

In our recognition approach, the participant is asked to perform the sequential arm gestures with short time intervals during which arm is being extended forward parallel to the ground such that the waveforms from the 3-D accelerometer vary subtly during the short time and the variance of the samples from the three-axis accelerations maintains relatively low. Afterwards, we adopt a sliding window to determine the starting point and ending point for each gesture. As the sliding window moves ahead from the starting points, we deduce the triple acceleration variances ($\mathrm{Var}(A_x)$, $\mathrm{Var}(A_y)$, $\mathrm{Var}(A_z)$) within the window which are then compared with the triple threshold pairs: (Δ_{Axs}, Δ_{Axe}) (Δ_{Ays}, Δ_{Aye}) (Δ_{Azs}, Δ_{Aze})to check whether the selected window represents the gesture starting point or ending point. We set the rules as follows:

Gesture starting point conditions:

$$\mathrm{Var}(A_x) > \Delta_{\mathrm{Axs}} \text{ or } \mathrm{Var}(A_y) > \Delta_{\mathrm{Ays}} \text{ or } \mathrm{Var}(A_z) > \Delta_{\mathrm{Azs}}$$

Gesture ending point conditions:

$$\mathrm{Var}(A_x) < \Delta_{\mathrm{Axe}} \text{ and } \mathrm{Var}(A_y) < \Delta_{\mathrm{Aye}} \text{ and } \mathrm{Var}(A_z) < \Delta_{\mathrm{Aze}}$$

The workflow of the gesture trace segmentation is shown in Figure2

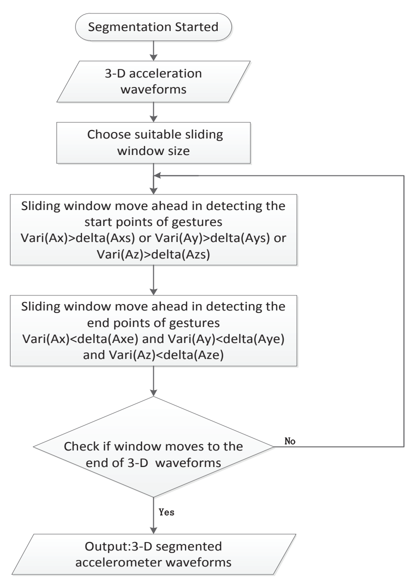

Figure2. Workflow of gesture trace segmentation

2.2 PCA for feature selection

Recall that the traditional Euclidean distance directly on the gesture traces is not applicable as a similarity measure for arm gesture traces suffer from inherent temporal variations. In our gesture recognition systems, we resort to PCA to reduce the dataset dimensionality and compute the waveform similarities by comparing the gesture eigenvalues and eigenvectors.

Jollifle [7] proved that PCA determines the direction along which the variability of the data is maximal. This orthogonal linear transformation method maps the original data into the eigenspace which characterizes the transformed data uncorrelated with each other. We apply PCA on the set of matrix I = $(\overrightarrow{A_x}, \overrightarrow{A_y}, \overrightarrow{A_z}, \overrightarrow{G_x}, \overrightarrow{G_y}, \overrightarrow{G_Z})$, where $(\overrightarrow{A_x}, \overrightarrow{A_y}, \overrightarrow{A_z})$ and $(\overrightarrow{G_x}, \overrightarrow{G_y}, \overrightarrow{G_Z})$ symbolize the outputs from triaxial accelerometer and triaxial gyroscope respectively. As covariance among variables only makes sense if they are measured in the same units, we standardize each variable to zero mean and unit variance before PCA feature selection. After subtracting the means from each vector and conducting nondimensionalizing, we acquire the matrix $\bar{I}$. Then we obtain the six by six covariance matrix as shown in formula 1.

