A Smartphone System: Providing a Shoe-Embedded Interface

Kaname Takaochi, Kazuhiro Watanabe, Kazumasa Takami
Information System Technology Department
Graduate School of Engineering, Soka University
Hachioji-shi, Tokyo, Japan
{e10m5222, kazuhiro, k_takami}@t.soka.ac.jp

Abstract— **Although a handsfree man-machine interface is useful when the user's hands are not free, existing handsfree input devices are not the type of device that are normally worn by people. We focus on a shoe as an input device because people normally wear it when they go out, and propose a shoe-embedded interface. The input device is a sensor shoe. Weight sensors are attached at three positions on a sole: the first metatarsal, the fifth metatarsal, and the calcaneal tuberosity. These positions have been selected based on the characteristics of the human foot skeleton. Two types of foot operation have been used: tap and push. By combining these operations, 10 commands have been defined. The sensor shoe houses an insole with hetero-core optical fiber sensor elements attached to it. These elements are sensitive to weight. We have built an experimental system that runs on a smartphone and provides the shoe-embedded interface, and conducted experiments with three test subjects to evaluate the system. The average rate of successful command identification was 89%.**

Keywords- shoe-embedded interface; heterocore optical fiber sensor; handsfree interface

I. INTRODUCTION

In the field of human-computer interaction, there are intensive studies on a man-machine interface [1]-[6]. A handsfree interface is useful for people who are in a public space and whose hands are not free, such as passengers holding baggage in an airport, parents holding small children, and golf players. Most handsfree interfaces with practical products already available use speech recognition. Speech recognition has been widely implemented in mobile information devices, such as smartphones, tablet terminals, and laptop PCs. Handsfree interfaces generally consist of an input device and a processing terminal. In cases where speech recognition is used, the input device is either built in a mobile information device or a microphone connected to the input port of a mobile information device, and the processing terminal is the mobile information device itself. The processing terminal conveys the user's intention to a given application by extracting a word from the waveforms sent from the input device, and identifying a pre-defined command that matches the word.

In cases where the user is in a public space and his/her hands are not free, a problem with conventional human interfaces is that the user needs to wear an input device just for the purpose of acquiring this interface whether the input device is an earphone-equipped microphone or a headset for voice input, a head-mounted display or eye-glass-like device for eye-tracking input, a cap-shaped input device for a brain-machine interface, or a camera to recognize a gesture or a motion. When the user does not need this interface, he/she has no need to wear such devices.

We have focused on a shoe because people always wear it when they go out. Specifically, we have chosen to use a shoe-embedded interface because it is suitable for use in a public space. The input device is a sensor shoe with a hetero-core optical fiber sensor element built in it [7]-[13]. The processing terminal is a smartphone. Section II gives an overview of the shoe-embedded interface. Section III describes three aspects of implementing the shoe-embedded interface: sensor shoe design, command input method, and command definition. Section IV describes the experimental system we have developed based on the proposed method. Section V reports on the experiment carried out using the experimental system, and evaluates the proposed method based on the experiment result. Finally, Section VI provides the conclusions and future work.

II. SHOE-EMBEDDED HUMAN INTERFACE

The shoe-embedded human interface is a wearable handsfree human interface. It consists of a sensor shoe and a processing terminal. A sensor shoe is a shoe with weight sensors. The weight sensors are attached to the insole of the user's shoe. The user puts his/her weight on the sensors to input an operation. The sensors are so thin that the shoe appears to be a normal shoe. The processing terminal identifies the user's operation from the weights measured by the sensors. The weight data is sent to the processing terminal using wireless communication. A smartphone is used as the processing terminal. Advantages of the proposed interface include resistance to noise, mobility and invisibility. The interface is highly resistant to noise because the user operates the sensor with his/her weight. It provides high mobility because the sensor shoe is not wired to the processing terminal. It is invisible to others because it requires only a sensor shoe and a smartphone. The user can use it without worrying about how he/she looks.

