# Evaluating interaction with Popie using coin size wireless sensor

Atia Ayman Daisuke Sato Shin Takahashi Jiro Tanaka

Department of Computer Science,University of Tsukuba

## 1. Introduction

Human hand gestures are typical way of communication between people. In ubiquitous environment where computers are in every where interaction with such computers can be done through various interfaces. Each interface has its own method of interaction.
Popie is a flow menu pen based styles interface that is used to facilitate entering of Japanese text [1]. The primary results of our previous research shows that interacting with Popie using a wireless sensor could be worth [2]. In this research we propose an evaluation experiments for interaction with Popie. We studied the accuracy and time of recognizing hand gestures while interacting with Popie.

## 2. System overview

### 2.1 Popie

Popie is a menu-selection-based Japanese input method for a pen device [1]. The system starts by allowing the user to select some characters constants then recommends the user a set of recommended words "Kanji" to select from.

The user can input consonants, and select kanji-form candidates fluidly, using FlowMenu. Figure 1(a) shows Popie menu. Popie is composed of eight basic directions and each direction has right and left direction flow so a total of 24 gestures.

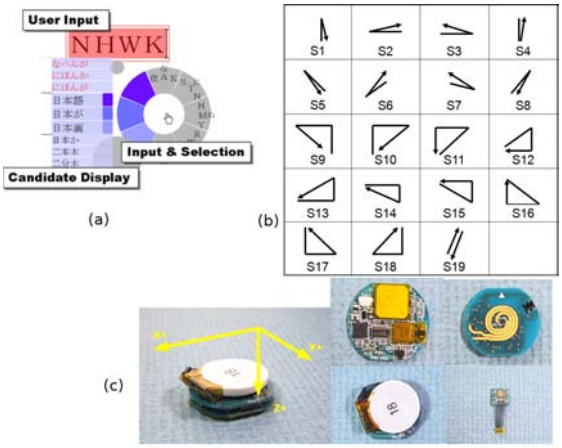


Figure 1: (a) Popie menu, (b)Popie gestures of operating, (c) Nokia Cookie Wireless Sensor

### 2.2 Nokia wireless coin size Sensors

In this paper we use Nokia cookie Figure 1(c) which is an experimental device developed to test ubiquitous context aware applications. The chassis contains 6 sensors 3-axis linear accelerometer, Compass 3-axis sensor, Ambient light sensor, Galvanic skin response sensor, Heart rate sensor and Skin temperature sensor. It transmit data using Bluetooth. The device is about the size of 5 stacked quarter coins [5].

### 2.3 Recognition Mmethod

Accelerometers have been used for activity recognition during assembly tasks [4]. There work focuses on extracting a set of activities to promote the context awareness of user's locations and activities. In order to allow users to define their own gestures we use a gesture recognition toolkit called GART which uses Hidden Markov Model (HMM) for pattern recognition. GART has been tested before by using a regular mouse device to recognize user's gestures [3].

## 3. Evaluating Hand Popie

It is required to measure two main features during Popie interaction. First, the time to enter sequence of characters with Popie. Second, the recognition error rate with respect to gestures shape.

To operate hand Popie we need to define 8 basic gestures in the X,Y plane which are moving towards Up, Left, Right, Down, Up Left, Down Left, Down Right and Up Right. Popie operates in 3 steps: (1) user selects one of the basic 8 directions (2) user choose either to move right or left with relative to the basic direction or go to step 3, (3) user return back to start point. So as a result we have 24 gesture. Figure 1(b) shows the directions of Popie hand gestures. We name these gestures S1, ... ,S18 respectively. S19 is used as a selector gesture and it is done on the X and Z plane. Some gestures have been ignored from the original 24 gestures of Popie as they are not practically used for Popie like Left then left as it is replaced by Down Left or Left then right as it can be replaced by Up Left and so on. In general the right part of Popie is frequently used.

### 3.1 Experiment setup

In a preliminary study we ask one subject non-native Japanese user who is one of the authors and well trained on interaction with Popie and GART toolkit. We ask the user to train the system by entering 8 trials for each 19 hand Popie gestures sequentially a total of 152 gestures. The user defines his gestures in front of 50 inch plasma display screen space.

To test the time to enter sequence of characters with Popie we ask the user to enter the sentence "I am a student." in Japanese language "watashi wa gakusei desu.": "私は学生です。" (7 Japanese characters). To enter this sentence using Popie the user need to enter 16 hand gestures as follows: (S7, S8, S2, S19, S16, S7, S6, S16, S8, S8, S9, S2, S3, S17, S5, S19). We ask the user to input the sentence twice.

To test the effect of location on the accuracy of recognition, we asked the user to repeat the same ex-

periment but in front of small size 15 inch display using different set of 19 gestures.

### 3.2 Time to enter sequence of characters

The system measures time required to enter the above sentence . The user takes an average of 2 minutes and 26 sec for the whole session in front of the 50 inch display screen. The results were dramatically decreased while interacting with 15 inch display. The average time for the whole session was 1 minute approximately. We think that the reduction of time was because the user have to make short distance gesture shapes while fixing his elbow during siting in front of small screen.

### 3.3 Recognition error rate

We observe the confusion matrix while entering the above sentence in front of the 50 inch and the 15 inch display screen.

While defining the gestures in front of the big display screen the user has to define big motion shape for gesture as a result he was uncomfortable because of hand tremor and pain while training the GART for 152 gestures. This tremor affect the sharp definition of hand gestures of Popie. Figure 2 shows a confusion with 11 gestures with a maximum average error rate of 5 counts.

