Proceedings of the 2008 IEEE
International Conference on Robotics and Biomimetics
Bangkok, Thailand, February 21 - 26, 2009

Towards HMM based Human Motion Recognition using MEMS Inertial Sensors

Guangyi Shi[1], Yuexian Zou[1*], Yufeng Jin[1], Xiaole Cui[1], and Wen J. Li[2]

[1]Advanced Digital Signal Processing Lab Shenzhen Graduate School of Peking University
[2]Centre for Micro and Nano Systems The Chinese University of Hong Kong

*Contact Author: zouyx@szpku.edu.cn

Abstract— This paper presents a new method of human motion recognition based on MEMS inertial sensors data. A Micro Inertial Measurement Unit (μIMU) that is 56mm*23mm*15mm in size was built. This unit consists of three dimensional MEMS accelerometers, gyroscopes, a Bluetooth module and a MCU (Micro Controller Unit), which can record and transfer inertial data to a computer through serial port wirelessly. Five categories of human motion were done including walking, running, going upstairs, fall and standing. Fourier analysis was used to extract the feature from the human motion data. The concentrated information was finally used to categorize the human motions through HMM (Hidden Markov Model) process. Experimental results show that for the given 5 human motions, correct recognition rate range from 90%-100%. Also, a full combination of 6 parameters (G_x, G_y, G_z, A_x, A_y, A_z) was listed and the recognition rate of each combination (total 63) was tested.

Index Terms— **MEMS, μIMU, Human Motion Recognition, HMM**

I. Introduction

Due to the availability of low-cost, small-size MEMS sensors, it is possible to build self-contained inertial sensors with an overall system dimension of less than 1 cubic inch and, at the same time, a sensing unit can track disorientation and other motions in real time [1]. With the aid of modern signal analysis and process method, the small-size, ubiquitous inertial measurement unit can be served for human motion recognition, which is widely studied for medical, artistic, or scientific purposes.

The world is facing with an increasingly aging population. With this increase, the proportion of the elderly who are frail and dependent is also likely to increase significantly [2]. This shift in demographic patterns will lead to an exponential increase in the number of elderly individuals who suffer injury from falls, as falls and fall-induced fractures are very common among the elderly. About 10% to 15% falls will cause serous injury in the elderly, and more than 1/3 of the persons aged over 65 will fall at least once per year [3, 4]. The early detection of fall is very important to rescue the subjects. Therefore, human motion recognition with wearable sensors, especially the recognition of falls, would be a hopeful solution for the elderly to avoid badly prognosis.

Some of the previous research into human fall detection make use of acceleration and tilt signals, set a few thresholds to these signals respectively, make decisions by detecting whether there is one or several data over the thresholds. Bourke A.K. etc. presents a threshold-based fall-detection algorithm using a tri-axial accelerometer [5]. Hwang et al [6] use tilt switch to trigger the detection program, when the tilt of the person's upper body over 70°, the program will start to process the acceleration signals to determine whether there is a fall occurred. These threshold methods are simple but not reliable. For instance, the rotations will affect accelerations which act as parameters of tilt. The bend of upper body could be more than 70° for normal motions, which will bring wrong trigger of fall detection.

Several research groups have reported the μIMU design and its applications. Tong Zhang et al developed a wearable sensor for fall detection with one-class SVM algorithm [7]. C.V. Bouten etc. reported a wearable accelerometric motion analysis system [8]. However, pure accelerometers will not result in accurate accelerations since the rotations affect accelerations a lot. A Micro Inertial Measurement Unit (μIMU) which consists of accelerometers and gyroscopes in 3 dimensions will cover these problems, since accelerometers used as inclinometers can measure the rotation about the vertical axis by the gyroscope. Figure 1 shows the working theory of a μIMU.

