International Journal of
Scientific Engineering and Technology Research

ISSN 2319-8885
Vol.02,Issue.16,
November-2013,
Pages:1838-1843

www.semargroups.org,
www.ijsetr.com

MEMS Accelerometer based Non-Specific – User Hand Gesture Recognition

M.LAVANYA[1], P.SURESH BABU[2]

[1]PG Scholar, Dept of ECE, Mother Teresa Institute of Science & Technology, Sathupally, AP-INDIA,
E-mail: mandalapu.lavanya02@gmail.com.

[2]Prof, Dept of ECE, Mother Teresa Institute of Science & Technology, Sathupally, AP-INDIA,
E-mail: popurisb@gmail.com.

Abstract: This paper presents three different gesture recognition models which are capable of recognizing seven hand gestures, i.e., up, down, left, right, tick, circle and cross, based on the input signals from MEMS 3-axes accelerometers. The accelerations of a hand in motion in three perpendicular directions are detected by three accelerometers respectively and transmitted to a PC via Bluetooth wireless protocol. An automatic gesture segmentation algorithm is developed to identify individual gestures in a sequence. To compress data and to minimize the influence of variations resulted from gestures made by different users, a basic feature based on sign sequence of gesture acceleration is extracted. This method reduces hundreds of data values of a single gesture to a gesture code of 8 numbers. Finally the gesture is recognized by comparing the gesture code with the stored templates. Results based on 72 experiments, each containing a sequence of hand gestures (totaling 628 gestures), show that the best of the three models discussed in this paper achieves an overall recognition accuracy of 95.6%, with the correct recognition accuracy of each gesture ranging from 91% to 100%. We conclude that a recognition algorithm based on sign sequence and template matching as presented in this paper can be used for non-specific-users hand-gesture recognition without the time consuming user-training process prior to gesture recognition.

Keywords: Gesture recognition, Interactive controller, MEMS accelerometer..

I. INTRODUCTION

The increase in human-machine interactions in our daily lives has made user interface technology progressively more important. Physical gestures as intuitive expressions will greatly ease the interaction process and enable humans to more naturally command computers or machines. For example, in telerobotics, slave robots have been demonstrated to follow the master"s hand motions remotely [2]. Other proposed applications of recognizing hand gestures include character-recognition in 3D space using inertial sensors [3, 4], gesture recognition to control a television set remotely [5], enabling a hand as a 3D mouse [6], and using hand gestures as a control mechanism in virtual reality [7]. Moreover, gesture recognition has also been proposed to understand the actions of a musical conductor [8].

In our work, a miniature MEMS accelerometer based recognition system which can recognize seven hand gestures in 3D space is built. The system has potential uses such as a remote controller for visual and audio equipment, or as a control mechanism to command machines and intelligent systems in offices and factories. Many kinds of existing devices can capture gestures, such as a "Wiimote", joystick, trackball and touch tablet. Some of them can also be employed to provide input to a gesture recognizer. But sometimes, the technology employed for capturing gestures can be relatively expensive, such as a vision system or a data glove [9]. To strike a balance between accuracy of collected data and cost of devices, a Micro Inertial Measurement Unit (μIMU) is utilized in this project to detect the accelerations of hand motions in three dimensions.

There are mainly two existing types of gesture recognition methods, i. e., vision-based and accelerometer and/or gyroscope based. Due to the limitations such as unexpected ambient optical noise, slower dynamic response, and relatively large data collections/processing of vision-based method [10], our recognition system is implemented based on an inertial measurement unit based on MEMS acceleration sensors. Since heavy computation burden will be brought if gyroscopes are used for inertial measurement [11], our current system is based on MEMS accelerometers only and gyroscopes are not implemented for motion sensing. Existing gesture recognition approaches include template-matching [12], dictionary lookup, statistical matching, linguistic matching, and neural network. For sequential data such as measurement of time series and acoustic features at successive time

Copyright @ 2013 SEMAR GROUPS TECHNICAL SOCIETY. All rights reserved.

frames used for speech recognition, HMM (Hidden Markov Model) is one of the most important models [7]. It is effective for recognizing patterns with spatial and temporal variation [8]. In this paper, we present three different gesture recognition models, which are 1) sign sequence and Hopfield based gesture recognition model, 2) velocity increment based gesture recognition model, and 3) sign sequence and template matching based gesture recognition model. In these three models, in order to find a simple and efficient solution to the hand gesture recognition problem based on MEMS accelerometers, the acceleration patterns are not mapped into velocity, displacement or transformed into frequency domain, but are directly segmented and recognized in time domain. By extracting a simple feature based on sign sequence of acceleration, the recognition system achieves high accuracy and efficiency without the employment of HMM.