It can be noted that S3 (Right) having the maximum average error rate in the confusion matrix with gesture S7 (Down Right). This is because both of the gestures are very near in shape while interacting with large screen display, the angle between the right and downright could be very small. As a result a similar confused shape. This error can be avoided if the user make his gestures in more precise and sharp shape.

When the user interacts with 15 inch display, the results show a confusion with 7 gestures with maximum average error rate of 2 counts. Figure 3 shows the average confusion matrix. Looking back to gesture S3,S5,S2 we recognized that in front of smaller display screen low error counts achieved. The explanation to this is doing gestures in sitting position can help the user to make an acute gesture shapes, which can be easily recognized by the GART engine.

Form Figure 2 and Figure 3 we can found that there are some gestures (S6, S7, S16, S17, S19) did not affected with the place of doing the gesture i.e. there error rate is 1 or below. The common thing between all of these gestures are that they all are movements towards user human body which might give them some acute shapes.

| | | Original Gesture | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $6 | $7 | $8 | $9 | $10 | $11 | $12 | $14 | $15 | $16 | $18 | Err |
| User Gesture | $7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.5 | 0 | **0.5** |
| | $8 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | **1** |
| | $2 | 1.5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1.5 | 0.5 | 0 | **3.5** |
| | $19 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | **1** |
| | $16 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | **0** |
| | $7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0.5 | 0 | **0.5** |
| | $6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | **0** |
| | $16 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | **0** |
| | $8 | 0 | 0 | 0 | 0 | 0.5 | 0 | 2.5 | 0.5 | 0 | 0 | 0 | **3.5** |
| | $8 | 0 | 0 | 0 | 0 | 1 | 0 | 0.5 | 0 | 0 | 0 | 0 | **1.5** |
| | $9 | 0 | 0 | 0 | 0 | 0.5 | 0 | 0 | 0 | 0 | 0 | 0 | **0.5** |
| | $2 | 0 | 0 | 0 | 0.5 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | **0.5** |
| | $3 | 0 | 3 | 0.5 | 0 | 0 | 0 | 0.5 | 0.5 | 0 | 0.5 | 0 | **5** |
| | $17 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | **1** |
| | $5 | 0 | 0 | 0 | 1.5 | 0.5 | 0.5 | 0 | 0 | 0.5 | 0 | 0 | **3** |
| | $19 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | **0** |

Figure 2: In front of 50 inch display screen

| | | Original Gesture | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | $6 | $7 | $11 | $12 | $14 | $15 | $17 | Err |
| User Gesture | $7 | 0 | 0 | 0 | 0 | 0.5 | 0 | 0 | **0.5** |
| | $8 | 0 | 0 | 0.5 | 1 | 0 | 0 | 0 | **1.5** |
| | $2 | 0.5 | 0 | 0 | 0 | 0 | 0 | 0 | **0.5** |
| | $19 | 0 | 0 | 0 | 0 | 0 | 0 | 0.5 | **0.5** |
| | $16 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | **0** |
| | $7 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | **0** |
| | $6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | **0** |
| | $16 | 0 | 0.5 | 0 | 0 | 0.5 | 0.5 | 0 | **1.5** |
| | $8 | 0 | 0 | 0 | 2 | 0 | 0 | 0 | **2** |
| | $8 | 0 | 0 | 0 | 1.5 | 0 | 0 | 0 | **1.5** |
| | $9 | 0 | 0 | 0 | 0.5 | 0 | 0 | 0 | **0.5** |
| | $2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | **1** |
| | $3 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | **0** |
| | $17 | 0 | 0 | 0 | 0 | 0 | 0 | 0.5 | **0.5** |
| | $5 | 0 | 0 | 0 | 0 | 0 | 0.5 | 0 | **0.5** |
| | $19 | 0 | 0 | 0 | 0 | 0 | 0 | 0.5 | **0.5** |

Figure 3: In front of 15 inch display screen

## 4. Conclusion and future work

Although more comprehensive and objective comparison is still required to be investigated, the results shows that using the accelerometers for interaction with Popie achieves about 7 characters per minute (cpm). The original Popie can input characters on an average of 12cpm on the first session. From our observations we think that there are two factors for low accuracy of recognition, one is related to the length of defining the hand gestures shape. The second reason is using 8 samples only for training of the engine which is not enough for accurate recognition. In order to enhance interaction with Popie we have to look for more simple way of interaction with Popie with less hand gestures. Also we have to test another hand gesture recognition engine.

## References


[1] D. Sato, B. Shizuki, M. Miura and J. Tanaka. "Menu-selection-based Japanese Input Method with Consonants for Pen-based Computers," In *Proceedings of APCHI 2004*, pp.399–408, (2004).

[2] A. Atia, D. Sato, S. Takahashi, and J. Tanaka. "Hands Popie using coin size wireless sensor" *WISS2007*, pp.175–176, (2007).

[3] K. Lyons, H. Brashear, T. Westeyn, J. Soo, and T. Starner. "GART: The Gesture and Activity Recognition Toolkit," *HCII2007*, pp.718–727, (2007).

[4] A. and Jamie, G. Troster. "Activity recognition of assembly tasks using body-worn microphones and accelerometers," *IEEE Transactions on pattern and machine intelligence*, vol. 28, no. 10, pp.1553–1567, (2006).

[5] H. Ohhashi "Nokia Cookie users manual,"Nokia research center, Nokia Japan Co.,Ltd., (2005).