

BackHand: Sensing Hand Gestures via Back of the Hand

Jhe-Wei Lin Chiuan Wang Yi-Yao Huang Kuan-Ting Chou
Hsuan-Yu Chen Wei-Luan Tseng Mike Y. Chen
Mobile HCI Research Lab, National Taiwan University
{r02944056, r03922001, b02901042, b01902094, b01902096}@ntu.edu.tw,
{r03944003, mikechen}@csie.ntu.edu.tw

ABSTRACT

In this paper, we explore using the back of hands for sensing hand gestures, which interferes less than glove-based approaches and provides better recognition than sensing at wrists and forearms. Our prototype, BackHand, uses an array of strain gauge sensors affixed to the back of hands, and applies machine learning techniques to recognize a variety of hand gestures.

We conducted a user study with 10 participants to better understand gesture recognition accuracy and the effects of sensing locations. Results showed that sensor reading patterns differ significantly across users, but are consistent for the same user. The leave-one-user-out accuracy is low at an average of 27.4%, but reaches 95.8% average accuracy for 16 popular hand gestures when personalized for each participant. The most promising location spans the 1/8~1/4 area between the metacarpophalangeal joints (MCP, the knuckles between the hand and fingers) and the head of ulna (tip of the wrist).

Author Keywords

Gesture recognition; wearable interface; back of the hand; hand gesture interface; strain gauge; machine learning; gestural interaction

ACM Classification Keywords

H.5.2 Information Interfaces and Presentation: User Interfaces

INTRODUCTION

Human-hand gesture is a fundamental part of human communication and a natural human-computer interface. Hand gesture recognition enables humans to be understood by the computer and interact naturally with lower learning efforts and lower cognitive load.

Many projects have explored wearable sensors to detect hand gestures, and can be classified into the following three approaches as shown in Figure 2: (1) sensing finger movement, (2) sensing physical properties of the wrist, (3) sensing physical properties of the forearm.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [Permissions@acm.org](http://Permissions.acm.org).

UIST'15, November 08-11, 2015, Charlotte, NC, USA.

Copyright ©2015 ACM. ISBN 978-1-4503-3779-3/15/11\$15.00
DOI: <http://dx.doi.org/10.1145/2807442.2807462>

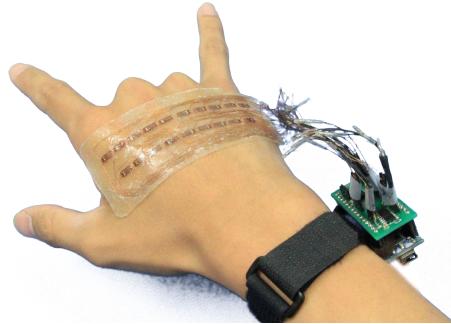


Figure 1. BackHand is a prototype that senses physical properties on the back of hand for hand gesture detection.

The first classification is to measure the overall flexion of the fingers. Data glove [8] places sensors directly on finger joints. It accurately captures finger movement, but the glove form factor reduces tactile feedback and interferes with finger movement. Digits [11] uses wrist-worn cameras and a vision-based approach to track fingers at high accuracy, but is constrained by occlusion and requires significant camera elevation from the wrist.

The second classification is the measurement of physical properties from the wrist. Capacitive sensing [16], force sensing approaches [6], and vision [9] have been used to sense changes in the physical properties of the wrist. Because of the reduced movement of the tendons compared to finger sensing, these approaches provide significantly lower accuracy even with a small number of gestures (2, 5, and 8 gestures respectively).

The third classification uses forearm sensing to extract movements and gestures of a user's finger. Arm-based sensing

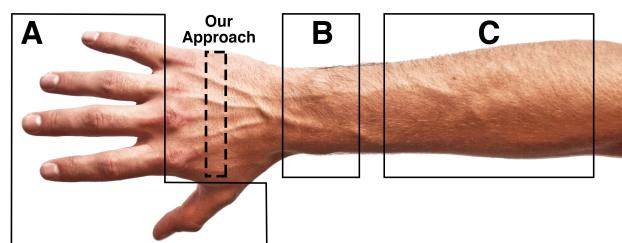


Figure 2. The three common body parts as sources of a sensor's input signal for a hand gesture interface include (A) the fingers, (B) the wrist, and (C) the arm.

measures electromyography (EMG) caused by hand gestures [3, 17]. It senses full hand movements (eg. clenched fist and splayed fingers) better than finer gestures such as moving a single finger.

Although the old colloquialism “like the back of your hand” suggests, we are not necessarily familiar with our own bodies. As shown in Figure 2, the back of the hand is closer to fingers, and the movement of tendons can be observed and sensed more clearly than at the wrists and forearm. Sensors placed at this location also do not interfere with finger movement nor reduce the sense of touch compared to glove-based approaches. In this work, we explore sensing the back of hands for hand gesture detection.