One of the criteria generally used to assess the ease of using a man-machine interface is usability. Usability is defined in ISO9241-11[14]-[16] as "extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.” ISO9241-11 goes on to define effectiveness as “accuracy and completeness with which users achieve specified goals,” efficiency as “resources expended in relation to the accuracy and completeness with which users achieve goals,” and satisfaction as “freedom from discomfort, and positive attitudes towards the use of the product.” Ishikawa [17] states that “usability is often evaluated in terms of the achievement of specified goals. In other words, it is evaluated with a defined evaluation task.” However, it is difficult to generalize the functions of the potential device that will be operated by the user. Therefore, we focus on the simple task of selecting a function when the user is in a public space and his/her hands are not free.

Figure 1 shows an example of the function selection task. In this example, the user in an airport selects the service of checking information about his/her reserved boarding pass from a list of services available. The user attaches his/her smartphone on the strap of his/her bag, and can see the display screen of the smartphone simply by looking down. First of all, the user inputs a start command to shift the system’s state from the walking state to the input state. Then, the processing terminal sends to the server an inquiry about services that are available at the airport, obtains a list of available services from the server, and displays it for the user. A unique command identifier is associated with each service. The user can recognize the associations between services and command identifiers from the positions of the identifiers on the screen. When the user inputs the command identifier associated with checking information about the reserved boarding pass, the processing terminal identifies it, and conveys this request to the server. In this case, effectiveness can be evaluated in terms of the probability at which the processing terminal correctly recognizes the command identifier for the service wanted by the user, or simply in terms of the rate of successfully identifying the intended command. Efficiency can be evaluated in terms of the amount of labor required to operate the user interface, or simply in terms of the number of input attempts. Since it is difficult to evaluate satisfaction in a general term, it is not addressed in this paper.

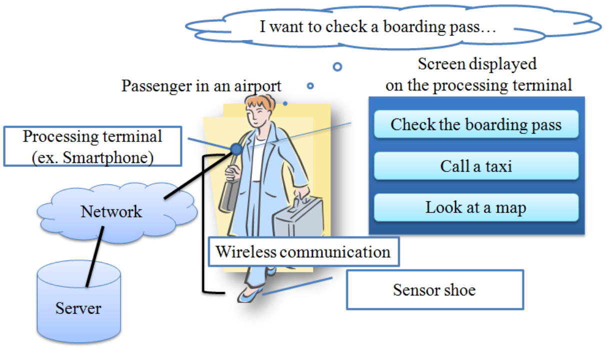


Figure 1. Example of the Use of the Proposed System.

III. PROPOSALS

A. Design of the Sensor Shoe

A sensor shoe has three weight sensors attached to the positions shown in Fig. 2. These positions have been determined based on the structure of the human foot skeleton. According to Noda [18], the plantar arch, a characteristic feature of the human foot skeleton, is made up of three arches linking three points: the first metatarsal (the base of big toe), the fifth metatarsal (the base of the little toe), and the calcaneal tuberosity (the heel area that touches the ground). Noda also states that, when the entire sole is touching the ground, the weight is distributed on the three points at the ratios of 2 on the first metatarsal, 1 on the fifth metatarsal, and 3 on the calcaneal tuberosity. The selection of the sensor positions is based on this finding.

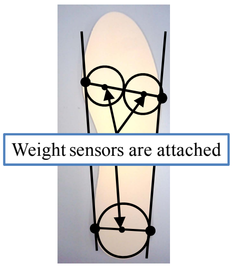


Figure 2. Positions of Weight Sensors.