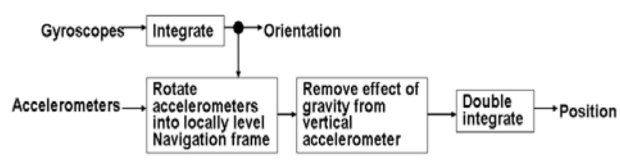

Fig. 1. The working principle of μIMU.

Our group has developed a Micro Input Device System (MIDS) based on MEMS sensors as a novel multi-functional interface input system that could potentially replace the mouse, pen, and keyboard as input devices to a computer [9]. They have also developed a μIMU that measures three-dimensional angular rates and accelerations based on MEMS sensors. We integrated a

978-1-4244-2679-9/08/$25.00 ©2008 IEEE

microcontroller and a Bluetooth module on the µIMU, and the overall size of the unit is less than 56 mm * 23 mm * 15 mm. We developed the module as a ubiquitous wireless digital writing instrument that interacts with humans and computers [10]. A human airbag system [1] is another application. Along with the µIMU, it will combine a SVM (Support Vector Machine) filter, an embedded DSP, and a mechanical system for airbag deployment.

SVM is feasible for binary classification. However, for multi-class classification, classifiers are complicated. Traditional approaches to solve time-varying problems are Dynamic Time Warping (DTW) and Hidden Markov Model (HMM). Signals scale along time or in amplitude (DDTW [11]) can be classified easily by DTW or its derivations, but rigorous templates are required. HMM also enjoys obvious advantages in recognition of time-varying sequence [12], such as speech recognition [13] and sign language recognition. In this paper, in order to distinguish different motions, FFT is used for feature extraction. The generated vectors are used to train HMM, and different motions can be categorized with little error by a sliding window method.

This paper is organized as follows. In Section II, the µIMU design and implementation is introduced. Section III will focus on concepts of FFT and HMM are introduced. Section IV shows the experimental results and conclusions are drawn in Section V.

II. MICRO IMU DESIGN AND HUMAN MOTION DATABASE

A. µIMU Design

MEMS sensors play a major role in the µIMU due to their low cost and miniature size. For our experiments, we use ADXL203 (AD Inc.) sensors as accelerometers and muRata ENC-03 angular rate gyros, respectively. They are low-cost and relatively high-performance sensors with analog signal output. The output signals of the accelerometers (A_x, A_y, A_z) and the angular rate gyros (ω_x, ω_y, ω_z) are converted directly by an A/D converter inside the microcontroller. We use an ATMEL ATmega32 microcontroller in the design. This microcontroller has 32 Kbyte flash, 2 Kbyte of SRAM, 8 channels of 10-bit ADC (here, we use 8-bit sampling rate), and a USART (Universal Synchronous and Asynchronous serial Receiver and Transmitter) port. The sampling rate of the microcontroller is 200 Hz, which ensures rapid reaction to human motion. We adopt a TDK Systems blu2i Module in our system to transfer data to a host system. This Bluetooth module provides easy integration to various host systems. The module is directly connected to the microcontroller via a USART port. The module is very small in size (45 mm * 23 mm * 15 mm) and can easily communicate with the microcontroller.

The accelerometers and the gyros act as a micro inertial measurement unit of the motion sensing system. These µIMU sensors and the Bluetooth module are housed in a small PCB, as shown in Fig.2. Two Li batteries of 3.6 V can power the unit for three hours. The computer receives 3 dimensional accelerations and angular rates, displays them in six plots in time domain (Fig.3), each represents one axis acceleration and rotation rate. At the same time, a txt file is generated consisting of six characters for later-on analysis. Thus, the µIMU can realize the function of data capture and transmission to the computer wirelessly.