Our prototype, BackHand (Figure 1), uses a total of 19 strain sensors in a 2-row configuration to measure the physical changes on the back of hands at a sampling rate of 75 Hz. A customized Arduino Nano shield circuit board is built for sensor signal conditioning. To better understand the areas that provide better gestural recognition accuracy, we conducted a user study to collect data by affixing our prototype to 8 specified rows evenly spaced from the metacarpophalangeal joints (MCP, the knuckles between the hand and fingers) to the head of ulna (tip of the wrist) while performing a number of 16 gestures. From the results of machine learning, the most promising location on the back of hand is found on the second and third row from the MCP joints. While an gestural recognition accuracy rate of 27.4% is low in a leave-one-user-out cross validation, the accuracy rate is 95.8% in a 10-fold cross validation. A visualization shown as heat maps assists our understanding towards the difference of sensor reading patterns among individuals and among various gestures.

Contributions of this paper are:

- We explored a new signal source, tendon activities on the back of hand, for hand gesture recognition.
- We developed a prototype using strain sensors that can sense a large variety of gestures per user at higher accuracy than wrist-based approaches.
- We identified the most promising location in row configuration on the back of hand to be near the metacarpophalangeal (MCP) joints.

RELATED WORK

The growing number of a new interaction modality: body-worn hand gesture interfaces which senses the human-hand gestures as a wearable platform has led to attempts in categorization of those. One of the taxonomies of body-worn hand gesture interfaces is based on the source of the sensor’s input signal which can be classified into three categories.

Finger-based Hand Gesture Interfaces

To gather high accuracy information of hand movements, a common technique is to measure the overall flexion of the fingers. One approach is to instrument the hand with a data glove-like hand gesture interface [8, 12, 15, 18] which is equipped with a number of sensors that capture physical data

about hand position, orientation, and bending angles of fingers. For instance, the CyberGlove [10] is currently considered one of the most accurate glove systems for 3D modeling, robot control, and so forth [4, 13], and the 5DT Glove [2] is famous in virtual reality. Although most of these finger-based systems provide high accuracy in measuring the degrees-of-freedom (DoF) of human hands, they suffered from several drawbacks. Major limitations originated from the cloth support and the mechanical structure [8] , which acted as a constraint on the users hand. For example, they have to be worn on the fingers to measure the flexion of the fingers. Such complications may not only reduce dexterity of the hand but blunt tactile sensitivity. Additionally, these systems can be complex, expensive, and cumbersome for users.

To overcome the discomfort of data glove, another approach uses wearable cameras to recognize finger gesture. Vision-based systems can be worn on some parts of the body such as Digits [11] and another system [19] which are wrist-worn sensors based on 3D infrared camera and time-of-flight (TOF) camera respectively, SixSense[14] which senses colored markers on the fingertips through wearing a lanyard necklace with a camera. By recovering 3D articulated hand poses, they provide high-fidelity sensing without wearing the glove. However, because of the fundamental limitation of cameras, the visible occlusion resulting from crossing finger, a significant elevation from the wrist, heavy computation requirement of 3D reconstruction technology, large power consumption and expensive cost are problematic for these vision-based systems.

Wrist-based Hand Gesture Interfaces

From another approach, a number of lightweight systems measuring physical properties from the wrist to detect gestures have been proposed. Although wrist-based hand gesture interfaces are much rare compared to finger-based hand gesture interfaces, they are the most friendly and acceptable interfaces due to their placement.

Instead of providing high-accuracy sensing, GestureWrist [16] uses capacitive sensors to recognize a small set of gestures. WristFlex [6] uses an array of force sensitive resistors (FSRs) supporting 5 gestures at an accuracy rate greater than 80%. Another system [9] places infrared transmitters and receivers around the wrist to recognize hand gestures. Hambone [7] employs bio-acoustic method performing hand motions through the wristband enhanced with piezoelectric sensors, supporting classification accuracies around 90% for 4 gestures. While all these devices overcome some of the limitations of the data glove-like systems (e.g., most have no cloth), they appear to be unsuitable for applications that require high accuracy measurement and the diversity of gesture types.

Arm-based Hand Gesture Interfaces

Arm-based hand gesture interface is another approach to explore hands-free and implement-free input techniques. Some research projects based on muscle-computer interface [17] use forearm electromyography to decode the signal from motor neurons associated with muscle activation and recognize

four-finger pinch gestures at an accuracy rate of 79%. A commercial product named Myo [3] also uses EMG for gesture detection. However, finer gestures involving single finger movements are not discernible to such devices. In addition, compared to the aforementioned interfaces, EMG systems require high data rate and have high power consumption due to extensive signal processing.

As mentioned above, these interfaces with a variety of sensing technologies in different places to assess finger movement are the most common approaches to recognize hand gestures. However, BackHand senses hand gestures by exploring the back of hand which is not in the three classifications. We achieved 95.8% average accuracy for 16 popular hand gestures when personalized for each participant. Additionally, we are not aware of any research that explored the back of hand to recognize hand gestures.

BACKHAND-PROTOTYPE DESIGN

Physical changes on the back of hand caused by finger motions and gestures are visible to the human eye. However, there is still difficulty on detecting it through most sensing techniques because such physical changes are very small in comparison to that of finger joint movements. Moreover, physical changes on the back of hand varies from person to person. Accordingly, we have two major requirements. First, a small and light device must be able to sense small physical changes. Second, a device must be able to extract individual information from multiple spots on the back of a human hand. Among a variety of sensing techniques, we conclude that strain gauge sensors best fit our requirements for this prototype.

