This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2017.2779466, IEEE Sensors Journal

Smart Wearable Hand Device for Sign Language Interpretation System with Sensors Fusion

B. G. Lee, *Member, IEEE*, and S. M. Lee

***Abstract*—Gesturing is an instinctive way of communicating to present a specific meaning or intent. Therefore, research into sign language interpretation using gestures has been explored progressively during recent decades to serve as an auxiliary tool for deaf and mute people to blend into society without barriers. In this paper, a smart sign language interpretation system using a wearable hand device is proposed to meet this purpose. This wearable system utilizes five flex-sensors, two pressure sensors, and a three-axis inertial motion sensor to distinguish the characters in the American Sign Language alphabet. The entire system mainly consists of three modules: a wearable device with a sensor module and a processing module, and a display unit mobile application module. Sensor data are collected and analyzed using a built-in embedded support vector machine classifier. Subsequently, the recognized alphabet is further transmitted to a mobile device through Bluetooth low energy wireless communication. An Android-based mobile application was developed with a text-to-speech function that converts the received text into audible voice output. Experiment results indicate that a true sign language recognition accuracy rate of 65.7% can be achieved on average in the first version without pressure sensors. A second version of the proposed wearable system with the fusion of pressure sensors on the middle finger increased the recognition accuracy rate dramatically to 98.2%. The proposed wearable system outperforms the existing method, for instance, although background lights, and other factors are crucial to a vision-based processing method, they are not for the proposed system.**

***Index Terms*—Gesture recognition, machine learning, mobile application, sign language recognition, wearable computing**

I. Introduction

SIGN language plays a vital role for deaf and mute people to communicate among themselves or with normal people in a non-verbal manner. Gestures are the primary method to convey messages, which are usually conducted in a three-dimensional space, known as a signing space [1], through an integration of manual and non-manual signals. Manual signals commonly correspond to hand motions and hand posturing, whereas non-manual signals correspond to an external appearance such as mouth movements, facial expressions, and body orientation [2]. Nevertheless, sign language has not been standardized globally. Each nation has developed its own sign language, such as the American Sign Language (ASL) and Germany Sign Language (GSL). However, each sign language varies slightly within different regions of the same country. Hence, it can be a challenge to develop a standardized sign language interpretation system for use worldwide. In a previous study, sign languages have been recognized using two major techniques, i.e., vision and non-vision approaches [3].

In fact, vision method is the major technique applied for sign recognition in the past decades. A system that uses a camera to observe the information obtained through hand and finger motions is the most widely adopted visual-based approach [4]. Tremendous effort and study has gone into the development of vision-based sign recognition systems worldwide. Indeed, vision-based gesture recognition systems can be subdivided into direct and indirect approaches.

A direct approach detects the hand gestures based on the RGB color spaces of the skin color. For instance, Goyal *et al.* [5] identified Indian Sign Language (ISL) using the scale invariance Fourier transform (SIFT) algorithm by searching the matched key points between the input image and images stored in a database. A similar method was also applied by More *et al.* [6] using the SIFT algorithm, which further reduces the dimensions of the feature vector using a principal component analysis (PCA) algorithm for speeding up the processing time. To detect the dynamic hand gestures used in Japanese Sign Language (JSL), Murakami *et al.* [7] proposed the use of recurrent neural networks capable of recognizing the JSL finger alphabet, which has 42 symbols. In contrast, Chowdary *et al.* [8] used a simple scanning method to compute the orientation and movement of fingers in binary converted images captured from a web camera. Khan *et al.* [9] proposed a more sophisticated gesture recognition system using digital images, including image filtering (pre-processing), image segmentation, color segmentation, skin detection (finger and hand detection using binary images), and template matching.

Meanwhile, an indirect approach identifies the fingers and hand gestures based on the RGB color spaces segmented based on different colors for each finger using a data glove. A possible segmentation method using RGB color spaces along with a hand glove for gesture detection was proposed by Siby *et al.* [10]. This method utilizes the RGB color space values extracted from a captured hand gesture image of a data glove, and compares the values with those stored in a database. Lamberti *et al.* [11] also proposed a real-time hand gesture recognition system using a learning vector quantization classifier to differentiate three colors of a wool glove

This work was supported by the Basic Science Research Program through the Ministry of Education of the Republic of Korea and National Research Foundation of Korea (NRF-2017-1678-2017R1D1A1A09000975).

