

Multi-sensor Based Gestures Recognition with a Smart Finger Ring

Mehran Roshandel¹, Aarti Munjal², Peyman Moghadam³,
Shahin Tajik¹, and Hamed Ketabdar⁴

¹ Deutsche Telekom Innovations Laboratories, Ernst-Reuter Platz, 10587 Berlin Germany
{mehran.roshandel, shahin.tajik}@telekom.de

² Department of Biostatistics and Informatics, University of Colorado Denver,
13001 E 17th Pl Aurora CO 80045
aarti.munjal@ucdenver.edu

³ Autonomous Systems, CSIRO Computational Informatics, 1 Technology Court,
Pullenvale, QLD 4069
peyman.moghadam@csiro.au

⁴ Quality and Usability Lab, TU Berlin Deutsche Telekom Innovation Laboratories
Ernst-Reuter-Platz 7, 10587 Berlin Germany
Hamed.Ketabdar@clickandbuy.com

Abstract. Recently several optical and non-optical sensors based gesture recognition techniques have been developed to interact with computing devices. However, these techniques mostly suffer from problems such as occlusion and noise. In this work, we present Pingu, a multi-sensor based framework that is capable of recognizing simple, sharp, and tiny gestures without the problems mentioned above. Pingu has been calibrated in the form of a wearable finger ring, capable of interacting even when the device is not in the vicinity of the user. An advanced set of sensors, wireless connectivity, and feedback facilities enable Pingu for a wide range of potential applications, from novel gestures to social computing. In this paper, we present our results based on experiments conducted to explore Pingu's use as a general gestural interaction device. Our analysis, based on simple machine learning algorithms, shows that simple and sharp gestures performed by a finger can be detected with a high accuracy, thereby, establishing Pingu as a wearable ring to control a smart environment effectively.

Keywords: Human Computer Interaction (HCI), Touch less gestural interaction, Wearable device, Finger ring.

1 Introduction

Gesture recognition is one of the important fields in Human Computer Interaction (HCI) as it enables users to interact with their computing devices more easily and naturally. In addition, use of multiple sensors for gestural recognition can extend the ability of accepting 3D inputs, which is not supported by conventional input devices.

Several applications, such as sign-language recognition, physical activity monitoring, and social interactions, can be developed based on simple gestures made by humans.

Several gesture recognition techniques based on optical and non-optical sensors were developed. Optical-based gesture recognition methods use optical sensors such as cameras [1, 2] or infrared (IR) sensors [3], to capture the movements of fingers and interpret them to commands. While the optical-based methods have problems such as occlusion, the non-optical methods try to use magnetometer [4, 5], accelerometer [7, 8], and proximity sensors [9] to overcome the limitations of optical gestural recognition techniques. However they have also their weaknesses, such as working in a limited space near the user device (MagiTact [4], MagiThings [17] and MagiSign [16]), accepting only 1D input (Nenya [5]), or in general they may be not socially acceptable.

In this work, we used our framework Pingu [6], a wearable and small finger ring that can interact with other electronic devices more naturally. While gestures made by any part of the body can be used for interacting with a computing device, previous research based on experiments conducted by Card et al. [10] shows that the information entropy of a finger-based interaction is much larger than the interaction based on any other human body parts. For instance, the interaction with the arm and wrist have the information rates of 11.5 and 25 bits/s respectively, while the information rate of finger is 40 bits/s. Therefore, Pingu is calibrated in the form of a finger ring with the following features:

1. It is composed of an advanced set of sensors (gyroscope, accelerometer and magnetometer).
2. It is equipped with wireless connectivity that makes it suitable for use in ubiquitous human-computer or human-human interaction.
3. The tiny size in the form of a finger ring makes Pingu wearable and, thus, socially acceptable.

Pingu is also capable of accepting 3D inputs from different gestures which are performed either in air or on surfaces such as a user's palm, top of the table. These gestures can be used in developing several interesting applications, including remote controlling or signature recognition. In this work, we analyze Pingu for recognizing a range of simple gestures. The main contribution of our work is to present results based on a multi-sensor interaction framework and effective classification algorithms. Our analysis shows that these generic gestures can be recognized with high accuracy.