II. GESTURE RECOGNITION METHOD

There are mainly two existing types of gesture recognition methods, i.e., vision-based and accelerometer and/or gyroscope based. Due to some limitations like ambient optical noise, slower dynamic response, and relatively large data collections/processing of vision-based method [2], our recognition system is implemented based on an inertial measurement unit based on MEMS acceleration sensors. If gyroscopes are used for inertial measurement [3] it causes heavy computational burden, thus our system is based on MEMS accelerometers only and gyroscopes are not implemented. Many researchers have focused on developing effective algorithms for error compensation of inertial sensors to improve the recognition accuracy. For few examples, [4] proposed a pen-type input device to track trajectories in 3-D space by using accelerometers and gyroscopes. An efficient acceleration error compensation algorithm based on zero velocity compensation was developed to decrease the acceleration errors for acquiring accurate reconstructed trajectory.

An extended Kalman filter with magnetometers (micro inertial measurement unit (μIMU) with magnetometers), proposed by employed to compensate the orientation of the proposed digital writing instrument. If the orientation of the instrument was estimated precisely, the motion trajectories of the instrument were reconstructed accurately. Proposed an optical tracking calibration method based on optical tracking system (OTS) to calibrate 3-D accelerations, angular velocities, and space attitude of handwriting motions. The OTS was developed for the following two goals: 1) to obtain accelerations of the proposed ubiquitous digital writing instrument (UDWI) by calibrating 2-D trajectories and 2) to obtain the accurate attitude angles by using the multiple camera calibration. However, in order to recognize or reconstruct motion trajectories accurately, the aforementioned approaches introduce other sensors such as gyroscopes or magnetometers to obtain precise orientation. This increases additional cost for motion trajectory recognition systems as well as computational burden of their algorithms.

.

In this paper, a portable device has been developed with a trajectory recognition algorithm. The portable device consists of a triaxial accelerometer, a microprocessor, and a zigbee wireless transmission module. The acceleration signals measured from the triaxial accelerometer are transmitted to a computer via the zigbee wireless module. Users can utilize this portal device to write digits and make hand gestures at normal speed. The measured acceleration signals of these motions can be recognized by the trajectory recognition algorithm. The recognition procedure is composed of acceleration acquisition, signal preprocessing, feature generation, feature selection, and feature extraction. The acceleration signals of hand motions are measured by the portable device. The signal preprocessing procedure consists of calibration, a moving average filter, a high-pass filter, and normalization. First, the accelerations are calibrated to remove drift errors and offsets from the raw signals. These two filters are applied to remove high frequency noise and gravitational acceleration from the raw data, respectively.

The features of the preprocessed acceleration signals of each axis include mean, correlation among axes, interquartile range (IQR), mean absolute deviation (MAD), root mean square (rms), VAR, standard deviation (STD), and energy. Before classifying the hand motion trajectories, we perform the procedures of feature selection and extraction methods. In general, feature selection aims at selecting a subset of size m from an original set of d features (d > m). Therefore, the criterion of kernel-based class separability (KBCS) with best individual N (BIN) is to select significant features from the original features (i.e., to pick up some important features from d) and that of linear discriminate analysis (LDA) is to reduce the dimension of the feature space with a better recognition performance (i.e., to reduce the size of m). The objective of the feature selection and feature extraction methods is not only to eradicate the burden of computational load but also to increase the accuracy of classification.

The reduced features are used as the inputs of classifiers. The contributions of this paper include the following: 1) the development of a portable device with a trajectory recognition algorithm, i.e., with the hardware module , can give desired commands by hand motions to control electronics devices anywhere without space limitations, and 2) an effective trajectory recognition algorithm, i.e., the proposed algorithm can efficiently select significant features from the time and frequency domains of acceleration signals and project the feature space into a smaller feature dimension for motion recognition with high recognition accuracy.