B. G. Lee, and S. M. Lee are with the Keimyung University, Daegu, 42601 Republic of Korea (e-mail: bglee@kmu.ac.kr).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2017.2779466, IEEE Sensors Journal

corresponding to the palm, fingers, and remaining parts of the glove, respectively, from captured images. An equivalent approach was also proposed by Iwai *et al.* [12], who divided a color glove into twelve regions with twelve different colors, i.e., ten finger regions, a wrist region, and a region made up of other parts of the hand. A decision tree pattern recognition method using these distinct image features is then applied to determine the regions of the hand.

The exploitation of a vision-based method is greatly affected by the processing of the images, such as image filtering, background cancellation, color segmentation, and boundary detection. For instance, diverse and uncontrolled background images can influence the skin color segmentation or movement detection. Indeed, many researchers have failed to address these complications, and no solid solutions have yet been proposed. Consequently, a non-vision based method is an alternative approach. This method typically utilizes flex and motion sensors to measure the flexion of fingers and the orientation of the hand, respectively. Dawane *et al.* [13] attached five flex sensors on a glove with respect to each finger to identify hand gestures by matching the motions with those in a stored motion database. Preetham *et al.* [14] also proposed a similar technique of using flex sensors, but mapped the sensor data to a character set, which was implemented using a minimum mean square error (MMSE) algorithm for gesture recognition. The results are displayed as text on an LCD screen. This technique was improved by Patil *et al.* [15]; here, the bending of each sensor is further divided into three flexions, namely, a complete bend (finger close), partial bend, and straightening (finger open). Each ASL alphabet is then mapped according to the bend flexions to be used for template matching.

On the other hand, inertial motions of hand or fingers are alternative approach for gestures recognition. Kim *et al.* [16] developed a glove-based 3-D hand motion tracking and gesture system that consists of three tri-axis accelerometer sensors placed on the thumb, middle finger, and back of the hand, respectively. The glove is able to detect simple rule-based hand gestures (e.g., scissor, rock, and paper) based on two different angular positions of the sensors: horizontal (z-axis) and vertical (x-axis) gestures. Lu *et al.* [17] implemented a YoBu glove with total of 18 inertial motion unit (IMU) sensors. Each finger consists of three IMMUs at each joint with a total of 15 for the five fingers. The remaining IMU sensors are placed on the arm, forearm, and upper arm. An extreme kernel-based learning machine is implemented to identify specific gestures based on a total of 54-dimension extracted features. On the other hand, Lim *et al.* [18] proposed a novel method that uses a small-sized infrared optic sensor, developed as a virtual button, to observe finger flexion patterns on the wrist caused by moving fingers. Five dynamic gestures were proposed: "bye," "hi," "hold," "release," and "wave," all of which are detected using a hidden Markov model (HMM) algorithm. Xie *et al.* [19] developed an accelerometer based smart ring to detect eight basic and twelve complex gestures in 2-D space. A segmentation algorithm is utilized to identify subject gestures which are further encoded by a Johnson code. Hsu *et al.* [20] presented an inertial-based digital pen to recognize the handwriting and gestures with dynamic time warping (DTW) method. The pen is hold by the user when writing the numerals or English lowercase letters and further transmitted wirelessly to a computer for online recognition. The similar hand gesture recognition approach with DTW method is proposed by Yin *et al.* [21] that adopted training-free method which did not require training samples. The gesture recognition process is carried out by first extracting the features followed by a robust template matching method. Besides that, Galka *et al.* [22] proposed an accelerometer glove for sign language recognition which consisted of seven active three-axis acceleration sensors with five located on the fingers (one sensor on each finger), one on the arm and one on the wrist. The sign recognition model is defined by a HMM and a parallel HMM approach. Cai *et al.* [23] also developed a wireless data glove type with four fingers button. The raw sensor data are converted into movement, acceleration, rotation and other features for gesture recognition. Meanwhile, Liu *et al.* [24] presented an interesting idea to integrate inertial and vision depth sensors with HMM model to recognize six hand gestures, including "wave", "hammer", "punch", "draw X", "circle" and "other" gestures. However, the proposed system is not feasible in practical usage as both sensors needed to present for high accuracy gestures recognition. Nevertheless, a simple data glove with three-axis accelerometer, magnetometers and gyroscopes is proposed by Kim *et al.* [25] which converted the sensor data into angle data for sign language recognition. Sousa *et al.* [26] presented a GyGSLA system, a wearable glove that aimed to help inexperienced people in learning the new Portuguese sign language alphabet which is tested with three completely inexperienced sign language subjects. The similar study is also performed by Caporusso *et al.* [27] by introducing dbGLOVE, a wearable device for supporting deaf-blind people to communicate with others.

