Capturing Basic Movements for Mobile Platforms Embedded with Motion Sensors

Aravind Kailas[1]
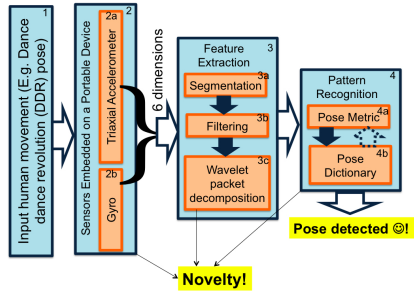
Fig. 1. Block diagram highlighting the main contributions of the paper.

Abstract— **A novel experimental setup using accelerometer and gyroscope sensors embedded on a single board along with a distance-based pattern recognition algorithm is presented for accurately identifying basic movements for possible application in gaming using a mobile platform. As an example, we considered some basic step sequences in the popular dance game (e.g., dance dance revolution), and could detect these movements with a reasonably high probability. We envision that the experimental results presented in this paper will motivate future research in the world of mobile gaming applications using advanced smart phones with a dual module design.**

I. Introduction: Evolution of Mobile Gaming

Mobile games are applications that simply use the device platform to run the game software. These games maybe downloaded on the fly and installed while mobile, or onto the handset using a cable, or come pre-installed by the original equipment manufacturer (OEM) or by the cellular operator as applications or services. More than 10 million people worldwide play games on mobile phones and handheld devices [1], and the world-wide mobile gaming revenue is expected to reach $11.4 billion by 2014. These are very important motivations for game developers and designers to create blockbuster games for mobile platforms.

The contributions of the paper can be summarized as follows:

1) Experimental Setup: A novel experimental setup using accelerometer and gyroscope sensors embedded on a single board (Block 2 in Fig. 1);
2) Data analysis and Feature Extraction: Our feature extraction using wavelet packet decomposition (WPD), thereby providing better resolution for analysis than the fast Fourier transform (FFT) (Block 3c in Fig. 1);
3) Algorithm Development: Pattern recognition-based algorithm based on the Mahalanobis distance, which is based on the correlations between the data points (Block 4a in Fig. 1);
4) Pose Recognition: Reliable pose recognition using two sensor boards (i.e., accelerometer and gyro combination) for games with a lot of human movements (e.g., dance dance revolution (DDR), a very popular game), is demonstrated with a high probability of success (Block 4 in Fig. 1) to motivate future research game changing phones with a dual module design such as the Fujitsu F-04B.

The rest of the paper is organized as follows. Section II briefly mentions relevant prior works. Section III outlines the methodology and experiments conducted. Experimental results are discussed in Section IV and conclusions are presented in Section V.

II. Prior Works

Combinations of accelerometers and gyroscope sensors have been used to (1) estimate human body orientation [2]; (2) detect human posture and walking speed [3], and (3) human recognition via gait recognition [4], to name a few. However, none of these implementations are for a portable device such as a mobile phone. There has been some work on developing systems that distinguish between activity and rest, measured the tilt angle of the body, and could detect walking and falling conditions, gait analysis, and other posture recognition algorithms for a mobile platform [5]–[7]. However, to the best of our knowledge, there have been no works that consider
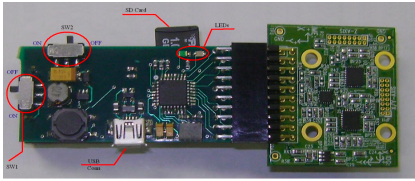
Fig. 2. The integrated sensor board mobile platform.

a combination of accelerometer and gyroscopes on a mobile device. To the best of our knowledge, there has been only one work that deploys accelerometers for video games such as the dance dance revolution (DDR). Crampton *et al.*, created and tested a wearable sensor network that detects the subjects body position as input for the video games [8]. The effects of multiple accelerometers on accurate pose detection support the use of accelerometer-based sensor networks as input in video games. This paper is different because, a combination of accelerometer and gyro sensors embedded on a portable dual mode mobile platform is considered for pose recognition resulting motivating future research in mobile gaming. Moreover, reasonable pose recognition is demonstrated using fewer sensors and our classification algorithm.