B. Tap and Push Operation

The user’s operations are defined using variations in the user’s weight. Weight sensors handle only two types of data: duration in when the weight is measured and the weight value. Operations can be defined in terms of either duration or weight. In a method focusing on duration, a threshold is defined regarding the weight value to determine whether the user has intended to make certain operation or not. Multiple types of operation can be defined depending on the duration in which the user continues this operation. In a method focusing on weight, a point in time, such as 5000 ms after the transition to the input waiting state, is selected for the identification of the user’s intention. Multiple types of operation can be defined depending on the weight measured at that time. However, this method requires delicate control of the weight the user applies. Controlling the weight is more difficult than controlling the duration. Therefore, we have adopted a method focusing on duration.

A command is defined by a combination of two types of operations: a tap operation and a push operation. A tap operation is tapping the sole of the user’s shoe on the ground. In this operation, the change in weight is expected to show a triangular wave, as shown in Fig. 3. By setting a threshold on the weight, it is possible to detect this operation through two steps:

Step 1: Measure the weight that exceeds the threshold value
Step 2: Measure the weight when it is below the threshold value over a certain duration.

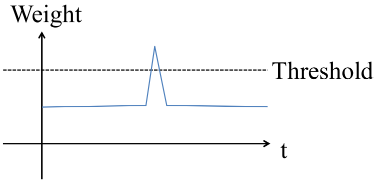


Figure 3. Change in Weight at the Time of Tap Operation.

A push operation is applying weight on the sensor casing as if the user is pressing the sensor. In this operation, the weight changes in the form of a trapezoidal wave, as shown in Fig. 4. By setting a threshold value, it is possible to detect this operation through three steps:

Step 1: Measure the weight that exceeds the threshold value
Step 2: The weight continues to exceed the threshold value for more than a certain time
Step 3: Measure the weight when it is below the threshold value

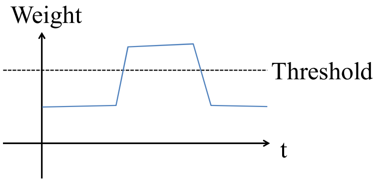


Figure 4. Change in Weight at the Time of Push Operation.

C. *Definition of Commands*

Ten commands have been defined by combinations of tap and push operations, as shown in Fig. 5. A unique command identifier is associated with each command.

Command ID	Operation Patterns
1	Push A -> Push A
2	Push A -> Push B
3	Push A -> Push C
4	Push B -> Push A
5	Push B -> Push B
6	Push B -> Push C
7	Push C -> Push A
8	Push C -> Push B
9	Push C -> Push C
10	Tap

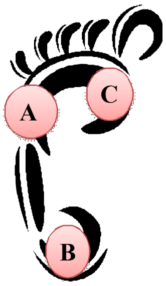


Figure 5. Definition of Commands.

Nine commands are represented by a push operation, and one command by a tap operation. The reason for defining the commands in this manner is as follows. Ishikawa [17] states that the task of selecting a function can be expressed as a hierarchical menu in a tree structure chart. If the number of commands to be defined for the task of selecting a function is small, the number of layers that must be crossed to reach the goal becomes large, resulting in an increased burden on the user because he/she needs to make a large number of input operations. Conversely, if the number of commands is large, it becomes a burden for the user to learn the required operations. One of the quantitative expressions of human's short memory capacity is Miller's Magical Number Seven, Plus or Minus Two [19][20]. He argues that the number of objects an average human can hold in working memory is around 7. This implies that if the number of commands used to select a function exceeds two digits, the burden of learning is large. Since such a burden reduces the efficiency of using the interface, it is necessary to minimize the number of commands. In the case of selecting one out of 50 functions, the relations between the number of commands and the calculated number of inputs are as shown in Table I. P stands for a "push operation," and T for a "tap operation." If the number of commands of push operation is 2, a tree that expresses 50 elements needs to have 6 layers. Just passing through each node requires one push operation (selection) and one tap operation (selection done). Therefore, to go through the 6 layers, a total of 12 operations are required. We have also studied other numbers of commands and found that, in cases where three weight sensor elements are attached to a sensor shoe, the number of required input operations is the smallest (i.e., the input operation is the most efficient) when the number of commands is 9. This is the reason why 9 commands are based on a push operation in this paper.