Meanwhile, physiological sensors such as sEMG are also another popular gesture recognition technique. Lu *et al.* [28] proposed a score-based sensor fusion scheme using four sEMG sensors and a three-axis accelerometer connected to a mobile application to realize gestures-based real-time interaction. A set of 4 small-scale gestures, 15 large-scale gestures and user defined personalized gestures are proposed in the study. A similar method is also proposed by Wu *et al.* [29] with four sEMG and an inertial sensor that placed on the wrist to detect 40 most commonly used ASL words. The study observed that only a single channel of sEMG located on the wrist is sufficient for the ASL recognition. Likewise, Wu *et al.* [30] fused the information from an inertial sensor and sEMG sensor which are placed on a wearable system to recognize 80 commonly ASL signs with selected feature subset and processed by a support vector machine classifier.

This study aims at the development of a sign language interpretation system by analyzing hand and finger gestures from a smart wearable device. The finger gestures are observed through the flexion of the flex sensors, whereas the hand gestures are examined based on the hand motion through the orientation derived from an inertial motion sensor. The gestures are recognized using a support vector machine (SVM) model

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2017.2779466, IEEE Sensors Journal

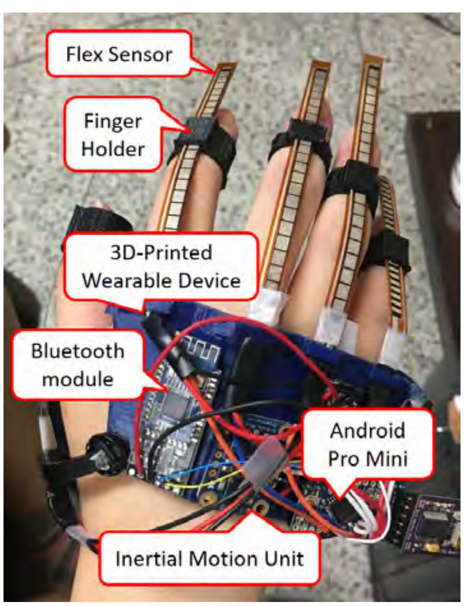
implemented in the wearable device. The gestures are then received using our developed mobile application through a wireless Bluetooth transmission, and text is displayed on the mobile device screen. Moreover, a text-to-speech service is also available in the mobile application, which instantly converts the received texts into audible outputs.

II. System Design

A. Hardware Design and System Flow

For this study, a custom-made wearable device was designed and built using a 3D printer to hold the hardware components, as illustrated in Fig. 1. The wearable device holder (see Fig. 2) is printed using flexible filaments with good elasticity. These filaments enable functional hinges, joints, and shaped parts, allowing the device to fit different hand sizes. Five finger holders were also designed using a flexible filament placed on the first joint of each finger to hold the flex sensors. Similarly, these flexible holders can also accommodate different finger sizes of different users.

Finger gestures are exploited through the flexion of flex sensors placed on the top of the finger. The flex sensors used in this study are either 4.5 inches [31] or 2.2 inches [32] in size. The shorter length flex sensor is suitable for the pinky finger, whereas the longer length flex sensor is used for the other four fingers. The flex sensor for the thumb has a longer connection distance to the microcontroller board (see Fig. 1) as compared to the other fingers, and thus a longer length flex sensor is used instead of a shorter length one. In fact, the flex sensor consists of both omnidirectional and bidirectional types. The omnidirectional flex sensor changes its resistance when it bends in one direction only, whereas the bidirectional type changes its resistance when it bends in both the upward and downward directions. This study utilizes a bidirectional type, and each flex sensor is connected with a 10 kΩ resistor. In addition, an Adafruit BNO055 absolute orientation sensor [33] is used as an inertial motion unit (IMU) for observations of the hand gesture movements. A 9-degree-of-freeom (9-DOF) IMU is integrated with an MEMS accelerometer, magnetometer, and gyroscope under a single die, and processed using a high-speed ARM Cortex-M0 processor. The device abstracts the sensor fusion and derives the orientation data in quaternions, Euler angles, or in a vector format.



(a)

Fig. 1. 3D printed wearable device that holds the hardware components, which include an Android Pro Mini microcontroller, a flex sensor, a motion sensor, and a Bluetooth low energy (BLE) module.

