An Intuitive Multi-touch Interface Using Finger Posture for Operating 3D Objects

Shingo Baba, Katsuhiko Kaji, and Nobuo Kawaguchi
Graduate School of Engineering, Nagoya University
Furo-cho, Chikusa-ku, Nagoya, Aichi, Japan

Toshiyuki Hagiya and Tsuneo Kato
User Interface Laboratory, KDDI R&D Laboratories, Inc.
2-1-15 Ohara, Fujimino-shi, Saitama, Japan

***Abstract*—Recently multi-touch user interfaces have been widely used. Multi-touch UI is becoming more and more important. We propose a new comfortable, multi-touch user interface which considers finger posture. It allows us to increase our input functions with smaller number of fingers. We collected the finger feature values, such as coordinates, angles, areas, and their differences of previous frame using FTIR touch-panel. We classified the finger postures using support vector machine, decision tree learning and multilayer perceptron. The maximum rate of correctly classified instance was 99.82%. Finally, we introduce the trial application of operating 3D objects using classified gestures.**

***Index Terms*—multi-touch, user interface, frustrated total internal reflection, machine learning**

I. Introduction

With development of computers, computational systems are used in many fields. Recently, multi-touch user interface have been widely used. Many devices have multi-touch display and multi-touch user interfaces. We can operate these devices to use some gestures, because the multi-touch displays can detect the multiple points. Recently, the share of smartphones, tablet devices and personal computer with multi-touch display is increasing. Due to this tendency, the importance of multi-touch user interface will increase.

We think of operating 3D object on touch-panel display. It requires large input functions. For example, select, move, rotate, expand any object and change the view point, and so on. In conventional user interface, we can increase input functions to combine gestures with number of the fingers. For example, if there are 5 variations per number of the finger, one finger makes 5 variations, and two fingers makes 10 variations and more in addition to some two-handed gestures. In this way, although we can get large input functions, it is too complicated to operate objects because users have to use many fingers to use various functions.

The problem we really want to focus on is to solve the complexity of operation. To solve this problem, we propose the solution that "By using recognition of finger posture, we can assign different roles to each posture". This allows us to avoid the complexity because we can operate object with less number of the finger, such as 1 or 2. Conventional UI uses coordinates, gesture and the number of fingers as features of fingers. However, there is not enough information for us to distinct the finger posture. For example, although we cannot estimate finger direction from only one point, if we could get a relationship of finger features, we can presume the finger posture from the relationship. Therefore, to detect finger posture, we need to find the relationship of features on the touch-panel, such as coordinates, angles, areas, and their differences of previous frame. If the finger posture can be recognized, we can apply for the different roles to each posture. In this way, we can reduce the complexity of inputs various functions, and we can get large gesture functions. For instance, to move the object is just to pinch with one hand and to change the view point is just to pinch with two hands. The remarkable point is that the proposed UI recognizes the different pair of fingers as different finger postures. In this way, gesture that considers the finger posture enables us to input various functions into devices.

In this research, we define the finger posture as the state of fingers. The examples of finger posture (Figure 1) are shown bellow.

- Which fingers do users use?
- How strong do users press the display?
- How fast do users' fingers move?
- What direction do user's fingers point?
- What kind of pairs do users use?

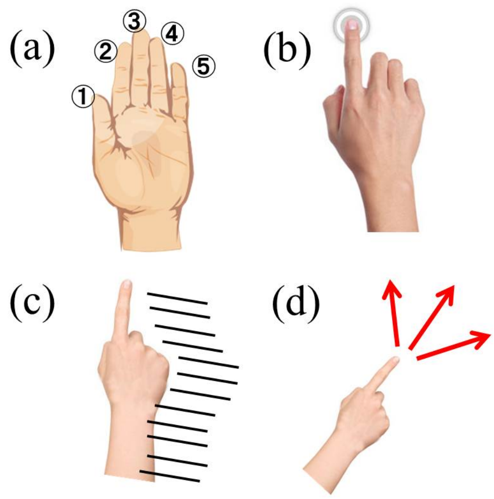

Fig. 1. The list of finger posture (a)Which fingers do users use?, (b)How strong do users press the display?, (c)How fast do users' fingers move?, (d)What direction do user's fingers point?

Some feature values are not measured from electric capacitance type touch-panel which is widely used for smartphones and tablet devices. Therefore we need to use the touch-panel system to be able to measure these feature values.