Hardware

Strain gauge sensors are sensitive devices used to measure very small strain and elongation of an object such as buildings, foundations, and other structures. In order to accomplish such measurement, a strain gauge must be adhesively attached to the object. Any tension or compression on the object leads to changes in the electrical resistance of the sensor.

In our prototype, the sensing stage of our prototype consists of a reusable hydrogel-based artificial skin and 19 120-ohm 2-mm strain gauge sensors. The artificial skin is attached to the back of hand and acts as a sensor carrier. As shown in

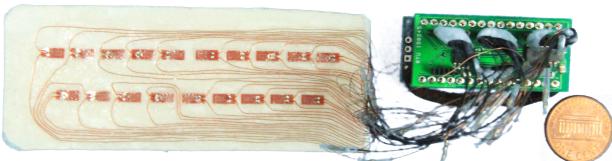


Figure 3. The layout shows a total of 19 sensors with a base length of 6.3 mm distributed into a 2-row configuration on an artificial skin. The strain sensors should avoid physical contact with each other by leaving a spaced interval of 2 mm between the neighboring sensors in a row. Therefore, a number of 10 sensors fits the hand width of a normal person. The second row includes 9 sensors to increase resolution and to measure the physical properties along the vertical direction between each spaced interval of the first row.

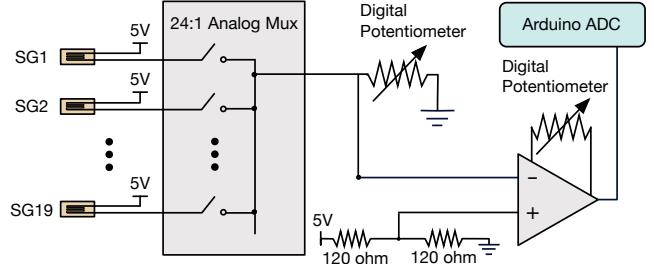


Figure 4. The complete circuit diagram. Note that SG stands for strain gauge. The 24:1 analog multiplexer is practically made up of 3 analog multiplexers.

Figure 3, the sensors are distributed into a 2-row configuration. The following sentences explain for such sensor layout. An average of human hand width is approximately 81 mm [1], and the base length of a 2-mm strain gauge is 6.3 mm. In order to avoid physical contact between any two neighboring sensors in a row causing strain interference, a spaced interval of 2 mm separates them, resulting in 10 strain gauges in a row across the hand width. Without the loss of strain measurement along the vertical direction between the two neighboring sensors, a second row with 9 sensors, each of which is vertically aligned with each center of the spaced intervals on the first row, is placed under the first row. The sensor array measures the deformations of multiple individual spots of the artificial skin caused by physical changes on the back of hand. Thus, our prototype is made possible to become a hand gesture interface for general users.

Next, the signal conditioning stage includes a Wheatstone bridge, 3 8-to-1 analog switches (MAX4617, Maxim Integrated), a dual digital potentiometer (MCP4251, Microchip), and an instrumental amplifier (AD623, Analog Devices). Finally, the processing stage includes an Arduino Nano used to sample the sensors and control chips in the signal conditioning stage. The complete circuit diagram is shown in Figure 4 and a customized circuit board is shown in Figure 5.

The 19 sensors are read sequentially through the analog switches, so each sensor is electrically connected at a time. Once a sensor is active, the sensor became one of the 4 resistors on the Wheatstone bridge. Since the resistance of each sensor varies under no applied external force, a digital potentiometer in series with the connected sensor is used for calibration. Another digital potentiometer serves the purpose of amplifier gain adjustment. In our prototype, the instru-

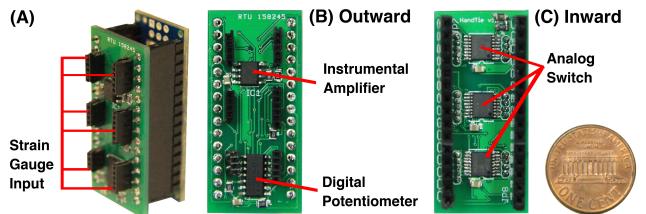


Figure 5. A customized 2-layer Arduino Nano shield circuit board used for signal conditioning is shown in (A) whose outward side is shown in (B) and inward side is shown in (C).

mental amplifier gain is set to be approximately 800 to 1000. Due to the dynamic response time of the amplifier, a 100-microsecond delay is added between each reading leading to a sampling rate of 75 Hz for our prototype.

Machine Learning

A support vector machine (SVM) is a novel supervised learning method that basically analyses the given data and recognizes patterns from it. We used LIBSVM tool [5], an open source library for SVM. We scale each feature to the range of 0 to 1 and then train a multi-class SVM classifier with a Radial Basis Function (RBF) kernel.

EVALUATION: MOST PROMISING LOCATION

The purpose of this evaluation is to investigate the accuracy of several row configurations ranging from back of the hand to wrist, and to uncover the most promising location that maximizes the performance of identifying hand gestures.

Based on our observation and assumption of the anatomy of finger tendon networks, a row configuration laying across the hand width is the most intuitive pattern to capture all tendon and muscle activities caused by movements of each finger. Although it is reasonable that the highest accuracy rate (e.g. >90%) may occur by covering the entire hand full of sensors, reaching high accuracy rate with less coverage area is worthy of being studied.