III. METHODOLOGY AND EXPERIMENTAL SETUP

A. Accelerometer-Gyro Board (Experimental Device)

The device we used for collecting data is shown in Fig. 2. The device integrates an IHU-500 board from InvenSense [1], an SD card with 2 GB capacity, and a USB port. The IHU-500 board is equipped with an integrated triaxis accelerometer (IME-3000), a two-dimension gyroscope (IDG-500), and a one-dimension gyroscope (ISZ-500). Therefore, the device can capture acceleration and rotation in three dimensions. In the futureThe sampling rate is set to be 500 Hz. Up to 8-hour sampling data with 16 bits/sample/dimension can be collected and stored in the in the 2 GB SD card. Up to 60 data sets could be recorded after which the data must be uploaded to a PC. Combined duration of the 60 datasets did not exceed eight hours of data recordings.

B. Pose Recognition Experiments

Three different experiments were conducted for pose recognition by using the DDR as an example, which depends on the placement of the accelerometer gyro on the body. It is vital to get the position and rotation of the "points of motion" to infer about the DDR step. In **Experiment 1**, every DDR song sequence can

be thought of as a combination of distinct sequences, each being a combination of forward, backward, right, and left movements. With this in mind, 48 poses were executed, each pose being a permutation of the four distinct movements without any repetition. A total of 100 readings are taken each time with a one-two second period between each movement. For this experiment, one sensor board was strapped tightly to the ankle, while the other was strapped around the torso (or the waist). In **Experiment 2**, one sensor board was strapped tightly to the ankle, while the other was positioned on the palm. A total of five poses were executed, and they were right, forward, backward, forward-backward, and backward-forward. The objective of this experiment was to distinguish the short movements or pose changes. Finally, in **Experiment 3**, one sensor board was strapped tightly to the ankle, while the other was strapped around the torso (or the waist). A total of five poses were executed, and they were right, forward, backward, forward-backward, and backward-forward. The objective here also was to distinguish the short movements or pose changes.

C. Training Phase

The purpose of the training phase is to create a raw data set of a given number of poses for an individual. In order to record a pose, after sufficient practice, the subject performs each pose 50 times. A reading consisted of the x, y, and z accelerations, and the pitch, roll, and yaw from the gyroscope from both the sensor boards. A data set for training comprised raw data from a single subject. In order to convert a set of raw data into recognizable pose pattern, two basic calculations must be performed. The mean acceleration for each pose must be computed. The acceleration values and standard deviations for each axis of the accelerometer for each pose are stored in a pose bank for subsequent tasks. This is repeated for every subject. The training phase and developing a customized pose dictionary are essential for pose classification. A high-level flowchart illustration the use of training data to generate a generic pose dictionary for the application is shown in Fig. 3.

D. Feature Extraction Phase

The proposed feature extraction framework has the three key steps. **Segmentation** is used mainly to automatically determine the beginning and end of the pose. Based on the experiments and observation of the data from the accelerometers, a pose is defined as follows. A pose begins with a fast acceleration, a continuous change in the direction throughout the duration of the pose, and it ends with a stop of the movement. One approach is to do this manually, i.e., every new set
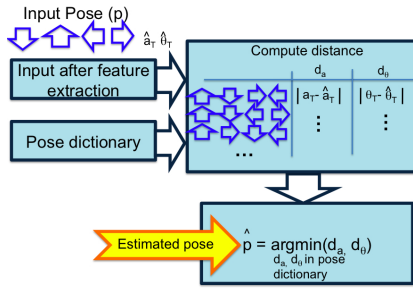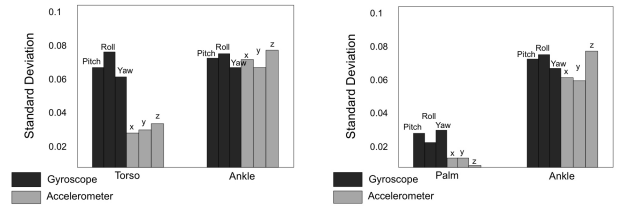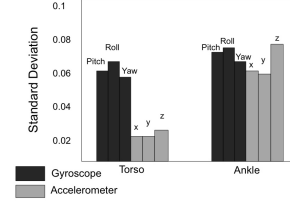
Fig. 3.   High-level flowchart of the classification algorithm.