TABLE I. COMPARISON OF DIFFERENT NUMBERS OF COMMANDS IN TERMS OF EFFICIENCY

Number of commands for selecting functions	Required operations	Number of inputs	Burden of learning
2	(P→T)×6 times	12	Small
3	(P→T)×4 times	8	Small
4	(P:2 times→T)×3 times	9	Small
8	(P:3 times→T)×2 times	8	Small
9	(P:2 times→T)×2 times	6	Small
16	(P:4 times→T)×2 times	10	Large
27	(P:3 times→T)×2 times	8	Large

IV. IMPLEMENTATION

We have built an experimental system based on the method proposed in Section III. As shown in Fig. 6, the system consists of a sensor shoe, an optical measurement instrument, and a user operation detection application running on a smartphone[21][22].

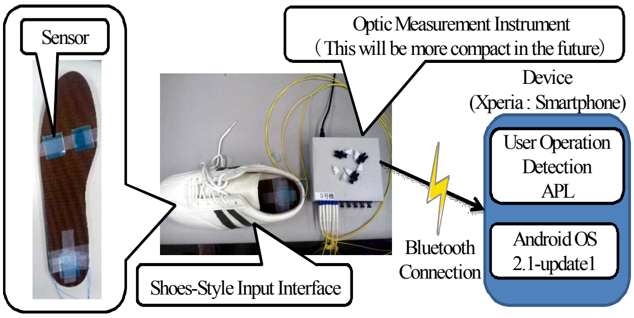


Figure 6. Structure of the Experimental System.

The sensor shoe is an insole. The parts of the insole where weight sensors are to be attached were removed and replaced with sensor casings. A hetero-core optical fiber weight sensor was used as a weight sensor. As shown in Fig. 7, a hetero-core optical fiber is composed of an optical fiber of a uniform core diameter with a small fiber segment with a different core diameter inserted. When the hetero-core optical fiber is bent, its optical loss increases. The hetero-core optical fiber weight sensor uses this property. It is highly sensitive to bending. When a weight is applied on the sensor casing, the fiber is bent, increasing its optical loss. The sensor detects how big the applied weight is by measuring the optical loss. The weight is actually measured by the optical measurement instrument. An LED/PD (Light Emitting Diode / Photo Diode) power meter was used as this instrument. The experimental system measures optical loss (in mV) and sends the measured value to the processing terminal every 33 ms. Since weight is expressed as optical loss, the larger the weight, the lower the level of the optical signal.

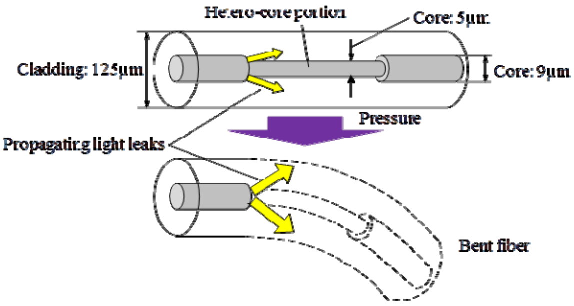


Figure 7. Hetero-core Optical Fiber.

The operation detection application software is configured as shown in Fig. 8, and is implemented on a smartphone. This application performs three functions: setting parameters for the experiment, calibration and operation detection. By calibration is meant the processing to equalize differences in users' weights and in sensors' sensitivities. The calibration was performed as follows. A push operation was applied to each element for 3 seconds a number of times, and the maximum and the minimum measurements were recorded. The difference between the maximum and the minimum values was multiplied by the value of a parameter we call a weight ratio. This value is subtracted from the maximum value. The result is used as the weight threshold. The $p_{threshold}$ can be expressed as

$$p_{threshold} = p_{max} - \{ Weight\ Ratio \times (p_{max} - p_{min}) \} \quad (1)$$

The other parameters used are the detection duration for a tap operation and that for a push operation.