The rest of this paper is organized as follows. In section 2 we review the related gesture recognition's solutions. Then in section 3 we explain the architecture of Pingu and its hardware. Experimentation and feature extraction via Pingu is discussed in section 4 and the results of gesture classification via different machine-learning algorithms are shown in the section 5. Finally we conclude the paper in the section 6.

2 Related Works

There are different gesture recognition approaches which have been developed in recent years and can be categorized into two groups: optical and non-optical gestural recognition techniques.

In optical-based gestural recognition approaches, optical sensors like cameras (e.g. SixthSense [1] and Gesture Pendant [2]) or infrared (IR) sensors (e.g. SideSight [3]) are the essential components to recognize the movements of fingertips and hands to interpret them to different commands. Although these approaches perform gesture recognition in some applications accurately, they do not support applications that are required to work with no direct line of sight (occlusion problem). Furthermore, optical data is sensitive to illumination conditions and, therefore, can only be used in certain circumstances. Finally, the user should wear additional cap or pendant which may be obtrusive and/or socially unacceptable.

On the other hand, non-optical gestural recognition methods use sensors such as magnetometer (e.g., MagiTact [4] and Nanya [5]), accelerometer [7, 8 and 11], and proximity sensors (e.g., Gesture Watch [9]) to mitigate the problems of optical-based methods.

Although proximity sensor solves the illumination problems, it still has the occlusion problem, as the gestures should be captured in the line-of-sight of sensors. Other methods based on accelerometer [7, 8, 11] do not have the occlusion and illumination problems, but since the acceleration data is very sensitive to noise, complementary sensors should be used. Techniques based on magnetometer send interaction commands when the magnetic field around the computing device is deformed. The advantage of this method is that there is no occlusion and illumination problem like previous approaches.

The gesture recognition techniques can also be categorized into types of wearable devices which they are embedded in. In some techniques, user should wear additional gloves such as Acceleration Sensing Glove [11] to interact with the computing device. The disadvantage of working with gloves is that they can be socially unacceptable or obtrusive. Other techniques like SixthSense [1] or Gesture Pendant [2] which require users to wear additional hat and pendant respectively, suffer from the same problems.

One possible solution is to develop the gestural recognizer as a ring or wristwatch, which may be socially more acceptable. Pinchwatch [12] is one of these systems which use a wristwatch for finger gesture recognition with the help of a camera. Users invoke functions by pinching and entering parameters by performing sliding and dialing motions. However, again this suffers from the problem of line of sight. Our previous work, MagiTact [4], involves interacting with a computing device equipped with an embedded compass (magnetic) sensor via a magnet placed on a finger. Coarse gestures made with the magnet affect the magnetic field around the device and, thus, used for gestural interaction. Although this approach has no occlusion problems, interaction is still limited to the immediate 3D space around the device.

More recently Nanya [5] a magnetically-tracked finger ring is developed which includes a permanent magnet in the form of a finger ring and worn-watch wireless tracking bracelet. While magnetometer is used to track the ring's position, a Bluetooth radio allows the bracelet to send ring input to the user's devices. Nanya supports only 1D input in comparison to Pingu which supports 3D input. Furthermore, it consists of two accessories in contrast to Pingu which includes all sensors and radio in only one ring. Magic Ring [13] is another finger-worn device which is developed for using static finger gestures and it uses accelerometer data to detect different gestures. Magic Ring is tested with six different finger gestures with doing some predefined task.

In our approach we classify nine finger gestures with four machine learning algorithms to derive the accuracy of gesture recognition.

3 Design

Using multiple sensors provides rich information related to motion and angular position of the device and, thus, results in a high accuracy in gesture recognition. Pingu is equipped with four sensors: accelerometer, magnetometer, gyroscope, and proximity sensor. A tri-axel accelerometer is used to detect the orientation and motion of the device along three axes x , y , and z . A tri-axel gyroscope detects the angular rate of movement of the ring along x , y , and z axis. Using gyroscope in addition to accelerometer provides six degree of freedom and can be used to detect the 3D trajectories of the ring. Additionally, the deformation of magnetic fields is useful in recognizing the coarse gestures made around the device. Moreover, Pingu has proximity-sensing plates installed, which allow sensing the proximity of other fingers. Figure 1 shows the prototype of Pingu and Table 1 lists the configuration of the sensors and radio which are used in the design of Pingu.