The purpose of this paper is to create a new comfortable, intuitive, multi-touch user interface by positively using recognition of finger posture. We used FTIR (Frustrated Total Internal Reflection)[1] touch-panel to get finger features on the touch-panel screen.

The process to predict the finger posture consists of two things. First, we collect the information about finger posture to do some gestures using FTIR touch-panel. Next we classify the gestures using machine learning[5]. We use the support vector machine, decision tree learning and multilayer perceptron as the tools of classification.

The composition of this paper is as follows. Chapter 2 provides related work. Next in Chapter 3 describes the FTIR touch-panel and its principle. In Chapter 4, we provide the information of contact area collected from FTIR touch-panel and In Chapter 5, we describe the evaluation of estimation method of finger angle. In Chapter 6, we provide the method of classification using machine learning. In Chapter 7, we provide the trial application for operating 3D object using classified gesture.

II. RELATED WORK

We will describe the researches of sensing method of multi-touch display and gesture recognition method. Multi-touch has been implemented in several different ways. The technologies of multi-touch mainly categorized into 4 methods, capacitive, resistive, optical, and wave technologies. Now focus on capacitive and optical technologies because they are popularly used for multi-touch displays.

In capacitive type, Rekimoto proposed “Smartskin”[2], the sensor architecture for making interactive surface that are sensitive to human hand and finger gestures. The sensor recognizes multiple hand positions, their shapes and distance between the hands and the surface by using capacitive sensing and mesh-shaped antenna. For SmartSkin, the shape of fingers and the number of fingers can be used for the functions to operate something. In contrast to camera-based gesture recognition systems, all sensor elements can be integrated within the surface, and this method does not suffer from lighting and occlusion problems.

In optical type, the methods that a camera detects diffused infrared at the finger contact areas are mainly utilized. For example, FTIR use the diffuse infrared in the transparent medium, such as acrylic panel. FTIR is constructed at a low cost and is easy to expand to widely size. The one of the optical method, Sakamoto proposed the “Winkle Surface”[3], the interface using the winkles of elastic material on the touch-panel. Winkle Surface recognizes the finger gestures, such as push, slide and twist. Input gestures are recognized by using the likelihood function of finger features, the intensity of luminosity value of finger contact areas.

In addition, we need to understand the relationship of finger features to make a comfortable multi-touch user interface. For that, we have to clear the relationships among the finger features. Kawai[7] revealed that the relationship between the intensity of the fingertip force and the finger contact area. We need to consider some relationships of finger features. Based on the knowledge from these researches, we assume that to collect the finger feature using FTIR touch-panel is best way in our research. Then, we aim to create a new multi-touch user interface utilizing the finger features by conducting the classification of finger posture.

III. FTIR TOUCH-PANEL

In this section, we describe FTIR touch-panel. We used an FTIR touch-panel to collect the finger information such as contact point, area, and angle. Why did we use the FTIR touch-panel? The reason is that we need to find the relationship of features on the touch-panel and the information is not measured from other multi-touch display. For example, a resistive membrane touch screen and an electric capacitance touch screen cannot get the information of finger posture because they detect contact area not as an image but as a coordinate. Therefore the other touch-panels were unsuitable, then we choose the FTIR touch-panel. We made the FTIR touch-panel system based on principle of FTIR. Figure 2 indicates the FTIR touch-panel system. Figure 3 indicates the captured raw image of infrared camera, we can make sure that infrared is reflect only at the finger contact areas.

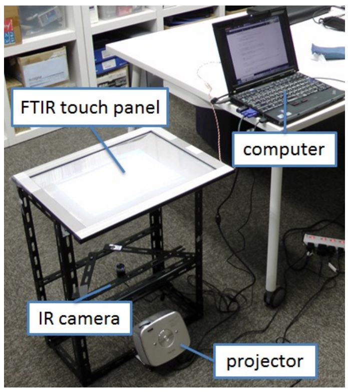

Fig. 2. FTIR touch-panel system

A. *Principle of FTIR touch-panel*

In this section, we describe a principle of FTIR touch-panel. Figure 4 shows the principle of FTIR touch-panel. FTIR touch-panel utilizes the phenomenon that total internal reflection occurs in transparent medium is diffused when a medium boundary contacts with other object. Total internal reflection usually happens when a ray of light strikes a medium boundary at an angle larger than a particular critical angle with respect to the normal to the surface. Under “no object on the acrylic panel”, total internal reflection of an infrared occurs on the whole panel. On the other hand, when there is an object on

Fig. 3. Captured raw image

the touch-panel, an infrared ray diffuse at the object contact areas. The infrared which entered from side of acrylic panel is reflected diffusely only at finger contact area and it is reflected totally at the other area.