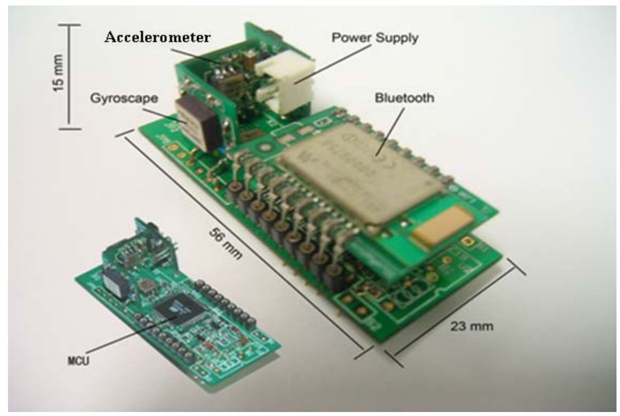

Fig. 2. Photograph of 3D motion sensing system consisting of 3 accelerometers and 3 gyros.

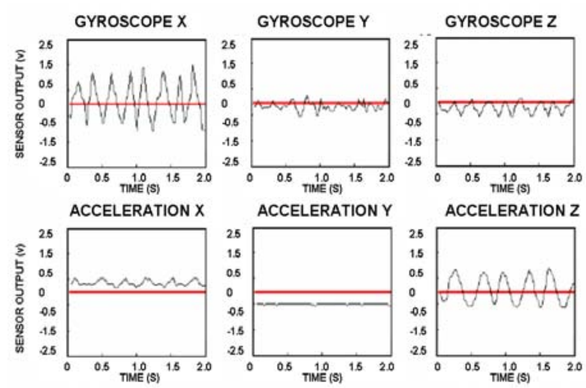

Fig. 3. Time domain accelerations and angular rate in 3 dimensional.

B. Human Motion Data Collection

Five groups of experiments were done. One hundred times of walking, stepping upstairs, lateral falling down, running and standing. Fig. 4 shows the original data of motions including 3D accelerations and angular velocity from experimental trials. Therefore, there are 6 elements in each vector from the µIMU. Elements 1, 2 and 3 are angular rates of X, Y and Z directions in the workspace (G_x, G_y and G_z). Axis of X is from back to front, axis Y is from left hip to right hip and axis Z is vertical to the ground. Elements 4, 5 and 6 are acceleration of X, Y and Z (A_x, A_y and A_z).Since our sampling rate is 200 Hz, number of X label in Fig. 4

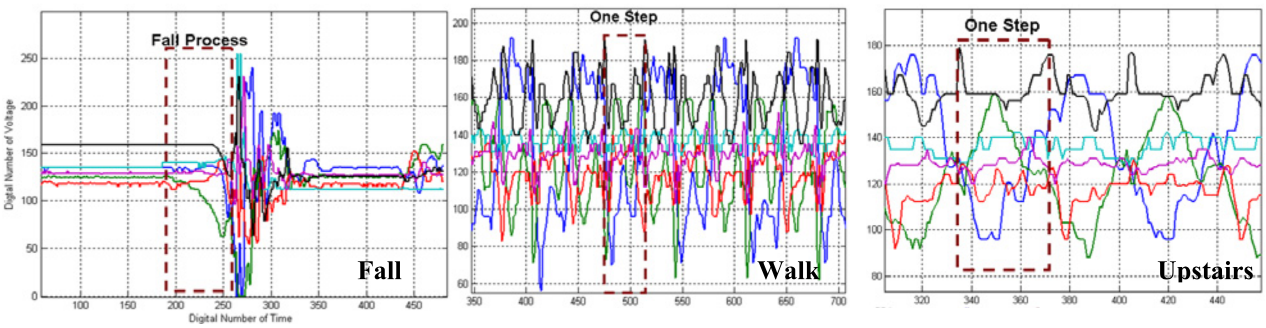

Fig. 4. Original motion data recording of fall, walking and stepping upstairs.

represents the time of 1/200 seconds. While the Y label in Fig. 4 is the sensor output voltage. 8-bit sampling rate was used, therefore, each number of Y label represent 1/256V.