$$C_{AG} = \bar{I}^T \bar{I} \tag{1}$$

The eigenvalues λ_i and the associated eigenvector $\vec{v_i}$, are computed by

$$C_{AG}\vec{v_i} = \lambda_i \vec{v_i} (\mathrm{i} = 1,2,\ldots,6) \tag{2}$$

Before conducting the PCA for feature extraction, we need to obtain the eigenvalues and eigenvectors for each gesture template. There are seven commonly used distinct gestures conducted in our experiments, namely, arm waving up and down, arm waving left and right, forearm flexion and extension, pronation and supination, ticking, crossing and circling. Each gesture is associated with six eigenvectors $\mu_{k,i,(k=1\ldots7,\ i=1,\ldots6)}$ and the corresponding eigenvalues $\overrightarrow{\omega}_{k,i,(k=1\ldots7,\ i=1,\ldots6)}$ after principal component analysis. These seven gestures are performed before gesture feature matching and each gesture features are stored in the template pattern.

After threshold segmentation, we obtain the starting point and ending point for each gesture. Then the waveforms are transformed into the segmented sample matrix $C_{A,G,j} = (\overrightarrow{A_{xJ}}, \overrightarrow{A_{yJ}}, \overrightarrow{A_{zJ}}, \overrightarrow{G_{xJ}}, \overrightarrow{G_{yJ}}, \overrightarrow{G_{ZJ}})$, where $\mathrm{j}(1 \le \mathrm{j} \le \mathrm{m})$ means the jth conducted gesture; m represents the mth segmented gestures. After PCA, we obtain the corresponding eigenvalues and eigenvectors.

Eigenvalues: $(\lambda_{j1}, \lambda_{j2}, \lambda_{j3}, \lambda_{j4}, \lambda_{j5}, \lambda_{j6})$

Eigenvectors: $(\vec{v}_{j,1}, \vec{v}_{j,2}, \vec{v}_{j,3}, \vec{v}_{j,4}, \vec{v}_{j,5}, \vec{v}_{j,6})$

For matching the similarity between specified segmented gesture and the original template gesture, we convert it into comparing the difference between the eigenvectors and eigenvalues, using the formula below:

$$D_{j,k} = \sum_{i=1}^{6} \left\| \vec{v}_{j,i} - \vec{\omega}_{k,i} \right\|_2 \left| \lambda_{j,i} - \mu_{k,i} \right| \ (1 \le \mathrm{j} \le \mathrm{m}, 1 \le \mathrm{k} \le 7) \tag{3}$$

$\vec{\omega}_{k,i}$ denotes the ith eigenvector associated with the kth template

$\vec{v}_{j,i}$ denotes the ith eigenvector associated with the jth segmented waveforms

$\mu_{k,i}$ denotes the ith eigenvalue associated with the kth template

$\lambda_{j,i}$ denotes the ith eigenvalue associated with the jth segmented waveforms

Assuming that $D_{j_0,k}$ is the minimum among the set $\{D_{j,k}\}$, then we label the jth inertial waveform segment as j_0 for the further HMM decoding.

2.3 Hmm recognition

HMM is a doubly stochastic process, in which the observable state sequence is generated from a hidden Markov state sequence [11]. In our arm gesture recognition system, there are 9 observation symbols (two observation symbols for arm waving Up and Down, two observation symbols for arm waving Left and Right, one for Pronation and Supination, one for forearm flexion and extension, one for ticking, one for crossing , one for circling). Thus, HMM is modeled as a triplet $\lambda = (\mathrm{A}_{7\times7}, \mathrm{B}_{7\times9}, \pi_7)$, where $\mathrm{A}_{7\times7}$ represents the State transition probability matrix; $\mathrm{B}_{7\times9}$ represents the observation probability matrix; π_7 represents the initial probability distribution. In our recognition system,

In HMM modeling, there are two classical methods, i.e., the Baum-Welch algorithm and supervised learning algorithm. Baum-Welch algorithm is a pervasively used unsupervised learning method but it suffers from the local maxima and is vulnerable to the ill-suited initial values. In contrast, the supervised learning method is independent of the initial estimation as well as independent of the local maximum. Thus we adopt the supervised learning to model the HMM.