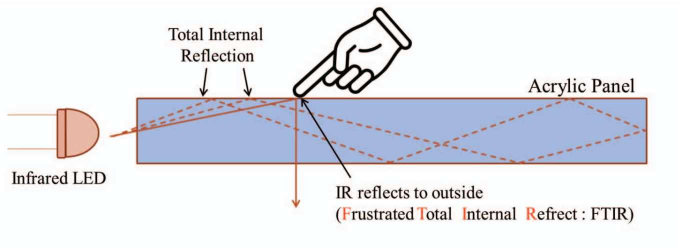

Fig. 4. Principle of FTIR panel screen

B. Structre of FTIR touch-panel

FTIR touch-panel system mainly consists of acrylic panel (300mm $\times$ 400mm), infrared LEDs, infrared detecting camera, projector, tracing paper as a diffuser, and computer. 20 infrared-LEDs (Toshiba TLN231) are placed at a 40mm interval on each long side of acrylic panel. The luminous intensity can be adjusted with volume controller. We use Point Frey FireFly MV (FMVU-03MTM-CS) as a camera to capture infrared and infrared optical filter (FIJIFILM IR-78) which can cut 50% of light whose wavelengths under 780nm. We were able to get enough cutoff of visible spectrum. In addition, we attached a piece of tracing paper to take a roll of the display. To project information on the touch-panel screen, we use a mobile projector (Kaga Electronics Co., Ltd.tKG-PL021).

IV. Information of Contact Area Collected from FTIR Touch-Panel

In this section, we describe the process of measure the finger features. To estimate the finger posture, we need to collect the finger feature values on the touch-panel, such as coordinates, angles, areas, and their differences of previous frame using FTIR touch-panel. We conducted the image processing[6] to collect these feature values.

First, we captured the reflected infrared from the acrylic panel, and got the source image (Fig.5(a)). The source image had some noises. Secondly we performed background subtraction to reduce the noises and smoothing using Gaussian filter, and got the smoothed image (Fig.5(b)). These processes are important to realize robust multi-touch system. Thirdly, in order to get the contours of finger contact area, we converted the smoothed image to greyscale image and then converted the greyscale images to the binary format image (Fig.5(c)). This process is a preparation to find the contours of contact area in the next step. Finally, we fitted an ellipse to each contours (Fig.5(d)). We treated the ellipse parameters as finger features. In this way, we can get the finger feature values from FTIR touch-panel.

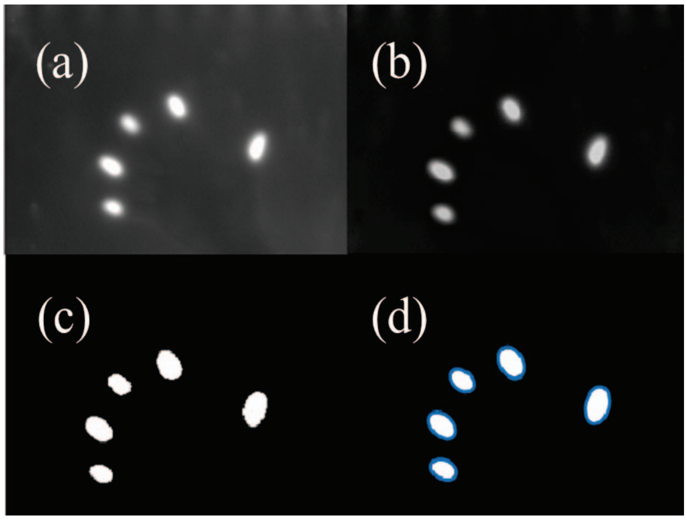

Fig. 5. Process images to measure finger feature values. (a)Source image, (b)Background subtracted image, (c)Binary format image, (d) Image of fitting ellipse to each contour

Then, we describe the feature values of finger collected from FTIR touch-panel. Figure 6 shows the feature values. We define the center of ellipse as the finger position (x_{ni}, y_{ni}), the area enclosed by contour as a contact area S_{ni}, where index n and i indicate a number of frame and a finger id respectively. The angle of finger is defined as the angle of major axis of ellipse with respect to horizontal axis. We define $\Delta x_{ni}, \Delta y_{ni}, \Delta S_{ni}, \Delta \theta_{ni}$ are the difference of previous frame.