In order to perform real-time detection, a sliding window method was used. Starting from the 32^{nd} frame, the previous 32 frames are detected, that is frame 1-32, 2-33, ..., n-32+n, ... We can also use a 64^{th} width window or 128^{th}. The width should be 2^n since we will use FFT for data generation later, 2^n input will improve FFT computation efficiency.

III. HMM Recognition Method

A. Feature Extraction

FFT is widely used for frequency domain processing and spectrum analysis. It is a computationally efficient discrete Fourier transform (DFT), which is defined as:

$$X(k)=\sum_{N=0}^{N-1} X_n W_N^{kn}, \quad k=0,\ldots,\mathrm{N-1} \tag{1}$$

where the twiddle factor is defined as:

$$\mathrm{W}_{\mathrm{N}}^{\mathrm{kn}} = \mathrm{e}^{-2\mathrm{j}\,\pi n k/N} \tag{2}$$

The reduced complexity of the radix-2 FFT algorithms is $(N/2)*log_2(N)$ complex multiplications and $(N)*log_2(N)$ complex additions. For each experimental result, we performed a FFT of 6 arrays respectively, and kept the first 10 coefficients of each FFT result. After 500 FFTs, we obtained 5 matrixes of 100 rows and 60 columns. Each row representing an experiment, which has 6*10 numbers in the sequence G_x, G_y, G_z, A_x, A_y, A_z. The feature extraction process can be seen in Fig. 5.

B. Hidden Markov Model

A Hidden Markov Model derives from Markov Chains, which is a doubly embedded stochastic process with an underlying stochastic process that is not observable, and can only be observed through another set of stochastic processes that produce the sequence of observations. An N-state $(S_1, S_2, \ldots S_N)$ of HMM can be described as $\lambda = (\pi, A, B)$. The parameters are as follows.

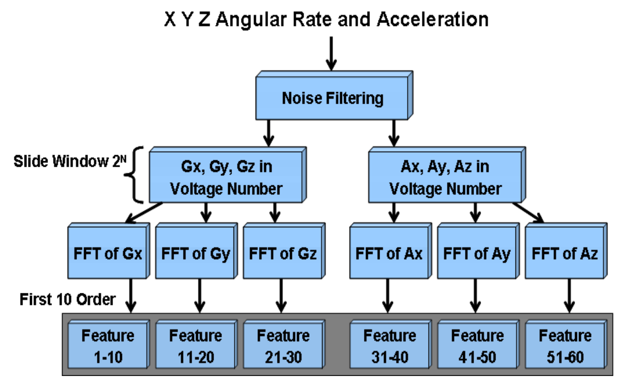

Fig.5. Feature extraction with FFT.

The initial state distribution,

$$\pi = \pi_1, \pi_2, \ldots \pi_N \tag{3}$$

which is used to describe the distribution of the state at time $t = 0$, and $\sum_{i=1}^{N} \pi_i = 1$.

The state transition probability distribution is:

$$A = \{a_{i,j}\}_{N \times N}, \quad i, j = 1, 2, \ldots N \tag{4}$$

$$a_{i,j} = P(q_j, t+1 \mid q_i, t), \quad 1 \le i, j \le N \tag{5}$$

$$\sum_{j=1}^{N} a_{ij} = 1 \tag{6}$$

In the special case where any state can reach any other in a single step, $a_{ij} > 0$ for any *i, j*. For other types of HMMs, $a_{ij} = 0$ for one or more *(i, j)* pairs.

The observation symbol probability distribution in state *j* can be described as $B = b_j(k)$, where

$$B = \{b_{jk}\}_{N \times N}, \quad k = 1, 2, \ldots, M \tag{7}$$

N denotes the hidden states, and *M* denotes the length of the codebook (discrete observation density). For continuous observation densities, it is $B = b_i(x)$.

Associating with the distribution of *B*, HMMs are categorized into Discrete Hidden Markov Model (DHMM with discrete probability density) and Continuous Hidden

Markov Model (CHMM, with continuous probability density).