(a) Experiment 1.

(b) Experiment 2.

(c) Experiment 3.

Fig. 4.   Standard deviations in the measurements of the motion sensors over multiple trials for a single user.

of data is preceded and followed by a sudden jerk. With sufficient data points, an empirical approach was to check the magnitude of the parameter, $\mathcal{D}$, defined as $\mathcal{D} = \sqrt{(x_k - x_{k-1})^2 + (y_k - y_{k-1})^2 + (z_k - z_{k-1})^2}$, where x_k, y_k, and z_k are the values of the accelerometer at instant k. If $\mathcal{D} > 2.8$, the segmentation begins, and stops when $\mathcal{D} < 1.2$, indicative of a pose ending. **Filtering** is used in order to eliminate some parts of the data stream that do not contribute to the gesture. This work uses two low pass filters to reduce noise effects in data sets, and used the Bayesian decision rule to set the thresholds. If $\mathcal{D} < 1.4$, the data is not included in the pose data stream. Again, cross validation was used to verify these empirical result based on the data points obtained from a single subject. The **Classifier** is used in order to identify the input gestures according to the database. Acceleration and gyroscopic measurements from the sensor boards are read in every time a pose is executed and compared to the pose signatures obtained from the training phase. Every pose is characterized by the mean and standard deviations of the components of its training feature vectors. The distance between an unknown sample M (input) and the mean of the features of class m (trained pose signature out of a total of L), $d(M, m)$, is then computed. The unknown input is then assigned to a class m^* (or move m^*) for which the distance, $d(M, m)$ is the minimum. Mathematically, $m^* = \min\{d(M, m)\}, \ i = 1, 2, \ldots, L$. The distance, $d(M, m)$ is the Mahalanobis distance metric, and is based on correlations between variables by which different patterns can be identified and analyzed. When the correlation between the patterns decreases, then the Mahalanobis distance metric reduces to the more widely used Euclidean distance metric. If the distance falls within an acceptable threshold, the pose is declared as being detected. A high-level flowchart of this classifier is summarized in Fig. 3.

IV. EXPERIMENTAL RESULTS

A series of experiments were performed to examine the capabilities and limitations of the two-sensor-based pose detection architectures in the mobile gaming context. The subjects executed a sequence of typical DDR steps multiple times consisting of 24 distinct poses for Experiment 1, and five distinct poses for Experiments 2 and 3. Tests were conducted under supervision with a consistent level of instruction and coaching given by a DDR professional.

A. Placement of Sensors

Figs. 4(a)–4(c) show the standard deviations in the measurements in the accelerometer and gyro (all six dimensions) at the different locations during each experiment (multiple trials). The larger the standard deviation in the measurements, the higher is the probability for successful pose recognition. In other words, very minuscule standard deviation in measurements is indicative of a poor choice in the location of the sensor board on the human body. From Fig. 4(b), because the standard deviation in the accelerometer and gyro measurements is very low, it implies that placing the sensor board on the palm is ineffective for pose recognition. In other words, the set up in Experiment 2 reduces to just a single sensor board (on the ankle) as the one on the palm is redundant. Hence, due to space constraints, we will only present results from Experiments 1 and 3.

B. Pose Difficulty

During the "learning" phase, it was noted that certain poses were consistently problematic. Training data from a single individual was tested over 20 trials using a

low recognition threshold determined by continually increasing the allowable recognition threshold in order to determine the level of difficulty among the different poses for the experiments. However, with sufficient training and practice, the recognition rates can be enhanced. Although it is desirable for the population to be capable of performing every pose, a difference in complexity among the poses can prove to be beneficial in some aspects of game development, allowing for an added level of difficulty in the game. The results from the experiments suggest a need for a larger training data set for the poses in order to accommodate a wider variety of body types and physical limitations in addition to expertise levels. This should not be surprising, because it shows that a large training sample is not necessarily required to create games that are playable by a range of people.