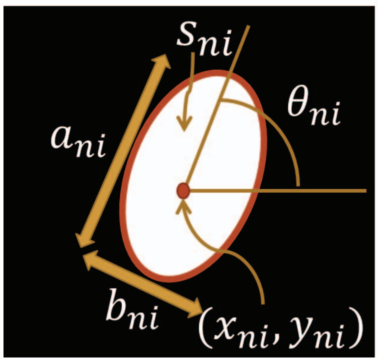

Fig. 6. Feature values of contact area

V. Evaluation of Estimation Method of Finger Feature

We have been able to measure the finger feature values. Then we conducted an experiment to evaluate the estimation method of finger features. Now we focus on "the angle of finger" because we assume it is an important role of to detect the finger posture. For example, we can get only coordinates of fingers it is hard to predict the kind of fingers. However if we could get finger angles besides the coordinates, we can compare the angles and structure of hand, we could make the prediction of finger posture easier. Therefore, we focus on "the angle of finger".

A. Process of measurement

The procedure of evaluation is as follows. We let the subjects to stand in front of FTIR touch-panel system and touch 3 times the target lines indicated on the touch-panel with their right index finger. The target lines included with 5 lines (-60, -30, 0, 30, 60 [degrees]) and they appeared counterclockwise on the touch-panel from -60 to 60 in turn. The measured values and angle of target lines are recorded together. First, we conducted this measurement without training. Next, after we let the subjects to train how to detect the finger features about for about 2 minutes, we repeat the same measurement. From this measurement, we evaluated the effects of the practice on the detection accuracy of finger angle. The number of subjects was 12 included of 10 men and 2 women in our laboratory.

B. Result of measurement

Figure 7 shows histogram that indicates the differences between measured angles and target angles. In Figure 7, the horizontal axis indicate the difference and vertical axis indicated the frequency. In this research, we defined the measured value whose difference between the target angle is greater than 20 degrees as the error value. The error rate is a proportion of number of error values to all measured values. The total number of each measured instances is 180 (12 subjects, 5 target angles, 3 times). The error rate of whole subjects without training was 18.9%, and one of after training was 6.1%, approximately 12.8% of error rate was decreased.

Figure 8 shows the error rate of each subject. Some subjects had a high rate of error, others had no error. The subject D, E had 60%, 47% of error rate and other most subjects had some errors before the training. On the other hand, most of subjects' error rate were decreased after the training. Therefore, we found that the error rate is improved with some training.

Those errors may be caused of disagreement between the direction of finger and the direction of the long axis of ellipse. The error happened when user had touched the panel with their finger standing. Even if the direction of the major axis of ellipse corresponds with the direction of finger, the finger is standing on the panel, the direction of major axis not always corresponds the direction of finger. Because the fitted ellipse looks like a circle, the direction necessarily agrees with the direction of finger. This problem must be solved. An idea to resolve it is to introduce the parameter that decides whether

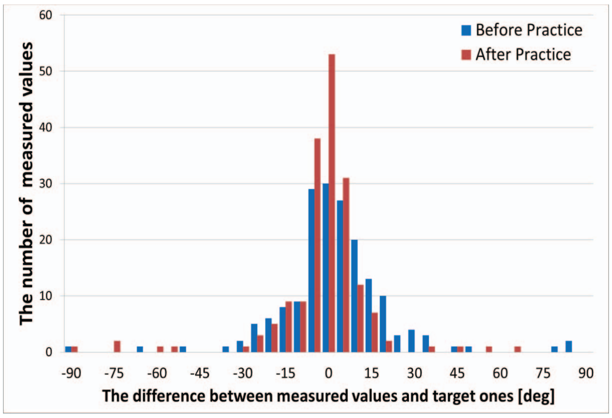

Fig. 7. The histogram of the difference between measured values and target ones

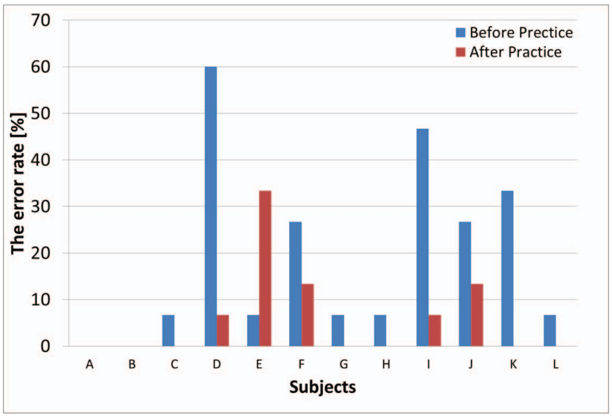

Fig. 8. The number of errors for each subject

current finger angle should be used according to contact area. This parameter utilizes the phenomenon that the contact area tends to become larger in the case with fingers lying and smaller the case with user's finger standing.