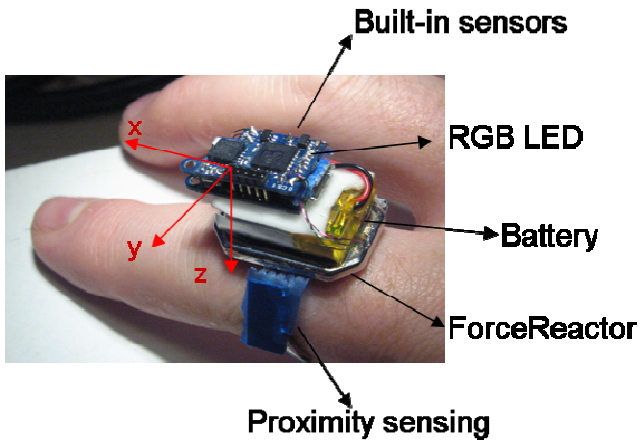


Fig. 1. Pingu, our multi-sensor framework, for interaction with a smart environment

Table 1. Sensors used in the design of Pingu and their specifications

| Sensor | Description |
|---------------|---|
| Accelerometer | $[-8g, 8g]$ |
| Magnetometer | $[-2\text{gauss}, 2\text{gauss}]$ |
| Gyroscope | $[-2000\text{deg/s}, 2000\text{deg/s}]$ |
| Bluetooth | Up to 2m |

4 Experiment

In this work, we evaluate Pingu for general gestural interaction. For this purpose, we have defined a set of nine general gestures shown in Figure 2. As shown, the gestures are highly general in nature and can be used to control smart environments. For example, gesture 1 and 2 can substitute the volume control buttons on a remote control and gesture 5 and 6 can change music tracks forward and backward respectively. To evaluate Pingu for general gestural interaction in smart environments, we perform all the nine gestures in three ways, as follows:

1. General Gestures in the air, in which a gesture is performed in the air.
2. General Gestures on the table, in which a gesture is performed on the top of the table.
3. General Gestures on the palm, in which a gesture is performed on user’s palm.

Setting the three medium of air, palm and desk provides a variety of surfaces for gesturing. In this way, the methodology can be tested under more variable yet practical scenarios. Both palm and desk are surfaces which are commonly available for users during the gesturing process. The air medium also provides the fantasy of writing in air for the user, when the two other mediums are not available.

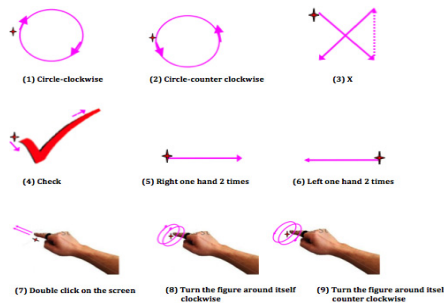


Fig. 2. A set of nine general gestures used in this work

Our results are based on a dataset collected from 24 users.

Table 2. User statistic Table

| Total Users | Male | Female | Right Handed | Left Handed |
|-------------|------|--------|--------------|-------------|
| 24 | 10 | 14 | 20 | 4 |

Feature Extraction Each of the nine gestures shown in Figure 2 is performed 15 times per user. The sensor readings specific to each of the nine gestures are then captured via a Java desktop application developed for Mac OS. To evaluate the interaction made by Pingu, we classify gestures based on the sensor readings collected for each gesture. In particular, we adopt the following approach:

We mix the data collected from all the 24 users and cross-validate. For this purpose, we form a feature vector containing data specific to each sensor. For example, a feature vector obtained from the accelerometer used in Pingu contains the following:

1. Mean of the linear acceleration along x, y, and z axis (3 features),
2. Variance of the linear acceleration along x, y, and z axis (3 features),
3. Mean of the Euclidian norm of the linear acceleration along x, y, and z axis (1 feature),
4. Variance of the Euclidian norm of the linear acceleration along x, y, and z axis (1 feature),
5. Standard Deviation of the linear acceleration along x, y, and z axis (3 features),
6. Piecewise correlation between linear acceleration along x, y, and z axis (3 features), and
7. Frequency features along x, y, and z axis (3 features).