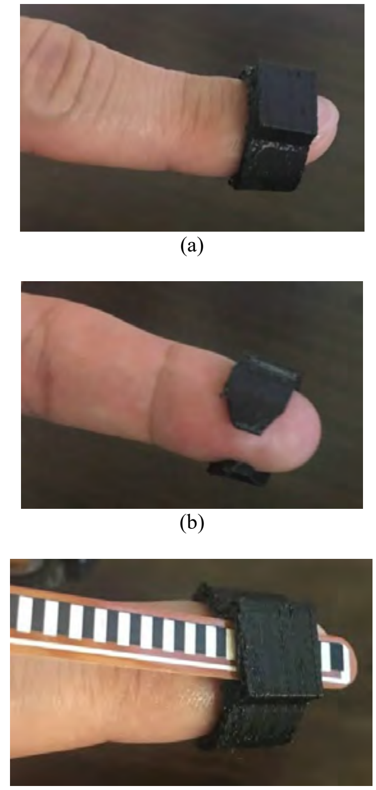

(a) (b) (c)

Fig. 2. 3D printed finger holder using a flexible filament that can accommodate different finger sizes, providing flexibility: (a) front view, (b) back view, and (c) holder with a flex sensor.

The proposed sign interpretation system is divided into three distinct modules: a sensor module, processing module, and application module, as shown in Fig. 3. The sensor and processing modules are implemented in the smart wearable

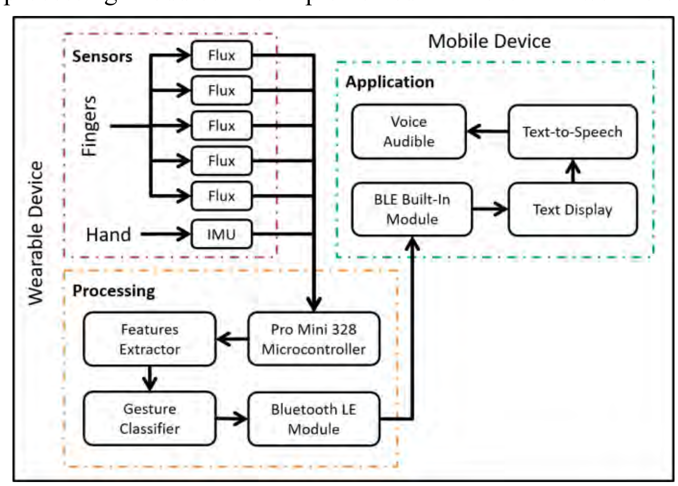


Fig. 3. Overview of sign interpretation system that consists of three modules, namely, sensors module, processing module, and application module.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2017.2779466, IEEE Sensors Journal

hand device, whereas the application module operates in an Android-based mobile device. The flex sensor and IMU data are collected using an Arduino Pro Mini 328 [34], which is operated at 5 V with an ATmega328 processor running at 16 MHz with an external resonator (0.5% tolerance). The features are extracted from the sensor data and serve as inputs to the built-in SVM classifier to determine the sign language alphabet letters. In this study, there are a total of 28 gesture patterns, which refer to the 26 alphabet letters of ASL, a "neutral" state (indicating no gestures to be observed), and an invalid sign. The detected sign is translated into text and transmitted to a mobile device using a Bluetooth 4.0 module [35], i.e., Bluetooth low energy. The received text is displayed on the mobile screen. Concurrently, the automated text-to-speech service translates the text into an audible voice to be played by the built-in speaker of the mobile device. Finally, all of these hardware components are powered using a 3.7 V 500 mAh ion lithium polymer battery [36].

B. *Experiment Setup*

Twelve subjects were recruited from a university campus and participated in the experiments voluntarily. The participating subjects do not possess any muscular diseases or neuromuscular disorders that would affect their sign gestures. The subjects were requested to perform all sign gestures for 20 times with each approximately 10 s each, as shown in Fig. 4(a). In addition, a "neutral" state was also recorded in which all of their fingers were wide open regardless of the hand orientation (see Fig. 4(b)). All data were recorded during the experiments using a desktop computer application. The data were saved in a text file format for easy accessibility. The computer application is not described in detail herein because it only served as a recording device.

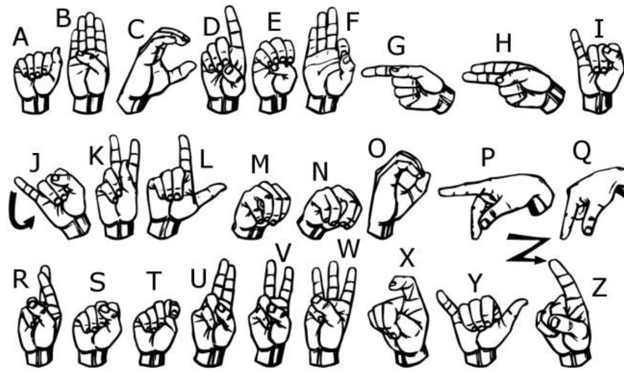
III. Methods

A. *Preprocessing*

Throughout the experiments, it was observed that the flexion values varied among the different subjects owing to the different hand sizes. A smaller hand size has a lower discrepancy in terms of the flex sensor values with respect to particular signs. In other words, the resistance of the flex sensors for a smaller hand has less variation as compared to the resistance with a larger hand. Thus, the differences among the signs are not discernable from the raw flex sensor values. To solve this issue, the sensor values were normalized based on the computed mean and standard deviation of each flex sensor for each subject into a range of [0, 1] for easier analysis as shown in (1),