The step in determining λ is as follows:

1. State transition probability matrix $\mathrm{A}=[\mathrm{a_{ij}}]_{7\times7}$

$$\mathrm{aij} = \frac{\mathrm{A_{ij}}}{\sum_{\mathrm{j}=1}^{\mathrm{N}} \mathrm{A_{ij}}} \tag{4}$$

where Aij is the number of times that the hidden state transits from state Si to Sj

2. Observation probability matrix $\mathrm{B}=[\mathrm{b_{ij}}]_{7\times9}$

$$\mathrm{bij} = \frac{\mathrm{B_{jk}}}{\sum_{\mathrm{k}=1}^{\mathrm{M}} \mathrm{B_{jk}}} \tag{5}$$

where $\mathrm{B_{jk}}$ is the number of times that the observation symbol is k and the hidden state is Sj

3. Initial probability distribution $\pi=[\pi_{\mathrm{i}}]_{7\times1}$

$$\pi_i = \Pr(\mathrm{q}_0 = S_i) \tag{6}$$

Given the HMM model λ and the gesture sequence after PCA feature selection, we take advantage of the Viterbi algorithm to find out the best hidden state sequence.

3. EXPERIMENTS

In our arm gesture recognition system, we design an IMU board that includes a module MPU6500 integrated with a MEMS three-axis accelerometer (full scale range: $\pm$ 16g), three-axis gyroscope (full scale range: ±2000°/sec) and three-axis magnetometer. An embedded processor MSP4300 and a Bluetooth module are also designed on this IMU board. The board size is $35\mathrm{mm} \times 30\mathrm{mm} \times 8\mathrm{mm}$. The sensor signals are sampled at 40HZ and interfaced to the computer by Bluetooth.

As is shown in Figure3, the collected inertial data corresponding to these seven gestures (i.e., arm waving arm up and down, arm waving left and right, pronation and supination, forearm flexion and extension, circling, ticking, crossing) is collected using IMU which is mounted on forearm. Each gesture trace is segmented using the triplet threshold pairs corresponding to x-, y- and z-acceleration $(\Delta_{\mathrm{Axs}}, \Delta_{\mathrm{Axe}})$, $(\Delta_{\mathrm{Ays}}, \Delta_{\mathrm{Aye}})$, $(\Delta_{\mathrm{Azs}}, \Delta_{\mathrm{Aze}})$. In the experiment, we set $\Delta_{\mathrm{Axs}} = \Delta_{\mathrm{Ays}} = \Delta_{\mathrm{Azs}} = 0.25$, $\Delta_{\mathrm{Axe}} = \Delta_{\mathrm{Aye}} = \Delta_{\mathrm{Aze}} = 0.15$, the sliding window size $\Delta = 25$. These parameters are chosen specifically for the user in reference to the experimental outcome. Seven template features are created and stored in the database that totally consists of 42 eigenvectors and 42 eigenvalues for the typical seven arm gestures.

In our experiment, the subject performed nine groups of arm movements and each group contains 130-160 gestures. We split the data into two parts. The former three are for the standard template features and Hidden Markov Modeling, the other six are for gesture testing. After PCA method, the features in the testing part for each gesture are compared with the template gestures. Then each gesture will be labeled with a specified number which represents the template with the highest similarity.

Figure3. Arm Gesture Sequence & IMU Board

The experimental results of our arm gesture recognition approach are listed in Table I. As is shown in the table, the recognition accuracy obviously gets better with the model size mounts up. The average recognition accuracy among the best three testing gesture models achieves up to 91%.

Sometimes, due to the unintended arm motions or the environmental vibrations, the system may recognize a gesture which is not an intended gesture input. These fake gestures are not taken into account as they never appear in the recognition process.

Table 1

Gesture Recognition Accuracy of Best Three Models

	Model Size	Average Accuracy
Model 1	130	88%
Model 2	150	92%
Model 3	160	93%

Table 2

Comparison of different Gesture Recognition Accuracy

Algorithm	No. of Gestures	Average Accuracy (%)
Proposed Approach	7	91%
Discrete HMM [5]	5	89.7
DCT & Discrete Wavelet Transform [6]	7	79%

Comparing the proposed recognition approach with another two methods [5] [6], shown in Table 2, we can see that the recognition algorithm based on PCA feature extraction and HMM classification is accurate in recognizing the arm gesture outputs from MEMS 3-D accelerometer and 3-D gyroscope.