In addition, we provide the accuracy rates of each row configuration and visualizations of physical changes on the back of hand caused by gestures that are collected from all participants. These can enable us to understand, illustrate, and provide insights into the physical properties of the back of hand.

Study Design

We recruited 10 participants (4 females, 9 right-handed) aged between 20 and 32 (mean: 23.2) to test 16 gestures, shown in Figure 6.

Gesture Set

We mainly picked a gesture set from American Sign Language (ASL) such as digits from 0 to 10. In addition to the ASL gestures, we chose 4 gestures possessing significant cultural meanings in Asia such as digit “six”, “seven”, and “zero”, and “I-love-you”, each of which also coincides with “Y”, “L”, “S”, and “ILY”, respectively, in ASL. Therefore,

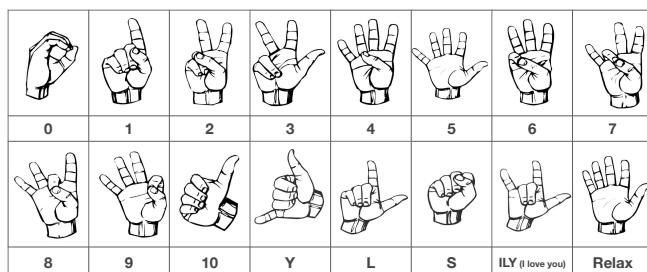


Figure 6. 16 gestures were chosen from American Sign Language and popular Asian gestures.



Figure 7. Study Setup: (A) monitor for on-screen instruction. (B) A button pushed by a participant to begin recording sensor readings while performing a gesture from the on-screen instruction. (C) A tilt platform for the arm to rest on. (D) Our prototype attached to the back of hand at specified locations.

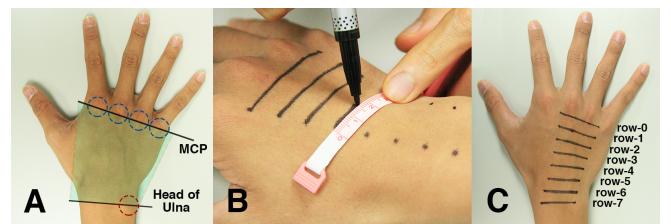


Figure 8. The procedures of marking the 8 rows on the back of hand. (A) Locate the MCP joint positions and the head of ulna. (B) Evenly mark 8 points from the index MCP joint to the imaginary horizontal line that goes through the head of ulna. Evenly mark another 8 points from the pinky MCP joint to the head of ulna. Draw 8 lines on the back of hand. (C) 8 rows evenly spaced between the MCP joint and the head of ulna are marked where our prototype aligns.

the gestures in Figure 6 are denoted using ASL letters. Finally, a neutral hand posture, relaxed gesture, along with 15 gestures adds up to a total of 16 gestures.

Procedure

Figure 7 shows the study setup. The participants rested their right arm and wrist on the tilt platform and were seated on a chair whose height they could adjust to perform the tasks with ease. The monitor in front of the participants displayed instructions during the study.

Initially, 8 rows (16 positions) across participants back of the hand were uniformly sampled and marked. The sampling scope ranged from the position of metacarpophalangeal joints (MCP, the knuckle between the hand and the finger) to the head of ulna (at the wrist). The markers served as affixing targets for artificial skin composed of arrays of 19 strain gauges are shown in Figure 8.

We then affixed the artificial skin composed of arrays of strain gauges to these 8-row markers on the right hand for 4 rounds (2 rows each time). At each round, each participant performed each of 16 gestures from on-screen instructions. Each gesture was collected 10 times. Each of them was sampled at 10 Hz in a duration of 2 seconds. The order of the gestures was randomly selected. Overall, the study took approximately 80 minutes per participants.

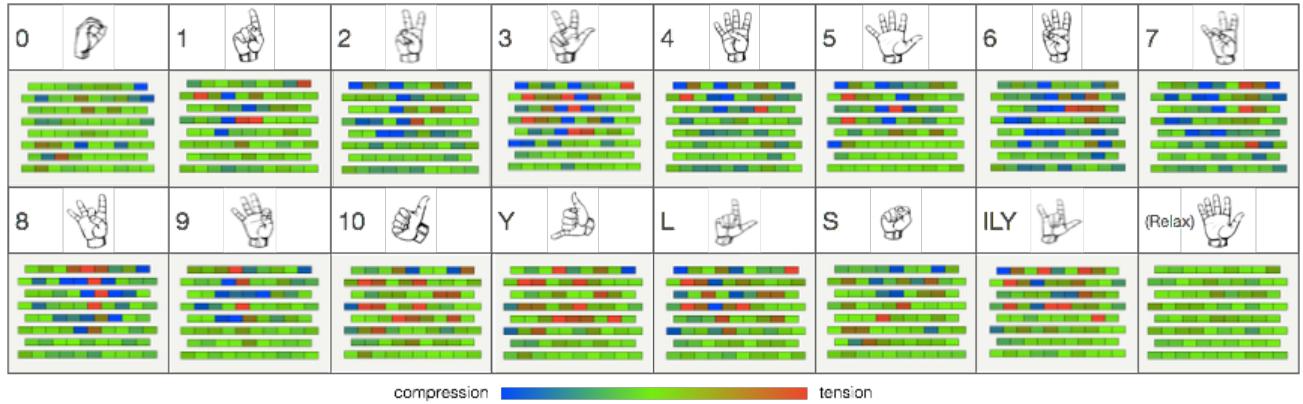


Figure 9. 16 heat maps display all 16 gestures performed by one of the participants. Each entry is an average heat map of 10 trials of a gesture. The sensor reading patterns are significantly different across 16 gestures.