Using these parameters, testing vectors can be interpreted properly. A testing vector is classified into the Model with the highest Maximum *A* Posteriori.

IV. Experiments and Analysis

A. Human Motion Modeling

We trained five 6-state left-right HMMs of the five human motions mentioned above. All HMM parameters including two model parameters (*M* and *N*) and three probability matrixes *A*, *B*, and π were initialized using a uniform segmentation for each training sequence. The Baum-Welch algorithm was used to iteratively reestimate the parameters according to the forward and backward variables. 40 iterations were run for each training process. Fig. 6 shows all 5 HMMs forward scores of the results for the Forward-Backward procedure, denoted as the log-likelihood versus the learning iteration index. Eventually the log-likelihood converges to a critical point. During the training process, the increase of log-likelihood indicated that the improvement of the model parameters. After the learning iterations, 5 HMMs representing the five human motions are retrieved based on the learning data samples.

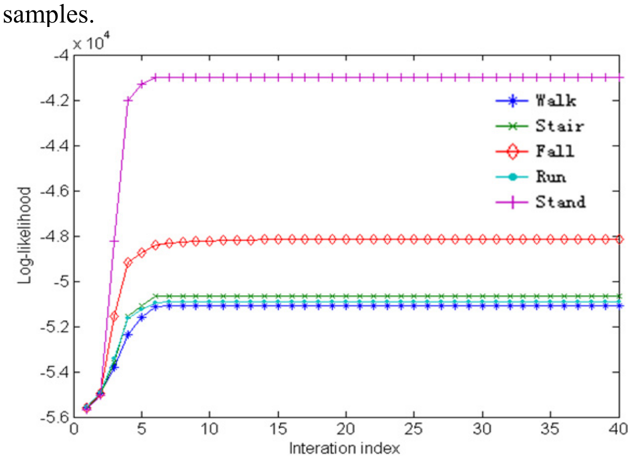

Fig.6. Log-likelihood versus learning iteration index in training process.

B. Model to Model Similarity Evaluation

In order to evaluate the 5 trained motion models (λ_1-λ_5), each model randomly generated 1 observation sequences which are as the same size as the ones for the training and testing targets. We applied those generated observation sequences O_i as the test data to all trained models λ_j (j = 1, 2, 3, 4, 5) one after another. The probability results $P(O_i | \lambda_j)$ are derived in the form of log-likelihood shown in table I. It is obviously to find that the log-likelihood reaches the maximum marked with shadow, when $i = j$. That is to say, each generated observation sequence has much more similarity with the model generating it than all the other models.

TABLE I.
Log-Likelihood of randomly generated sequences to the trained HMMs

Log-Likelihood	**walking**	**Up-stairs**	**fall**	**running**	**standing**
GS #1-1	-0.5120	-0.5135	-0.9944	-0.5179	-1.2643
GS #2-1	-0.5114	-0.5040	-0.5470	-0.5048	-1.2644
GS #3-1	-0.5116	-0.5020	-0.4796	-0.5105	-1.2643
GS #4-1	-0.5147	-0.5150	-0.7405	-0.5091	-1.2628
GS #5-1	-0.5211	-0.5069	-0.4923	-0.5180	-0.4045

C. Human Motion Recognition

As introduced before, we have extracted data of five motions, each have 100 samples. We divided the samples into two groups. One is for HMM training process, the other is for testing. Table Ⅱ shows the recognition result with different iterations (20, 40, 60), therefore, the correct rates of recognition range from 90%-100%.