C. Threshold

It is unreasonable to expect a user to be able to achieve a "perfect" pose, in other words achieve a recognition distance of zero, even if they are a part of the training sample. Simple variations are inevitable due to the core mechanics of the human musculature system. Therefore, an allowable range, defined by a threshold, is experimentally determined to allow reasonably similar poses to be detected while avoiding false alarms. The threshold for an individuals data set was determined by continually increasing the allowable Mahalanobis distance after each trial until the lowest possible value is reached where all user poses could be recognized. It should be noted that the threshold for one individual data set will not necessarily be the same as anothers since their standard deviations will vary based on how they perform the pose in the training phase and how steady they are during recording. From Figs. 5(a) and 5(b), it can be observed that thresholds of 17.8 and 21.2 are acceptable because they achieve a pose recognition rate of close to 98% and 80% for Experiments 1 and 3, respectively. Once an acceptable threshold was found, a confusion matrix was created to determine the likelihood of false positives among other poses. The results indicate little possibility for the occurrence of false positives.

V. CONCLUSIONS

In this work, it has been shown that using a combination of six sensors (three accelerometers and three gyros), one can detect and discern human dance movements pertaining to a DDR-like game with a reasonably high probability. These initial results obtained using two sensor boards with the form factor of mobile phones, but without the processor serves as a motivation for future
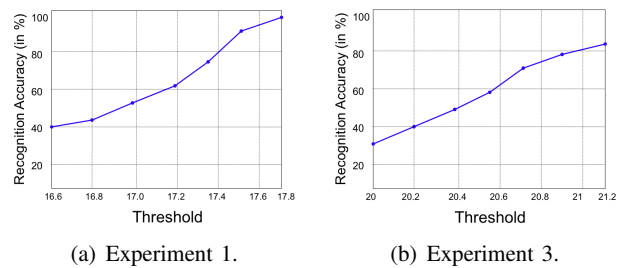
(a) Experiment 1.   (b) Experiment 3.

Fig. 5.   Success rates and thresholds for a single user.

research in mobile gaming using dual mode design phones. We have created a relatively inexpensive way of accurately identifying specific poses for gaming on a mobile platform. Though a new user may not achieve 100% pose recognition, it has been shown that with sufficient practice, improvement will be seen. The effect is desirable in game development to help maintain user interest over time. Although this new and original human pose recognition technology was tested for DDR, the same technology may be used to help the elderly stay at home. This technology can be applied to track and monitor the movements of elderly people to guarantee they are getting up, to detect falls, or to monitor electric wheel chairs so that they may go outside alone thereby assisting the elderly while they live independently in their own homes where they feel safe.

REFERENCES

[1] J. O. B. Soh and B. C. Y. Tan, "Mobile gaming," *Commun. ACM*, vol. 51, no. 3, pp. 35–39, 2008.

[2] H. J. Luinge, P. Veltink, and C. Baten, "Estimation of orientation with gyroscopes and accelerometers," *First Joint BMES/EMBS Conf. Serving Humanity, Advancing Technology*, 1999, pp. 844.

[3] K. Motoi *et al.*, "Evaluation of a new sensor system for ambulatory monitoring of human posture and walking speed using accelerometers and gyroscope," *Proc. SICE Annual Conf.*, Fukui, 2003, pp. 1223–1235.

[4] M. F. Nowlan, "Human recognition via gait identification using accelerometer gyro forces." Available: cs-www.cs.yale.edu/homes/mfn3/proj/mfn_gait_id.pdf

[5] T. Iso and K. Yamazaki, "Gait analyzer based on cell phone with a single three-axis accelerometer," *Proc. 8th Conf. on Human-Comput. Interaction with Mobile Devices and Services*, 2006, pp. 141–144.

[6] D. Karantonis *et al.*, "Implementation of a real-time human movement classifier sing a triaxial accelerometer for ambulatory monitoring," *IEEE Trans. Inf. Technol. Biomed.*, pp. 156–167, 2006.

[7] K. -T. Song and C. -W. Chen, "Multi-person pose recognition using a zigbee sensor network," *Proc. 11th World Congress the Intl. Federation of Automatic Control*, Jul. 2008, pp. 14976–14981.

[8] N. Crampton *et al.*, "Dance, dance, evolution: Accelerometer sensor networks as input to video games," *Proc. IEEE Intl. Workshop on Haptic Audio Visual Environments and their Applications*, Oct. 2007, pp. 107–112.