C. Tendency about the directions of finger

We found the tendency about the directions of finger on the touch-panel. Figure 9 displays the average of measured angles. Each average angles had a little differences before or after practice. For -60, -30, 30 and 60 degrees, the absolute values of the averages were smaller than values of the target angle because of the direction of subjects and easiness of moving the fingers. Figure 10 indicates the variance of measured angles. For -30, 0, 30 degrees data, the variances were decreased. On the other hand, for the target angle of -60 and 60 degrees had a little change and kept the error rate at higher values. From these results, we can conclude that the measured angle have some differences according to the target angle, for example, in figure 9, for angle 60 degrees, there is a little differences between target and measured. Namely the finger angle was recognized with a little bias to the direction of the vertical direction with respect to the direction of subjects.

Fig. 9. The average of measured angles

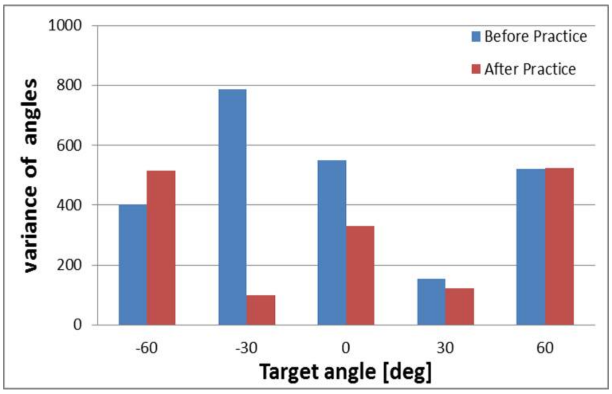

Fig. 10. The variance of measured angles

VI. Gesture recognition using features of finger

We have described the finger features to predict the finger posture. In this section, we describe the way of finger posture recognition using finger features. We focused on pinch gesture for recognition as a basic gesture using multiple fingers. We classified gestures as either one-handed pinch gesture or two-handed pinch gesture. One-handed pinch gesture use sum and index finger of right hand, and two-handed pinch gesture use the index fingers of both hands.

Fig. 11. (a)One-handed pinch, (b)Two-handed pinch.

A. *Classification methods*

We describe the classification methods. First of all, we describe an input vector. We collect the finger feature values using FTIR touch-panel and make a training set that consists of a set of N the input vectors $\{\boldsymbol{x_1}, \boldsymbol{x_2}, ..., \boldsymbol{x_n}, ..., \boldsymbol{x_N}\}$, where N is a number of frames. Each $\boldsymbol{x_n}$ has the gesture id g_n, process time t_n, finger id f_{ni}, finger coordinates (x_{ni}, y_{ni}), finger angle θ_{ni}, finger contact area S_{ni}, the length of ellipse major and minor axes (a_{ni}, b_{ni}), each differences of coordinates, angle, area $(\Delta x_{ni}, \Delta y_{ni}, \Delta\theta_{ni}, \Delta S_{ni})$. The index i $(i = 1, 2, ...)$ indicates the number of fingers. The finger id is an index to indicate one-handed, two-handed and more handed gestures. The finger id is assigned to each contact areas according to their position of horizontal axis. The further position, the larger number of id is assigned.

We use machine learning to classify the gestures. We use support vector machine, decision tree and multilayer perceptron.

B. *The method to classify the gestures*

The method to classify the gestures is described bellow. We collect the data about each gesture on the FTIR touch-panel. We conduct the classification of gestures using SMO (support vector machine algorithm in Weka), j48 decision tree, and multilayer perceptron of Weka[4]. Weka is a collection of machine learning algorithms for data mining tasks.

We used 3 kinds of machine learning methods to classify. We introduce them bellow. A support Vector Machine (SVM) is a supervised learning models with associated learning algorithms that analyze data and recognize patterns, used for classification and regression analysis. The basic SVM takes a set of input data and predicts, for each given input, which of two possible classes forms the output. SVM is basically two classification method, but it is capable of multi-classification.