As shown in Figure 2, mean and variance of the sensor readings obtained for gestures 1 and 2 may not be able to differentiate between these two gestures. Therefore, we have included the piece-wise correlation and the frequency features specific to each sensor. The feature vector for the angular rate movement of the ring (i.e., from gyroscope) is obtained in a similar manner. Feature vector for each sensor, therefore, contains 17 elements. Since multiple windows provide more detailed information in gesture classification, our results are based on 4 windows. Feature vectors obtained from each window are concatenated to form a new feature vector of 68 (=17×4) features. To further validate that Pingu is effective in gestural interaction in a smart environment, we do not use the magnetometer readings in this analysis.

5 Gesture Classification

The feature vectors obtained for each of the three experiments are then used as an input to a classification algorithm for gesture classification. Specifically, we have classified the gestures with four classifiers: (a) Decision Tree (DT), a decision tool that uses graphs and model of decisions to derive the outcomes and consequences, (b) Multi-Layer Perceptron (MLP), a feed forward artificial neural network that models the relationship of inputs and outputs to find the patterns, (c) Naïve Bayes (NB), a probabilistic classifier that uses Bayes' theorem with strong independence assumptions, and (d) Support Vector Machines (SVM), a set of hyperplanes in high dimensional space for using classification and regression. Our analysis is based on the implementation of these classifiers in the Weka machine learning toolkit [14, 15]. Tables 2-4 list the classification results obtained for all the three experiments. As shown in Tables 2 and 3, MLP classifier outperforms all the other three classifiers for gestures performed in air and on table, with more than 97% and 93% accuracy, respectively. Table 4 shows that SVM has better results with more than 77% accuracy. To illustrate further how accurate different gestures can be distinguishable, confusion matrices obtained from MLP classifier for General Gestures performed in air are shown in Table 5. As shown, all nine gestures are most of the times distinguishable, but more specifically the confusion matrices indicate that gesture 1 and gesture 2 are somewhat more difficult to classify. Similarly, we note that gesture 8 and gesture 9 are classi-

fied with lower accuracy due to the inherent similarity in performing these gestures. On the other hand, gesture 4 is the easiest to be classified, as it's easily distinguishable from other gestures. These results show that generally, the gesture recognition by Pingu is trustworthy.

Table 3. Gesture Classification Results for General Gestures in the air

| Algorithm | Accuracy |
|-----------|----------------|
| MLP | 97.879% |
| DT | 87.043% |
| NB | 57.389% |
| SVM | 96.443% |

Table 4. Gesture Classification Results for General Gestures on the table

| Algorithm | Accuracy |
|-----------|----------------|
| MLP | 93.689% |
| DT | 73.080% |
| NB | 49.238% |
| SVM | 83.991% |

Table 5. Gesture Classification Results for General Gestures on the palm

| Algorithm | Accuracy |
|-----------|----------|
| MLP | 71.130% |
| DT | 71.160% |
| NB | 33.415% |
| SVM | 77.541% |

Table 6. Confusion matrix obtained from MLP for the results shown in Table 3

| Gesture | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 1 | 333 | 3 | 0 | 0 | 0 | 1 | 0 | 3 | 0 |
| 2 | 3 | 334 | 2 | 0 | 0 | 0 | 0 | 0 | 1 |
| 3 | 2 | 1 | 336 | 0 | 1 | 0 | 0 | 0 | 0 |
| 4 | 1 | 1 | 0 | 343 | 0 | 0 | 1 | 1 | 0 |
| 5 | 0 | 0 | 0 | 1 | 338 | 2 | 0 | 0 | 0 |
| 6 | 1 | 0 | 0 | 0 | 3 | 324 | 0 | 0 | 0 |
| 7 | 0 | 1 | 0 | 0 | 0 | 0 | 340 | 2 | 0 |
| 8 | 7 | 0 | 1 | 0 | 0 | 0 | 0 | 320 | 12 |
| 9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 14 | 332 |