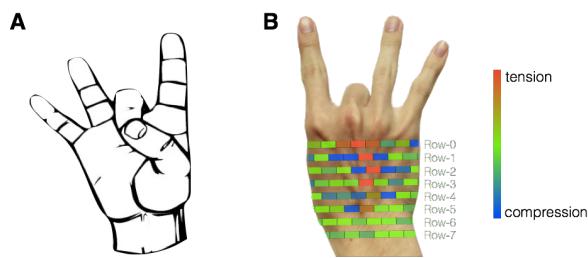


Figure 10. (A) A gesture of “8” from ASL is performed. (B) A heat map overlapped with a hand is illustrated to show the sensor values of each strain gauge resulting from the gesture of “8” performed by a user.

Results

Our results consist mainly three parts. The first part is a visualization of physical changes on the back of hand. The second part is a leave-one-user-out cross validation to test the feasibility of building a general model across participants. The last part is the determination of the most promising location through a personalized gestural recognition accuracy.

Heat Map Visualization

The visualization is shown in heat maps which display a block with highest tension sensor value as red and highest compression value as blue. The color green represents the neutral state (without tension and compression) of a strain gauge. Other blocks with colors in-between the extremes of red and blue on the color spectrum feature sensor values between the highest tension value and highest compression value.

Figure 9 showing all 16 gestures performed by a participant implies that every pattern on the back of the hand has its unique gesture model inside it. Figure 10 and Figure 11 indicate heat maps of a gesture of “8” performed by participants. From Figure 10(B), the strain sensors attached along with the prominent, superficial tendon of middle finger are stretched, while sensors in-between tendons are compressed. Figure 11(B) indicates the consistent reading patterns for a single user concluding the stability to utilize such signal source. The sensor reading patterns are significantly different across 10 participants in Figure 11(A) implying the difference of both the prominence of tendons and the physi-

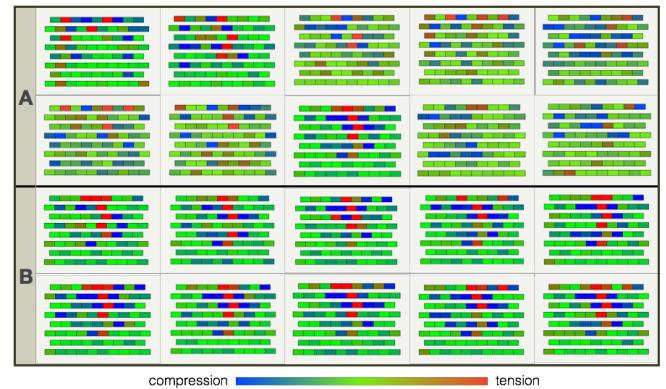


Figure 11. (A) Each entry is an average heat map of 10 trials of gesture “8” performed by a single user. The color distributions of each heat map differ quite much from person to person. (B) 10 heat maps display all 10 trials of gesture “8” carried out by a participant. The color distributions remain fairly consistent throughout the trials.

cal appearance of finger tendon networks among individuals. The measurements of tendon activities contribute the visualization which leads to a better understanding of the back of hand that in turn possesses a potential for hand gesture recognition.

Accuracy Rates

We tested all 15 possible row configurations for determining minimum strain sensor coverage area on the back of hand including 8 single-row configurations (from row-0 to row-7) and all 7 two-adjacent row configurations (row-01, row-12, row-23, row-34, row-45, row-56, and row-67) of sensing locations from each participant.

To understand how physical changes on the back of hand vary from person to person, the results of a leave-one-user-out cross validation are shown in Figure 13 where the value of each bar is an average of test results of each participant validated by the other 9 participants’ training model of a row configuration on the back of the hand. The highest accuracy rate of 27.4% occurs at the location of row-23. Hence, the low accuracy rates imply that physical patterns on the back of hand are likely to be in low similarity among individuals.