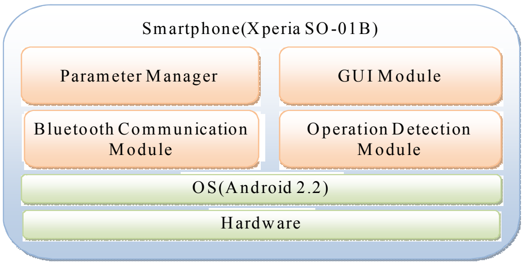


Figure 8. Configuration of the Operation Detection Application.

V. EXPERIMENT AND EVALUATION

A. *Detection Duration of a Tap Operation, that of a Push Operation, and Threshold Value*

Experiments were carried out using the experimental system. Parameters were set before starting the experiment. We measured changes in weights measured respectively by elements A, B and C in Fig. 5 when a tap operation and a push operation were respectively applied in order to determine the duration needed to detect an operation successfully. As show in Fig. 9, 85% to 90% of the measured weight data concentrated on either the range where the optical measurement value was between the maximum value and that minus 25% or the range where it was between the minimum value and that plus 25%. We found that the weight fell in one of these ranges only when an operation was applied, and therefore, it is possible to detect an operation by setting the weight threshold in the middle of these two ranges. When the threshold was calculated with the weight ratio at 0.7, it took 100 ms to 200 ms for the user to perform a tap operation, as shown in Fig. 10, and 600 ms to 1100 ms to perform a push operation, as shown in Fig. 11. Therefore, we decided that operations can be detected correctly if we set the detection duration for a tap operation to around 200 ms, and the detection duration for a push operation to slightly below 600 ms.

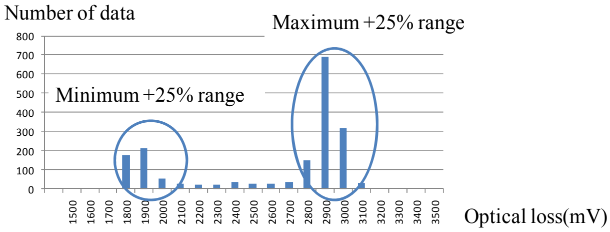


Figure 9. Polarization of Measured Values.

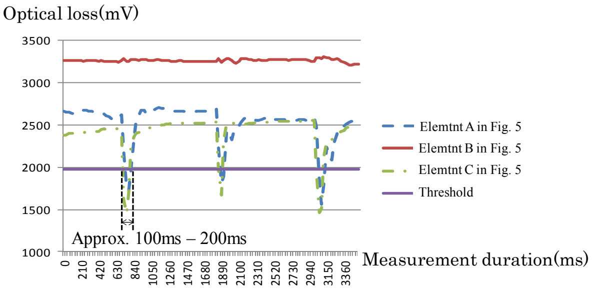


Figure 10. Measurement for a Tap Operation.

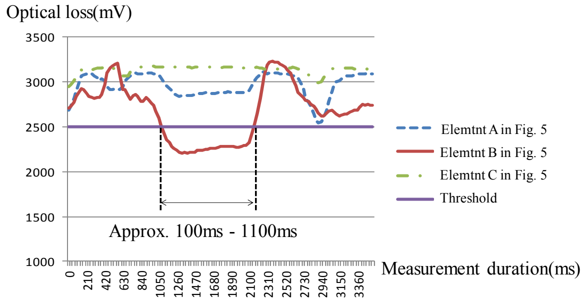


Figure 11. Measurement for a Push Operation.

B. *Experiment Conducted using the Experimental System*

We carried out an experiment to examine how the experimental system behaves. It was conducted in 4 steps:

Step 1: Set the parameters
Step 2: Establish a Bluetooth connection
Step 3: Set calibration and threshold values
Step 4: Select a function.