6 Conclusions and Future Work

In this work, we have presented our results for gestural recognition using a multi-sensor based framework called Pingu. Our results are based on a set of nine pre-defined general gestures that can be used to interact in a smart environment. Pingu is a socially wearable, small finger ring that is equipped with multiple sensors to provide rich information about the general gestures made by a user. Our analysis is based on a large dataset of 24 users. We have shown that with simple classification algorithms, different gestures can be distinguished from each other with high accuracy. Therefore, we can trust Pingu to be involved in many interesting applications such as remote controlling, signature recognition, physical activity analysis and sign-language recognition.

References

1. Mistry, P., Maes, P.: SixthSense: A wearable gestural interface. In: ACM SIGGRAPH ASIA 2009 Sketches. ACM (2009)
2. Starner, T., et al.: The gesture pendant: A self-illuminating, wearable, infrared computer vision system for home automation control and medical monitoring. In: The Fourth International Symposium on Wearable Computers. IEEE (2000)
3. Butler, A., Izadi, S., Hodges, S.: SideSight: Multi- “touch” interaction around small devices. In: Proc. UIST, pp. 201–204 (2008)
4. Ketabdar, H., Yüksel, K.A., Roshandel, M.: MagiTact: Interaction with mobile devices based on compass (magnetic) sensor. In: Proceedings of the 15th International Conference on Intelligent User Interfaces. ACM (2010)
5. Ashbrook, D., Baudisch, P., White, S.: Nanya: Subtle and eyes-free mobile input with a magnetically-tracked finger ring. In: Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems. ACM (2011)
6. Ketabdar, H., Moghadam, P., Roshandel, M.: Pingu: A new miniature wearable device for ubiquitous computing environments. In: 2012 Sixth International Conference on Complex, Intelligent and Software Intensive Systems (CISIS). IEEE (2012)
7. Wu, J., Pan, G., Zhang, D., Qi, G., Li, S.: Gesture recognition with a 3-d accelerometer. In: Zhang, D., Portmann, M., Tan, A.-H., Indulska, J. (eds.) UIC 2009. LNCS, vol. 5585, pp. 25–38. Springer, Heidelberg (2009)
8. Fukumoto, M., Tonomura, Y.: “Body coupled FingerRing”: wireless wearable keyboard. In: Proceedings of the SIGCHI Conference on Human Factors in Computing systems. ACM (1997)
9. Kim, J., et al.: The gesture watch: A wireless contact-free gesture based wrist interface. In: 2007 11th IEEE International Symposium on Wearable Computers. IEEE (2007)
10. Card, S.K., Mackinlay, J.D., Robertson, G.G.: A morphological analysis of the design space of input devices. *ACM Trans. Inf. Syst.* 9(2), 99–122 (1991)
11. Perng, J.K., Fisher, B., Hollar, S., Pister, K.S.J.: Acceleration sensing glove (ASG). In: The Third International Symposium on Wearable Computers (ISWC 1999), pp. 178–180 (1999)
12. Loclair, C., Gustafson, S., Baudisch, P.: PinchWatch: A wearable device for one-handed microinteractions. In: Proc. MobileHCI (2010)

13. Jing, L., et al.: Magic Ring: A Finger-worn device for multiple appliances control using static finger gestures. *Sensors* 12(5), 5775–5790 (2012)
14. Weka3: Data Mining Software in Java,
<http://www.cs.waikato.ac.nz/ml/weka/>
15. Witten, I.H., Frank, E., Hall, M.A.: *Data Mining: Practical Machine Learning Tools and Techniques: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann (2011)
16. Ketabdar, H., Moghadam, P., Naderi, B., Roshandel, M.: Magnetic signatures in air for mobile devices. In: *Mobile HCI 2012*, pp. 185–188 (2012)
17. Ketabdar, H., Abolhassani, A.H., Roshandel, M.: MagiThings: Gestural Interaction with Mobile Devices Based on Using Embedded Compass (Magnetic Field) Sensor. *IJMHCI* 5(3), 23–41 (2013)