Decision tree learning is a method popularly used in data mining. A tree can be learned by splitting the input data set into subsets based on an attribute value test. This process is repeated on each derived subset in a recursive way. The recursion is completed when the subset at a node has all the same value of the target variable, or when splitting no longer adds value to the predictions.

A multilayer perceptron (MLP) is a feedforward artificial neural network model that maps sets of input data onto a set of appropriate output. MLP utilizes a supervised learning technique called backpropagation for training the network. MLP is a modification of the standard linear perceptron and can distinguish data that is not linearly separable.

C. *The process and result of classification*

We describe the process of the classification and the result. First, to collect the data, we let subjects to conduct each gesture 50 times. The image capture rate is about 16 frames per second. Each gesture operation contained several pairs of finger feature values. Next, we perform 10 fold cross-validation. Finally, We calculate the rate of correctly classified instances using each machine learning methods.

Figure 12 shows the rate of correctly classified instances for each classification methods. For each subject, the rates of correctly classified instances were 92.0%, 85.9% and 97.7%. The maximum rate of correctly classified instance was 99.82%

by using MLP. On the other hand, the worst rate of correctly classified instance was 91.84% by using SVM. We had good classification of the gestures of one-handed pinch gesture and two-handed pinch gesture using MLP.

	MLP	J48 Decision Tree	SVM
subject a	99.92%	99.03%	91.97%
subject b	99.84%	98.73%	97.69%
subject c	99.69%	98.47%	85.87%
average rate	99.82%	98.74%	91.84%

Fig. 12. The rate of correctly classified instances for each classification methods.

VII. Application Implementation

We made a simple application to operate 3D objects using gesture that had been classified by machine learning. In this application, we could realize dynamically switching operation of objects according to one-handed or two-handed gesture. For instance, bending the object is just to pinch with one hand and changing the view point is just to pinch with two hands. In this way, we can increase our gesture functions by applying the different roles to the different pairs of fingers.

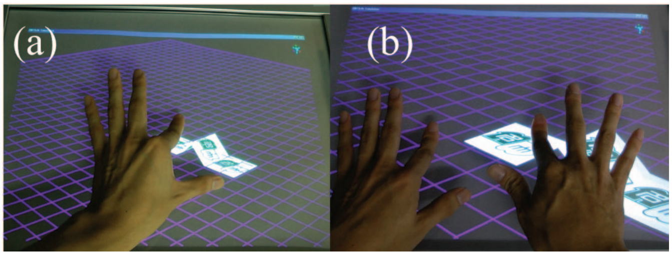

Fig. 13. (a):One-handed pinch for bending object. (b)Two-handed pinch for changing veiw point.

VIII. Conclusion and Future works

We proposed a multi-touch UI using finger feature values. We found that if we train the subjects to use the touch-panel, we can get the better recognition rate. And the tendency about the directions of finger on the touch-panel was confirmed. We conducted the classification of pinch gestures with different pair of fingers using machine learning. We also made a trial application using classified gestures.

We are going to apply for other gestures. We need to collect more subjects' data and optimize the classification method to detect more finger postures.

References

[1] JeffersonY. Han. Low-cost multi-touch sensing through frustrated total internalreflection. In Proceedings of the 18th annual ACM symposium on User interfacesoftware and technology, UIST' 05, pp. 115-118, New York, NY, USA, 2005.

[2] Jun Rekimoto. Smartskin: an infrastructure for freehand manipulation on interactivesurfaces.In Proceedings of the SIGCHI conference on Human factors incomputing systems: Changing our world, changing ourselves, CHI ' 02, pp. 113-120, New York, NY, USA, 2002.

[3] Y. Sakamoto, et al., "Constructing an Elastic Touch Panel with Embedded IR-LEDs using Silicone Rubber", Proceedings of IADIS International Conference Interfaces and Human Computer Interaction 2012 (IHCI 2012), July 21-23, 2012.

[4] Weka 3 -Data Mining with Open Source Machine Learning Software in Java: 11 Nov. 2012. http://www.cs.waikato.ac.nz/ml/weka/

[5] Stephen Marsland. Machine Learning:An Algorithmic Perspective (Chapman & Hall/Crc Machine Learning &Patrtern Recognition). Chapman and Hall/CRC. 2009.

[6] Gary B., Adrian K., trans. Koichi Matsuda. *Learning OpenCV*. O'Reilly Japan. 2009.

[7] Satoru K., Hisayuki K., Hideo N., Contact-Surface-Area Changes in Relation to Finger Force. Tezukayama College Bulletin. 1994.