III.HARDWARE DESIGN OF PORTABLE DEVICE

The portable device consists of a triaxial accelerometer (MMA2240), a microcontroller (C8051F206 with a 12-b A/D converter), and a wireless transceiver (nRF2401, Nordic). The triaxial accelerometer measures the acceleration signals generated by a user's hand motions. The microcontroller collects the analog acceleration signals and converts the signals to digital ones via the A/D converter. The wireless transceiver transmits the acceleration signals wirelessly to a personal computer (PC).The MMA2240 is a low-cost capacitive micro machined accelerometer with a temperature compensation function and a g-select function for a full-scale selection of +_}2 g to +_}6 gand is able to measure accelerations over the bandwidth of 0.5 kHz for all axes. The accelerometer's sensitivity is set from −2 g to +2 g. The C8051F206 integrates a high-performance 12-b A/D converter and an optimized signal cycle 25-MHz 8-b microcontroller unit (MCU) (8051 instruction set compatible) on a signal chip. The output signals of the accelerometer are sampled at 100 Hz by the 12-b A/D converter. Then, all the data sensed by the accelerometer are transmitted wirelessly to a PC by an zigbee transceiver at 2.4- GHz transmission band with 1-Mb/s transmission rate. The overall power consumption of the digital pen circuit is 30 mA at 3.7 V. The block diagram of the portable device is shown in Fig. 1.

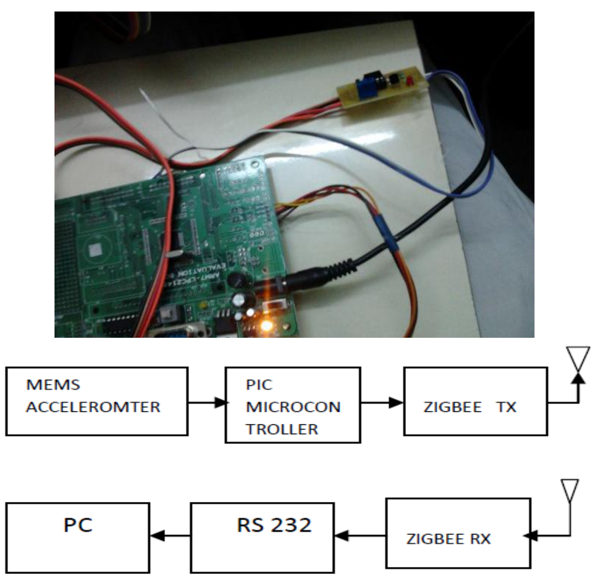

Fig.1. Block diagram of the portable device.

IV. TRAJECTORY RECOGNITION ALGORITHM

The proposed trajectory recognition algorithm consisting of acceleration acquisition, signal preprocessing, feature generation, feature selection, and feature extraction. In this paper, the motions for recognition include Arabic numerals alphabets. The acceleration signals of the hand motions are measured by a triaxial accelerometer and then preprocessed by filtering and normalization. Consequently, the features are extracted from the preprocessed data to represent the characteristics of different motion signals, and the feature selection process based on KBCS picks *p* features out of the original extracted features. To reduce the computational load and increase the recognition accuracy of the classifier, LDA is utilized to decrease the dimension of the selected features. The reduced feature vectors are then fed into a PNN classifier to recognize the motion to which the feature vector it belongs.

A. Signal Preprocessing

The microcontroller collects the acceleration signals of hand motions which are generated by the accelerometer. Due to slight tremble movement of hand certain amount of noise is generated. The signal preprocessing consists of calibration, a moving average filter, a high-pass filter, and normalization. First, the accelerations are calibrated to remove drift errors and offsets from the raw signals. The second step of the signal preprocessing is to use a moving average filter to reduce the high-frequency noise of the calibrated accelerations, and the filter is expressed as

$$y[t] = \frac{1}{N}\sum_{i=1}^{N-1} x[t+i] \tag{1}$$

Where $x[t]$ is the input signal, $y[t]$ is the output signal, and N is the number of points in the average filter. In this paper, we set $N = 8$. The decision of using an eight-point moving average filter is based on our empirical tests. Then, a high-pass filter is used to remove the gravitational acceleration from the filtered acceleration to obtain accelerations caused by hand movement. In general, the size of samples of each movement between fast and slow writers is different. Therefore, after filtering the data, we first segment each movement signal properly to extract the exact motion interval. Then, we normalize each segmented motion interval into equal sizes via interpolation.