Figure 12 shows a test subject wearing the sensor shoe of the experimental system. Table II shows screenshots of the smartphone at each step, and the user's state that can be inferred from it.

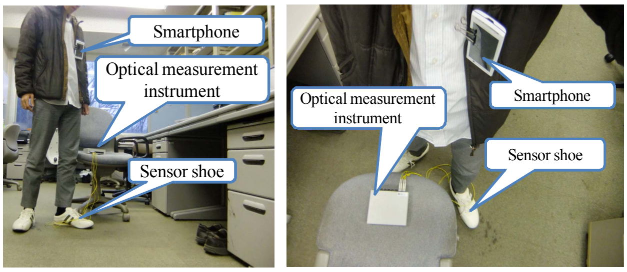


Figure 12. User wearing the Sensor Shoe.

C. *Evaluation of the Successful Command Identification Rate*

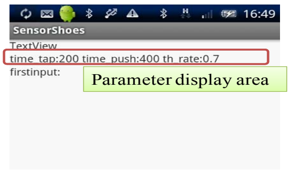
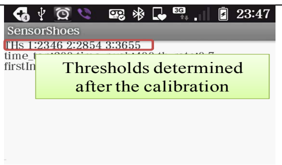
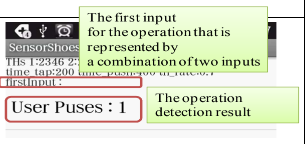
We had three test subjects. They learned how to operate the system for about 10 minutes before starting the experiment. They input the ten commands in sequence from command identifier 1 to 10. They tried these several times so that we could examine the probability at which the commands they intended to input were identified correctly. The parameter values used in this experiment were 200 ms for the tap operation detection duration, 400 ms for the push operation detection duration, and 0.7 for the weight ratio. The result of the experiment is shown in Table III. A screenshot of the experimental system taken during the experiment is shown in Fig. 13.

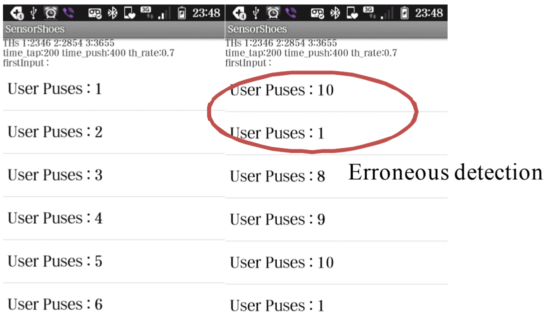
TABLE III. RESULT OF THE COMMAND IDENTIFICATION EXPERIMENT

Commands Item	1	2	3	4	5	6	7	8	9	10	Number of successful detection	Success rate
Subject 1	10	8	10	10	9	10	6	10	8	10	91	0.91
Subject 2	10	9	8	9	8	4	9	9	10	10	86	0.86
Subject 3	2	2	2	2	2	2	2	1	2	2	19	0.95
Avarage or Total	1.00	0.86	0.91	0.95	0.86	0.73	0.77	0.91	0.91	1.00	196	0.89

The three subjects conducted the experiment a total of 220 times, of which their input commands were identified correctly 196 times. The rate of successful identification for the three subjects ranged from 86% to 95%. The average rate was 89%.

TABLE II. SCREENSHOTS OF THE SMARTPHONE

Application State	Step1-1：Parameters setting	Step1-2：Display parameters	Step2：Establish a Bluetooth Connection
Screenshot	 16:51 SensorShoes Tap Detection Time 200 Push Detection Time 400 Threshold Rate 0.7	 16:49 SensorShoes TextView time_tap:200 time_push:400 th_rate:0.7 firstinput: Parameter display area	 16:49 SensorShoes Bluetoothのペア設定リクエスト Instrument ID 「BLUET_SOUK1」とペア設定するためのPINを入力してください（0000か1234を試してください）。 0000 OK キャンセル
User State	Before putting on the shoe-embedded interface	Shoe-embedded interface worn	
Application State	Step3-1：Calibrate and set thresholds	Step3-2：Display the thresholds	Step4：Selects a function
Screenshot	 please push element no.1	 23:47 SensorShoes THs 1:2346 2:2854 3:3655 Thresholds determined after the calibration	 The first input for the operation that is represented by a combination of two inputs SensorShoes THs 1:2346 2: time_tap:200 firstInput : User Puses : 1 The operation detection result
User State	Several attempts of a push operation		Selects a function