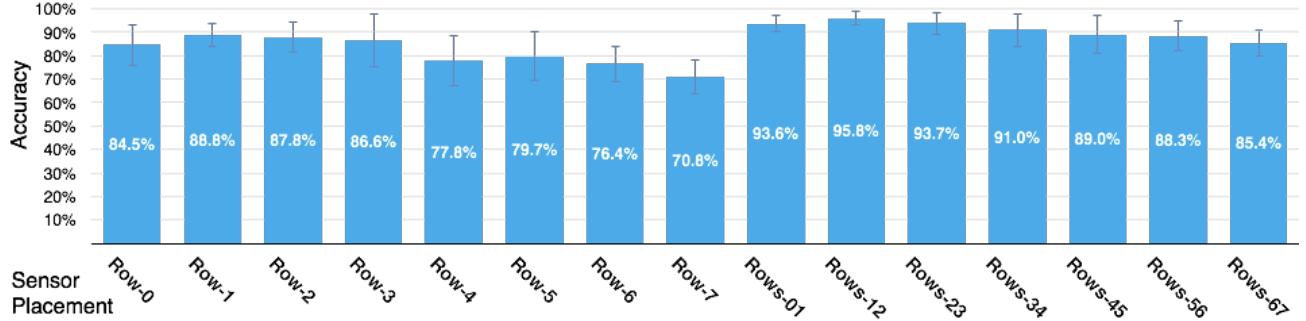


Figure 12. The result of the average of 10-fold cross validations of all 10 participants. The value of each bar is an average of personalized accuracy rate in 10-fold cross validation of a row configuration across all participants.

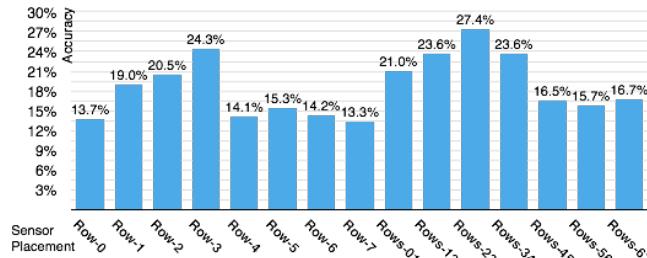


Figure 13. The result of the average of leave-one-user-out cross validations of all 10 participants. The value of each bar is an average of test results of each participant validated by the other 9 participant's training model of a row configuration on the back of the hand.

Based on the results from the leave-one-user-out cross validation, constructing a generalized model across general users is highly unlikely. Therefore, we delve into personalized accuracy rates and the aim for a most promising location. Each row configuration was evaluated based on a 10-fold cross validation where each fold contains data from a single trial of a single participant.

First, we average the accuracy rates of each of the 10 folds in a cross validation for a row configuration of a single participant. Second, we take the average of the accuracy rates of a row configuration across 10 participants, and the overall accuracy rate of a row configuration is obtained. Such computational process is repeated to retrieve the overall accuracy rates of all 15 row configurations. The best results are obtained with a selection of highest mean accuracy rate of a row configuration across all participants. In 10-fold cross validation shown in Figure 12, the mean accuracy rate of the best result among 15 row configurations across all participants is 95.8% (SD: 3.14%).

The results show that (1) the configuration row-12 has the highest accuracy rate, indicating the most promising location is not on the closest row to the finger joints but second to it. (2) The accuracy rate of each single-row configuration is lower than that of its adjacent 2-row configuration. For example, the accuracy rate of configuration row-1 is lower than that of row-01 and row-12. Hence, the greater sensing resolution, the greater the accuracy rate is obtained. (3) The accuracy rates decrease along the rows as approaching to the

wrist where physical signals become unrecognizable among hand gestures.

DISCUSSION

Our discussion is divided into three subsections. First of all, we show all the classification rates for 16 gestures through confusion matrix for better insight into the differences and similarities among hand shapes. Secondly, we reduce the number of gestures to 5 in order to present its performance under a different scenario for small numbers of gestures. Thirdly, prototype limitations are listed to address some main factors for improvement.

		Predicted Class (%)															
		0	1	2	3	4	5	6	7	8	9	10	Y	L	S	ILY	(Relax)
Actual Class (%)	0	98.5	0.5	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1
	1	0.0	94.0	1.0	0.6	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.9	1.0	0.6
2	2	1.3	0.0	96.4	0.0	1.1	0.0	0.2	1.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	3	0.0	0.0	1.0	95.8	0.0	1.8	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.4	1.0	0.0
4	4	0.7	0.0	0.8	0.0	97.4	1.2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	5	0.0	0.0	0.0	1.0	0.1	96.9	0.7	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.4
6	6	0.0	1.0	0.0	0.0	0.0	0.0	98.1	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	7	0.0	0.0	0.5	0.0	1.1	0.0	0.5	98.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
8	8	0.5	0.0	0.5	0.0	0.0	0.0	0.0	0.0	98.1	1.0	0.0	0.0	0.0	0.0	0.0	0.0
	9	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	99.0	0.0	0.0	0.0	0.0	0.0	0.0
10	10	1.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	90.8	2.0	0.0	5.2	0.0	0.0
	Y	0.0	0.8	0.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.4	98.2	0.0	0.0	0.0	0.0
L	L	1.9	0.2	1.0	0.0	0.1	0.0	0.0	0.0	1.0	0.0	0.0	90.2	0.0	5.1	0.7	0.0
	S	1.0	2.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.5	4.4	0.0	0.0	89.6	0.0	1.0
(Relax)	ILY	0.0	0.0	0.0	0.0	0.0	0.8	0.0	0.0	0.0	0.0	0.7	2.0	0.0	96.1	0.6	0.0
	(Relax)	0.4	1.1	1.0	0.0	0.0	1.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	95.6	0.0

Figure 14. Confusion matrix for accuracies using training data from the row configuration of row12.