B. Feature Generation

The characteristics of different hand movement signals can be obtained by extracting features from the preprocessed *x*-,

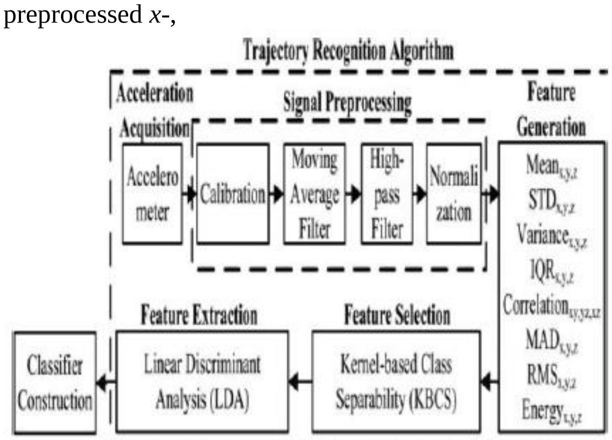

Fig.2. Block diagram of the trajectory recognition algorithm.

Y-, and *z*-axis signals, and we extract eight features from the triaxial acceleration signals, including mean, STD, VAR, IQR [7], correlation between axes [8], MAD, rms, and energy [9] . They are explicated as follows.

1) Mean: The mean value of the acceleration signals of each hand motion is the dc component of the signal

$$Mean = \frac{1}{|W|}\sum_{i=1}^{|W|} x_i \tag{2}$$

Where *W* is the length of each hand motion.

2) STD: STD is the square root of VAR

$$STD = \sqrt{\frac{1}{|W|-1}\sum_{i=1}^{|W|}(x_i - m)^2}. \tag{3}$$

3) VAR

$$VAR = \frac{1}{|W|-1}\sum_{i=1}^{|W|}(x_i - m)^2 \tag{4}$$

Where *xi* is the acceleration instance and *m* is the mean value of *xi* in (3) and (4).

4) IQR: When different classes have similar mean values, the interquartile range represents the dispersion of the data and eliminates the influence of outliers in the data.

5) Correlation among axes: The correlation among axes is computed as the ratio of the covariance to the product of the STD for each pair of axes. For example, the correlation (*corrxy*) between two variables *x* on *x*-axis and *y* on *y*-axis is defined as

$$corr = \frac{\mathrm{cov}(x,y)}{\sigma_x\sigma_y} = \frac{E\left((x-m_x)(y-m_y)\right)}{\sigma_x\sigma_y} \tag{5}$$

Where *E* represents the expected value, ***σx*** and ***σx*** are STDs, and *mx* and *my* are the expected values of *x* and *y*, respectively.

6) MAD

$$MAD = \frac{1}{|W|}\sum_{i=1}^{|W|}|x_i - m|. \tag{6}$$

7) rms

$$rms = \sqrt{\frac{1}{|W|}\sum_{i=1}^{|W|} x_i^2} \tag{7}$$

where *xi* is the acceleration instance and *m* is the mean value of *xi* in (6) to (7).

8) Energy: Energy is calculated as the sum of the magnitudes of squared discrete fast Fourier transform (FFT) components of the signal in a window. The equation is defined as

$$Energy = \frac{1}{|W|}\sum_{i=1}^{|W|}|F_i|^2 \tag{8}$$

Where *Fi* is the *i*th FFT component of the window and |*Fi*| is the magnitude of *Fi*.

C. Feature Selection

Feature selection comprises a selection criterion. The KBCS can be computed as follows: Let (**x**, *y*) (R*d* × ***Y***) represents a sample, where R*d* denotes a *d*-dimensional feature space, *Y* symbolizes the set of class labels, and the size of *Y* is the number of class *c*. This method projects the samples onto a kernel space, and **m** *i* is defined as the mean vector for the *I* th class in the kernel space, *ni* denotes the number of samples in the *i*th class, **m** denotes the mean vector for all classes in the kernel space, **S** *B* denotes the between-class scatter matrix in the kernel space, and **S**/ *W* denotes the within-class scatter matrix in the kernel space. Let (•) be a possible nonlinear mapping from the feature space R*d* to a kernel space *κ* and tr(**A**) represents the trace of a square matrix **A**.

The following two equations are used in the class separability measure:

$$\begin{aligned}\mathrm{tr}\left(\mathbf{S}_B^{\emptyset}\right) &= \mathrm{tr}\left[\sum_{i=1}^{c} n_i\left(\mathbf{m}_i^{\emptyset}-\mathbf{m}^{\emptyset}\right)\left(\mathbf{m}_i^{\emptyset}-\mathbf{m}^{\emptyset}\right)^T\right]\\ &= \sum_{i=1}^{c} n_i\left[\left(m_i^{\emptyset}-m^{\emptyset}\right)(m_i^{\emptyset}-m^{\emptyset})^T\right]\end{aligned} \tag{9}$$