SensorShoes
THs 1:2346 2:2854 3:3655
time_tap:200 time_push:400 th_rate:0.7
firstInput :
User Puses : 1
User Puses : 2
User Puses : 3
User Puses : 4
User Puses : 5
User Puses : 6

SensorShoes
THs 1:2346 2:2854 3:3655
time_tap:200 time_push:400 th_rate:0.7
firstInput :
User Puses : 10
User Puses : 1
User Puses : 8
User Puses : 9
User Puses : 10
User Puses : 1

Figure 13. Screenshot taken in the Experiment.
(Each subject input commands in sequence from 1 to 10. An erroneous identification occurred in the seventh input attempt)

VI. CONCLUTION AND FUTURE WORK

We have focused on foot as a user interface that provides high mobility, a high command identification rate, and ability to complement information, hence the proposal of a shoe-embedded interface. This interface was implemented by attaching sensors at three points on the insole of a shoe. These points were selected based on the characteristics of the shape of a human foot. Commands were defined by a combination of two operations: push and tap. We have devised a method of indicating one of 10 alternative commands by a single foot operation. We have developed an experimental system that implemented the proposed method, and used it to evaluate the probability at which input commands are identified correctly. In our experiment with three subjects, the average rate of successful identification was 89%.

In the future, it will be necessary to examine the rate of correct identification and possible occurrences of system failures in cases where the user is running or walking. It is also necessary to study how providing recommended options to users can increase the amount of information the user can input without impairing usability.

REFERENCES

[1] Masafumi Nishimura and Gakuto Kurata, “Recent Advances and Possibilities of Innovation in Speech Interface Technology(<Special Feature>Toward Developing Practical Automatic Speech Recognition Technology),” Journal of Information Processing Society of Japan, vol. 51, no. 11, pp. 1434-1439, Nov. 2010.

[2] Kiyohiko Abe, Mikio Ohuchi, Shoichi Ohi, and Minoru Ohyama, “Eye-gaze Input System Based on the Limbus Tracking Method Using Image Analysis,” vol. 57, no. 10, pp. 1354-1360, Oct. 2003.

[3] Yuki Ebina, Shuhei Nishida, and Masahiro Nakagawa, “On the Chaos and Fractal Properties in EEG and NIRS Signals and their Applications to Robotics,” Technical Report on NLP, The Institute of Electronics, Information and Communication Engineers, vol. 107, no. 267, pp. 41-46, Oct. 2007.

[4] H.-K.J. Kuo and Lee Chin-Hui, “Discriminative training of natural language call routers,” Speech and Audio Processing, IEEE Transactions on , vol.11, no.1, pp. 24- 35, Jan 2003.

[5] K. Tsukada and M. Yasumura, “Ubi-Finger: Gesture Input Device for Mobile Use,” Proceedings of APCHI 2002, Vol. 1, pp.388-400, 2002.

[6] M. Bacchiani, F. Beaufays, J. Schalkwyk, M. Schuster, and B. Strope, “Deploying GOOG-411: Early lessons in data, measurement, and testing,” Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on, vol., no., pp.5260-5263, March 31 2008-April 4 2008.

[7] Michiko Nishiyama, “Unconstrained measurement of respiration rhythm focusing on weight movement using hetero-core fiber optic sensors,” IEICE Technical report Vol. 109, No. 429, pp. 49-52, Feb. 2010.