Confusion Matrix

A confusion matrix shown in Figure 14 describes the misclassification rates of the gestures using the configuration row-12. The occurrence of false detection is usually resulted from the similarities of handshapes. For instance, the “S” gesture has the lowest true positive rate (89.6%) among 16 gestures and has a chance of 4.4% to be classified as the “10” gesture because the two gestures only differ by the thumb posture. Another reason for misclassification is associated with weak physical signals. For example, the “L” gesture has a true positive rate (90.2%) and a false negative rate of 5.1% for being classified as the “ILY” gesture likely because the physical changes caused by movements of the pinky finger are less obvious compared to that of other fingers. Therefore,

an increase in the amplifier gain leads to a better effectiveness in extracting weak signals. Overall, the true positive recognition rates are greater than 90% for 15 gestures and 95% for 12 gestures out of 16 gestures. From such result, it is reasonable to infer that the back of hand is an excellent signal source for a hand gesture interface.

Signal Source Comparison

To verify our assumption that the closer the signal source is to the fingers, the greater the accuracy rate of finer gesture recognition is obtained, a comparison is made with a related work [6] as in which we use the exact same setup except for the on-screen feedback. Currently, WristFlex [6] is a work with the highest accuracy rates (>80%) supporting the ample number of gestures (5 gestures) among other works using wrist-based signal source. We choose the 4 pinching gestures and a relaxed gesture [6] which are also included in our gesture set. We pick the training data from the most promising location. The methods to compute the accuracy rates are exactly the same as [6]. The accuracy rate obtained from the back of hand is statistically greater than that from the wrist in a 10-fold cross validation (98.2% vs 96.3%) and in a real-time cross validation with 3 training trials and 6 testing trials without feedback (90.6% vs 69.3%). The results indicate that sensing from the back of hand leads to a higher accuracy rate due to its close proximity to the fingers.

Limitations

Aforementioned in the leave-one-user-out cross validation, a small likelihood of a generalized model to fit all users is one of the limitations of utilizing such signal source.

Our prototype employs strain sensors which are meant to measure very small strain and elongation of a rigid structure. The sensors can be damaged by applying an excessive stress while removing the artificial skin off the user. High elongation strain gauges can be adopted to resolve this problem.

Although we picked a highly reusable hydrogel-based artificial skin that can be used repeatedly across 10 participants in our user study, the artificial skin absorbs perspiration which decreases the adhesion for fixation on skin. We continue to look for alternative solutions such as "smart skin" systems that integrates this sensing technique with other modalities that warrant or even require to be located on the back of the hand.

The signals from the sensors must undergo stages of signal processing and conditioning leading to a number of chips required in the hardware taking up space in our prototype. Thus, the current size of our hardware is not suitable for an on-body hand gesture interface. Moreover, the system requires a USB serial cable for power supply and data transmission. It will be future work to redesign a small sized, battery powered, and wireless hardware.

EXAMPLE APPLICATIONS

In addition to gesture recognition for general applications, we broadened the application of BackHand by showing a musical performance application and smartwatch control using static and dynamic gestures, respectively.

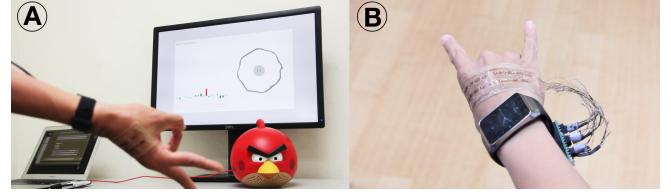


Figure 15. (A) A musical performance application turns the hand into a musical instrument. Gestures are translated into musical notes. (B) A smartwatch control provides a more user friendly, intuitive hand gesture interface.

Musical Performance

Since BackHand can distinguish finer gestures, we developed an application which turns the hand into a musical instrument whose musical notes are translated from designated gestures. The pinch gestures such as the "9", "8", "7", and "6" gestures from ASL are translated into musical notes of "Do", "Re", "Mi", and "Fa", respectively in C major. Similar to the pinch gestures, the thumb makes physical contact with the bottom segment of the index, middle, and ring finger to generate the rest of the musical notes of "Sol", "La", and "Si", respectively, in a C major scale. BackHand can be turned into a wearable musical instrument that can increase productivity and be entertaining.

Smartwatch Control

Due to the high data rate and a variety of gestures supported by BackHand, a smartwatch can be operated by a single hand. Users are able to customize the input gestures. For example, sliding the thumb across the index finger skips to the next page, and the "9" gesture from ASL makes a confirmation command. The application makes smartwatch control more user friendly by providing an intuitive gestural input and without the need for touching the small screen.

CONCLUSION

We made a classification for hand gesture interfaces based on the three common input signal sources, each of which was briefly introduced with its principle and its several characteristics such as gestural recognition rate, computational overhead, finger dexterity, comfort, limitations, and so forth. Based on our classification, in general, the closer the source is to the fingers, the greater the accuracy rate of gestural recognition is obtained. The back of the hand is a new category of input signal source and is closer to fingers where obvious tendon movements are located. Sensing at this location causes less interference with finger agility. We implemented a prototype, BackHand, whose sensing stage consists of a reusable artificial skin and an array of 19 strain gauge sensors in a 2-row configuration which measure physical changes at multiple spots on the back of hand.