$$\begin{aligned}\mathrm{tr}\left(\mathbf{S}_W^{\emptyset}\right) &= \mathrm{tr}\left[\sum_{i=1}^{c}\sum_{j=1}^{n_i}\left(\emptyset(\mathbf{x}_{ij})-\mathbf{m}_i^{\emptyset}\right)\left(\emptyset(\mathbf{x}_{ij})-\mathbf{m}_i^{\emptyset}\right)^T\right]\\ &= \sum_{i=1}^{c}\sum_{j=1}^{n_i}\left[\left(\emptyset(x_{ij})-m_i^{\emptyset}\right)^T\left(\emptyset(x_{ij})-m_i^{\emptyset}\right)\right].\end{aligned} \tag{10}$$

The class separability in the kernel space can be measured as

$$J^{\emptyset} = \frac{\mathrm{tr}\left(\mathbf{S}_B^{\emptyset}\right)}{\mathrm{tr}\left(\mathbf{S}_W^{\emptyset}\right)}. \tag{11}$$

To maintain the numerical stability in the maximization of $J^{\emptyset}$, the denominator tr($\mathbf{S}^{\emptyset}_W$) has to be prevented from approaching zero.

V. EXPERIMENTAL RESULTS

The experimental results of the three recognition models discussed above are listed in Table 1. As shown in the table, Model III (based on sign sequence and template matching) achieves the highest accuracy among the three models, while the performance of Model II is the worst of the three. Besides, Model II is not as robust as the other two methods, i.e., variations of the input gestures are more likely to affect the outcome of the gestures recognized, so this model should not be preferred when MEMS accelerometers are used for gesture recognition. Since Model I and Model III have similar gesture encoding the recognition mechanism, only the evaluation result of Model III is provided in more detail in this paper.

The test results shown in Table 1 are based on 72 test samples, totaling 628 single gestures, i.e., each test sample consists of a sequence of input gestures in a particular order. They are collected in two kinds of gesture sequences: 1) in the order of up – down – left – right – tick – circle – cross, and 2) 10 same single gestures in one sequence, e.g., circle – circle – circle – circle – circle – circle – circle – circle – circle – circle. To increase data diversity and simulate variations in gestures made by different persons, gestures were made in different speeds and intensities; the trajectories of some gesture motions were made with some variation, e.g., an ellipse was made instead of a circle. Model III has an overall mean accuracy of 95.6%, with the recognition accuracy of each gesture above 90%. Table II shows the detailed recognition results by using Model III, which shows the total number of input for each gestures and how many of the input are correctly recognized. We note here that, during experiments, some input gestures were not detected at all (i.e., due to loss of wireless transmission).

TABLE I
COMPARISON OF GESTURE RECOGNITION ACCURACY (%) OF THREE MODELS

	Up	Down	Left	Right	Tick	Circle	Cross	Mean
Model I	95.0	86.0	91.0	84.0	64.0	75.0	61.0	79.0
Model II	87.0	19.0	63.0	94.0	25.0	0.0	88.0	54.0
Model III	94.8	91.1	96.7.0	100	94.4	97.7	94.4	95.6

For example, if the order of input gestures in one experimental sample is up-down-left-right-tick-circle-cross, the detected gestures may only be down-left-right-tick-circle-cross, i.e., only the last six gestures were detected. Moreover, sometimes a „ghost" gesture may be detected, i.e., due to environmental vibrations or unintended hand motions, the algorithm may „recognize" a gesture even though there was no intended gesture input. These "missing" or "ghost" gestures were not taken into account when "recognition accuracy" is determined, because they did not go through recognition process at all. We note here that the recognition performance using Model III is higher than the performance obtained by our group"s prior work using HMM in [9]. A comparison of the results discusses in this paper and in [9] is provided in Table III. The experimental result proves that an algorithm based on sign sequence and template matching is efficient in recognizing gesture data from MEMS accelerometers without using a time consuming training process.