[8] Kazuhiro Watanabe, “Macrobending Characteristics of a Hetero-Core Splice Fiber Optic Sensor for Displacement and Liquid Detection,” IEICE Trans. Electron., Special Issue on Optical Fiber Sensors, vol.E83-C, no.3, pp.309-314, 2000.

[9] Masako Sonobe, Michiko Nishiyama, and Kazuhiro Watanabe, “Unconstrained Pulse Monitoring for On-Site Usage Using a Hetero-core Fiber Optic Nerve,” 2009 International Symposium on Smart Sensing and Actuator System, Proceedings of ISSS’09,pp.48-51, 2009.

[10] Mitsuo Miyamoto, Michiko Nishiyama, and Kazuhiro Watanabe, “Hetero-core Fiber Optic Nerve Weight Sensors for Unconstrained Respiration Monitoring during Sleep,” 2009 International Symposium on Smart Sensing and Actuator System, Proceedings of ISSS’09, pp52-55, 2009.

[11] Kaname Takaochi, Mitsuaki Shimono, Michiko Nishiyama, Kazuhiro Watanabe, and Kazumasa Takami, “An information communication system operated by a sole placed on a hetero-core optical fiber sensor mat,” IET Conf. Pub. 2010, 25 (2010), DOI:10.1049/cp.2010.

[12] Shinichi Nose, Michiko Nishiyama, Kazuhiro Watanabe, and Kazumasa Takami, “A Personal Web Service based on a User-Friendly Personal Identification Method in a Ubiquitous Environment,” in proceedings of PDPTA'07-The 2007 International Conference on Parallel and Distributed Processing Techniques and Applications. 25-28 June 2007.

[13] Shinichi Nose, Mituaki Shimono, Michiko Nishiyama, Tetuya Kon, Kazuhiro Watanabe, and Kazumasa Takami, “Personal identification based on sole pressure distribution using a hetero-core optical fiber sensor network for personal web services,” in proceedings of 2009 IEEE Congress on Services (SERVICES 2009) July 6-10, 2009.

[14] ISO 9241-11:1998 Ergonomic requirements for office work with visual display terminals (VDTs) -- Part 11: Guidance on usability.

[15] Timo Jokela, Netta Iivari, Juha Matero, and Minna Karukka, “The standard of user-centered design and the standard definition of usability: analyzing ISO 13407 against ISO 9241-11,” CLIHC '03 Proceedings of the Latin American conference on Human-computer interaction, ACM New York, NY, USA, 2003.

[16] Alain Abran, Adel Khelifi, Witold Suryn, and Ahmed Seffah, “Usability Meanings and Interpretations in ISO Standards,” SOFTWARE QUALITY JOURNAL, Volume 11, Number 4, 325-338, DOI: 10.1023/A:1025869312943, 2003.

[17] Yasushi Ishikawa, “UI Design and Usability : Subjects of Speech Interface, ” Technical Report on SLP, Information Processing Society of Japan, vol. 2007, no. 103, pp. 35-40, Oct. 2007

[18] Yuji Noda, “All about plantar arch,” A Body on Viewpoint of Sole, pp. 70-77, Kohdansya Ltd., Tokyo, 1998.

[19] G. A. Miller, “The magical number seven, plus or minus two: Some limits on our capacity for processing information,” Psychological Review, vol. 63, no. 2, pp. 81-97, March. 1956.

[20] G. A. Miller, “The cognitive revolution: a historical perspective, ” Trends in Cognitive Sciences 7 (3): 141–4. doi:10.1016/S1364-6613(03)00029-9. PMID 12639696, 2003.

[21] Ed Burnette, “Hello, Android: Introducing Google's Mobile Development Platform, ” 2009.

[22] Margaret Butler, “Android: Changing the Mobile Landscape,” IEEE Pervasive Computing, vol. 10, no. 1, pp. 4-7, Jan.-Mar. 2011.