4. CONCLUSION

In this paper, we present an implementation of arm gesture recognition approach using IMU that has a three-axis accelerometer and three-axis gyroscope. Unlike the traditional gesture recognition methods, our recognition system is intentionally designed to recognize the meaningful gesture sequence where each gesture state relates to the contiguous states with a proportional distribution. This recognition system is low cost and energy efficient which can be applied to the occasions like the traffic officers directing the traffic, wheelchair-propulsion upper-limb kinetic researches and the arm-injured patients in rehabilitation. It's noticeable that the segmentation algorithm is the critical issue in improving the performance of arm gesture system, and the other effective and robust segmentation methods will be investigated in the future work.

5. ACKNOWLEDGEMENT

This work was supported by the National Science Foundation of China under contact (61174026, 61172145, 61202418), the National High Technology Research and Development Program of China (863 Program: 2011AA040101), Foundation of Chinese Academy of Sciences under contact (KGCX2-EW-104).

6. REFERENCES

[1] S. Mitra, T. Acharya, 2007,"Gesture Recognition: A Survey" IEEE Transactions on Systems, Man and Cybernetics. Vol. 37, No.3. pp. 311-324

[2] B. Wang, T. Yuan, 2010, "Accelerometer-Based Chinese Traffic Police Gesture Recognition System", Chinese Journal of Electronics, Vol. 19, No.2, pp.270-274 April

[3] M.B.Morrow, J. Hurd, R.Kaufman, 2009 "Upper-limb joint Kinetics expression during wheelchair propulsion". *Journal of Rehabilitation Research & Development*. Vol. 46, No.7. pp. 939-944

[4] X. Zhang, X. Chen, Yun Li, Vuokko Lantz, 2011, "A Framework for Hand Gesture Recognition Based on Accelerometer and EMG Sensors". *IEEE Transactions on Systems, Man and Cybernetics*. Vol.41, No.6 pp.1064-1076

[5] T. Scholmer, B. Poppinga, N. henze, S. Boll, 2008, "Gesture Recognition with a Wii Controller" *In Proceedings of the second international Conference on Tangible and Embedded Interaction*. Bonn Germany. pp. 11-14.

[6] S. Zhou, Q. Shan, F. Fei, 2009 "Gesture Recognition for Interactive Controllers Using MEMS Motion Sensors" *In Proceedings of the 2009 t4h IEEE International Conference on Nano/Micro Engineered and Molecular Systems,* Shenzhen, China. pp.935-940

[7] S. Calinon, A. Billard, 2005 "Recognition and Reproduction of Gestures using a Probabilistic Framework combining PCA, ICA, and HMM" *In Proceedings of the 22nd International Conference on Machine Learning*, Bonn Germany, pp. 105-112.

[8] T. Kanungo, D. M. Mount, 2002 "An Efficient k-Means Clustering Algorithm: Analysis and Implementation" *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol.24, No.7, pp. 881-892

[9] Allen, Y.Yang, S.Iyengar, 2008, "Distributed Segmentation and Classification of Human Actions Using a Wearable Motion Sensor Network" *in Proceedings of Comput. Soc. Conf. Computer Vision and Pattern Recognition Workshops,* Jun.2008, pp.1--8

[10] R. Xu, S. Zhou, Wen J. Li, 2012,"MEMS Accelerometer Based Nonspecific-User Hand Gesture Recognition". *IEEE Sensors Journal*, Vol.12, No.5, pp.1166-1173.

[11] Lawrence, R. Rabiner, 1989 "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition" *IEEE Proceedings of the Spoken Language Process*, Vol.77, No.2, pp. 257-286.

[12] IMU: MPU 6500 Six-Axis (Gyro + Accelerometer) [Online]. http://www.invensense.com/mems/gyro/mpu6500.html[1]

Corresponding author: Wei Liang, email: weiliang@sia.cn