A user study was conducted to better understand gesture recognition accuracy and the effects of sensing locations. A set of 16 gestures was chosen from both American Sign Language and popular Asian gestures. We collected 8 rows of data from the participants' back of hand and employed SVM machine learning technique to discern 16 gestures. Our results showed that the most promising location spans from the

second row to the third row starting off from the metacarpophalangeal joints. The highest leave-one-user-out average gestural recognition accuracy rate is 27.4% which becomes one of the limitations due to the significant different patterns on the back of hand across participants. However, the personalized average gestural recognition accuracy rate using the data from the most promising location reaches 95.8% in a 10-fold cross validation. A visualization shown as heat maps indicating sensor reading patterns facilitates our understanding towards such phenomenon. We showed two example applications to inspire people with its potential use. Our results provide future developers an alternative choice for input sources of hand gesture interface.

REFERENCES

1. 2002. The effect of hand preference on hand anthropometric measurements in healthy individuals. *Annals of Anatomy - Anatomischer Anzeiger* 184, 3 (2002), 257 – 265.
2. 2015. Data Gloves — 5DT . http://www.5dt.com/?page_id=34. (2015).
3. 2015. Thalmic Labs. Myo. <https://www.thalmic.com/en/myo/>. (2015).
4. Grigore C. Burdea and Philippe Coiffet. 2003. *Virtual Reality Technology* (2 ed.). John Wiley & Sons, Inc., New York, NY, USA.
5. Chih-Chung Chang and Chih-Jen Lin. 2011. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology* 2 (2011), 27:1–27:27. Issue 3. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
6. Artem Dementyev and Joseph A. Paradiso. 2014. WristFlex: Low-power Gesture Input with Wrist-worn Pressure Sensors. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology (UIST '14)*. 161–166.
7. T. Deyle, S. Palinko, E.S. Poole, and T. Starner. 2007. Hambone: A Bio-Acoustic Gesture Interface. In *Wearable Computers, 2007 11th IEEE International Symposium on*. 3–10.
8. L. Dipietro, A.M. Sabatini, and P. Dario. 2008. A Survey of Glove-Based Systems and Their Applications. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on* 38, 4 (July 2008), 461–482.
9. Rui Fukui, Masahiko Watanabe, Tomoaki Gyota, Masamichi Shimosaka, and Tomomasa Sato. 2011. Hand Shape Classification with a Wrist Contour Sensor: Development of a Prototype Device. In *Proceedings of the 13th International Conference on Ubiquitous Computing (UbiComp '11)*. 311–314.
10. N.Y.Y. Kevin, S. Ranganath, and D. Ghosh. 2004. Trajectory modeling in gesture recognition using CyberGloves reg; and magnetic trackers. In *TENCON 2004. 2004 IEEE Region 10 Conference*, Vol. A. 571–574 Vol. 1. DOI: <http://dx.doi.org/10.1109/TENCON.2004.1414484>
11. David Kim, Otmar Hilliges, Shahram Izadi, Alex D. Butler, Jiawen Chen, Iason Oikonomidis, and Patrick Olivier. 2012. Digits: Freehand 3D Interactions Anywhere Using a Wrist-worn Gloveless Sensor. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology (UIST '12)*. 167–176.
12. Yoon Sang Kim, Byung Seok Soh, and Sang-Goog Lee. 2005. A new wearable input device: SCURRY. *Industrial Electronics, IEEE Transactions on* 52, 6 (Dec 2005), 1490–1499.
13. Joseph J. LaViola, Jr. 1999. *A Survey of Hand Posture and Gesture Recognition Techniques and Technology*. Technical Report.
14. Pranav Mistry and Pattie Maes. 2009. SixthSense: A Wearable Gestural Interface. In *ACM SIGGRAPH ASIA 2009 Sketches (SIGGRAPH ASIA '09)*. Article 11, 1 pages.
15. J.K. Perng, Brian Fisher, S. Hollar, and K.S.J. Pister. 1999. Acceleration sensing glove (ASG). In *Wearable Computers, 1999. Digest of Papers. The Third International Symposium on*. 178–180.
16. Jun Rekimoto. 2001. GestureWrist and GesturePad: Unobtrusive Wearable Interaction Devices. In *Proceedings of the 5th IEEE International Symposium on Wearable Computers (ISWC '01)*. 21–.
17. T. Scott Saponas, Desney S. Tan, Dan Morris, Ravin Balakrishnan, Jim Turner, and James A. Landay. 2009. Enabling Always-available Input with Muscle-computer Interfaces. In *Proceedings of the 22Nd Annual ACM Symposium on User Interface Software and Technology (UIST '09)*. 167–176.
18. Koji Tsukada and Michiaki Yasumura. 2001. Ubi-Finger: Gesture Input Device for Mobile Use. In *Ubicomp 2001 Informal Companion Proceedings*. Citeseer, 11.
19. D. Way and J. Paradiso. 2014. A Usability User Study Concerning Free-Hand Microgesture and Wrist-Worn Sensors. In *Wearable and Implantable Body Sensor Networks (BSN), 2014 11th International Conference on*. 138–142.