TABLE II
GESTURE RECOGNITION RESULTS FOR MODEL III

Experimental Results of Gestures		Recognized Gestures						
		Up	Down	Left	Right	Tick	Circle	Cross
Input Gestures	Up	91						
	Down	2	82			5		
	Left			87			1	
	Right			2	90			
	Tick	1				85	1	
	Circle						85	5
	Cross	2	8	1				85

TABLE III
COMPARISON OF GESTURE RECOGNITION ACCURACY (%) OF SEVERAL DIFFERENT ALGORITHMS

Algorithm	Average Accuracy (%)	No. of Testing Samples
DCT (Discrete Cosine Transform) [9]	94.6	370
DCT & Average Filtering [9]	87.6	370
DCT & Discrete Wavelet Transform [9]	90.5	370
Model I	79.0	628
Model II	54.0	628
Model III	95.6	628

VI. CONCLUSION

This paper describes a non-specific person gesture recognition system by using MEMS accelerometers. The recognition system consists of sensor data collection, segmentation and recognition. After receiving acceleration data from the sensing device, a segmentation algorithm is applied to determine the starting and end points of every input gesture automatically. The sign sequence of a gesture is extracted as the classifying feature, i.e., a gesture code. Finally, the gesture code is compared with the stored standard patterns to determine the most likely gesture. Since the standard gesture patterns are generated by motion analysis and are simple features represented by 8 numbers for each gesture, the recognition system does not require a big data base and needs not to collect as many gestures made by different people as possible to improve the recognition accuracy. We note here, however, to enhance the performance of the recognition system; we will need to improve the segmentation algorithm to increase its accuracy in finding the terminal points of gestures. Moreover, other features of the motion data may be utilized for pattern classification, i.e., more recognition methods will be investigated in our future work.

VII. REFERENCES

[1] Ruize Xu, Shengli Zhou, And Wen J. Li, Fellow, IEEE, "MEMS Acclerometer Based Non - Specific – User Hand Gesture Recognition", Sensors-5584-2011.

[2] T. H. Speeter, "Transformation human hand motion for telemanipulation," Presence, Vol. 1, No.1, pp.63-79, 1992.

[3] S. Zhou, Z. Dong, W. J. Li, and C. P. Kwong, "Hand-Written Character Recognition Using MEMS Motion Sensing Technology," in Proc. of the 2008 IEEE/ASME International Conference on Advanced Intelligent Mechatronics, pp. 1418 – 1423, 2008.

[4] J. K. Oh, S. J. Cho, W. C. Bang et al., "Inertial Sensor Based Recognition of 3-D Character Gestures with an Ensemble of Classifiers," in Proc. of the 9th Int'l Workshop on Frontiers in Handwriting Recognition, 2004.

[5] W. T. Freeman and C. D. Weissman, "TV Control by Hand Gestures," IEEE Intl. Wkshp. on Automatic Face and Gesture Recognition, Zurich, 1995.

[6] L. Bretzner and T. Lindeberg, "Relative Orientation from Extended Sequences of Sparse Point and Line Correspondences Using the Affine Trifocal Tensor," in Proc. of the 5th European Conference on Computer Vision, vol. 1406 of Lecture Notes in Computer Science, pp. 141 – 157, Springer Verlag, Berlin, June 1998.

[7] D. Xu, "A Neural Network Approach for Hand Gesture Recognition in Virtual Reality Driving Training System of SPG," 18th International Conference on Pattern Recognition, 2006.

[8] H. Je, J. Kim, and D. Kim, "Hand Gesture Recognition To Understand Musical Conducting Action," in Proc. of. IEEE International Conference on Robot & Human Interactive Communication, August 2007.

[9] T. Yang, Y. Xu, A "Hidden Markov Model for Gesture Recognition," CMU-RI-TR-94 10, Robotics Institute, Carnegie Mellon University, May 1994.

[10] S. Zhou, Q. Shan, F. Fei, W. J. Li, C. P. Kwong, C. K. Wu, et al, "Gesture Recognition for Interactive Controllers Using MEMS Motion Sensors," in Proc. IEEE International Conference on Nano/Micro Engineered and Molecular Systems, pp. 935-940, January 2009.

[11] S. Zhang, C. Yuan, and V. Zhang, "Handwritten Character Recognition Using Orientation Quantization Based on 3D Accelerometer," the 5th Annual International Conference on Ubiquitous Systems, July 25th, 2008.

[12] J. S. Lipscomb, "A trainable gesture recognizer," Pattern. Recognition, Vol. 24, No. 9, pp. 895-907, 1991.

Author's Profile:

1st Author:

Name: M.LAVANYA

Qualification: B.Tech., (M.TECH),

Email ID: mandalapu.lavanya02@gmail.com.

2nd Author: -

Name: Prof. P SURESH BABU
Qualification: M. Tech., (Ph. D),
Designation: Professor & HOD
Research Area: VLSI
Experience : 12 years
Industry : 2 years
Teaching : 14 Years
E-mail ID: popurisb@gmail.com.