TABLE Ⅱ.
Recognition result with different iterations

Iterations	**walking**	**Up-stairs**	**fall**	**running**	**standing**
20	48/50	45/50	49/50	50/50	50/50
40	48/50	48/50	49/50	50/50	50/50
60	50/50	50/50	49/50	50/50	50/50

V. Conclusion

In this paper, we present a novel MEMS-based human motion recognition method. A µIMU that consists of 3 dimensional accelerometers and 3 dimensional gyroscopes was developed and used for the human motion data collection. 5 human motions including walking, stepping upstairs, fall, running and standing were recoded, each have 100 samples. With the FFT process for the data extraction, the 500 samples were divided into two groups: training samples and testing samples. Through HMM training process, the 5 human motions can be recognized from each other with correct rates range from 90% to 100%.

References

[1] Guangyi Shi, Cheung Shing Chan, Yilun Luo, Guanglie Zhang, Wen J. Li, Philip H. W. Leong and Kwok-Sui Leung, "Development of a Human Airbag System for Falling Protection Using MEMS Motion Sensing Technology", IEEE International Conference on Intelligence Robots and Systems (IEEE IROS) 2006, Beijing, China.

[2] Myo Naing Nyan, Francis Eng Hock Tay, Teck Hong Koh, Yih Yiow Sitoh and Kwong Luck Tan, "Location and sensitivity comparison of MEMS accelerometers in signal identification for ambulatory monitoring", Electronic Components and Technology, June, 2004. Vol. 1, 1-4, Page(s):956-960.

[3] Noury N., "A Smart Sensor for the Remote Follow up of Activity and Fall Detection of the Elderly", Proceedings of 2nd Annual International IEEE-EMBS Special Topic Conference on

Microtechnologies in Medicine & Biology, Madison, Wisconsin (2002) 314-317.
[4] Luo S., Hu Q., "A Dynamic Motion Pattern Analysis Approach to Fall Detection", IEEE International Workshop on Biomedical Circuits & Systems, Singapore (2004) S2.1_5-S2.1_8.
[5] Bourke A. K., O'Brien J. V., Lyons G. M., "Evaluation of a Threshold-based Tri-axial Accelerometer Fall Detection Algorithm", *GAIT & POSTURE*, 26 (2): 194-199 JUL 2007.
[6] Hwang J.Y., Kang J.M., Jang Y.W., Kim H.C., "Development of Novel Algorithm and Real-time Monitoring Ambulatory System Using Bluetooth Module for Fall Detection in the Elderly", Proceedings of the 26th Annual International Conference of the IEEE EMBS, San Francisco (2004) 2204-2207.
[7] Zhang T., Wang J., Xu L., et al. "Fall Detection by Wearable Sensor and One-class SVM Algorithm", LECTURE NOTES IN CONTROL AND INFORMATION SCIENCES, 345: 858-863 2006.
[8] C.V. Bouten, K. T. Koekkoek, M.Verduin, R. Kodde and J. D. Janssen, "A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity", IEEE Trans. on Biomedical Engineering, Mar. 1997. Vol. 44, no. 3, Page(s):136-147.
[9] A.H.F. Lam, R.H.W. Lam, W.J. Li, M.Y.Y. Leung and Yunhui Liu, "Motion Sensing for Robot Hands using MIDS", ICRA '03. IEEE International Conference on Robotics and Automation, Sept. 2003. Vol. 3 ,14-19, Page(s):3181-3186.
[10] Guanglie Zhang, Guangyi Shi, Yilun Luo, H. Wong, Wen J. Li, P.H.W. Leong and Ming Yiu Wong, "Towards an ubiquitous wireless digital writing instrument using MEMS motion sensing technology", IEEE/ASME International Conference on Advanced Intelligent Mechatronics 2005, Page(s):795-800.
[11] Eamonn J. Keogh, Michael J. Pazzeni, "Derivative Dynamic Time Warping", First SIAM International Conference on Data Mining (SDM'01), 2001.
[12] Chard. O. Duda, *Pattern Classification, Second Edition,* John Wiley, 2003.
[13] Lawrence R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition", IEEE Proceedings of the Spoken Language Process, pp 257-286, Vol.77, No.2, 1989.