$$\widehat{fs_i} = \frac{(fs_i - \overline{fs})}{\sigma_{fs}} \quad (1)$$

where fs_i, $\overline{fs}$, and σ_{fs} are the *i*-th sensor reading, mean and SD of flex sensor value respectively. The normalized values are computed individually for each sensor and for each test subject.



(a)

(b)

Fig. 4. (a) American sign language with a total of 26 letters and a (b) neutral state indicating relaxation or no gestures to be observed.

In other words, due to the different hand size, the grasping of fingers into fist shape generate different maximum flex sensors values. Thus, the mean and SDs of flex sensors are computed separately for each subject based on the subject's sensors readings. This method eliminated the necessity for the new user to perform all the sign gestures before the new user started to use the proposed system.

Meanwhile, each hand movement was derived from the IMU in three orientations: pitch, roll, and yaw. The calculation of each orientation is similar depending on the axis used. The angle agl_i of each axis is computed using a complementary filter method at the *i*-th time, as illustrated in (2),

$$agl_i = 0.98 \times (agl_{i-1} + gyr_i/gyr_{HZ}) + aglc_i \times 0.02, \quad (2)$$

$$aglc_i = \arctan(A_u, A_v) \times 180/\pi, \quad (3)$$

where gyr_i, gyr_{HZ}, and $aglc_i$ are the raw gyroscope sensor reading, gyroscope sensor sampling rate, and the angular acceleration speed (refer to (3)) at the *i*-th time, respectively. In addition, A_u, and A_v are the coordinates of the exclusive raw linear accelerometer readings that do not correspond to the angles being computed, e.g., to compute the pitch angle, A_u and A_v are denoted as the y-axis and z-axis of the raw linear accelerometer readings, respe ctively. The multiplier constant in (2) converts the angle from radians into degrees (°) for ease

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2017.2779466, IEEE Sensors Journal

of analysis. Further details of this computation can be found in our previous related study [37]. The flex sensors and IMU data are gathered at a sampling rate of 100 Hz.

B. Feature Extraction

To simplify and optimize the coding implementation, a vector of the flexion degree in tabular format was considered. The flexion degree is split into three regions in each vector. The first region is denoted as "no bend" or "slight bend," which is associated with a normalized flexion value within the range of [0.0, 0.3). The second region is considered as a "partial bend" with the associated normalized flexion value within the range of [0.3, 0.7), and the last region is a "complete bend" with associated normalized flexion value within the range of [0.7, 1.0]. These regions are abbreviated in order as OR (open region), PR (partially open or closed region), and CR (closed region). Table 1 shows a mapping of these regions for all 26 alphabet letters in ASL, as well as a "neutral" gesture.

It can be seen that some of the signs exhibit the same regions, for instance, the letters "U" and "V," "K" and "P," and "I" and "J," as well as "G," "L," "Q," and "Z." To further distinguish these signs, sensor level fusion [38] is adopted. In this study, recognition with only flex sensors and inertial sensor are denoted as 1st version and addition of pressure sensors for sensor level fusion are designated as 2nd version. To support the sensor fusion, two Flexiforce pressure sensors were considered. The resistance of a Flexiforce pressure sensor [39] is reduced when the surface is pressed, but the resistance does not change while being flexed. In fact, the resistance changes only when the pressure is applied to the round area at the end of the sensor, and ranges from 0 to 25 lbs of pressure. These sensors are placed below and on the left side (right side for a left-handed user) of the first joint of the middle finger, as illustrated in Fig. 5. The pressure sensors are connected to the digital inputs in processing module which produced reading of "0" if pressure sensor surface is not pressed, and produced reading of "1" if the pressure is sensed on the surface of the sensor. The inclusion of the pressure sensors on the middle finger managed to solve the issues for the letters "U" and "V," whereas the second pressure sensor resistance value is high for letter "U" and low for the letter "V". Moreover, the letters "G", "L," and "Q" can also be distingui shed based on the resistance of the respective pressure sensors.

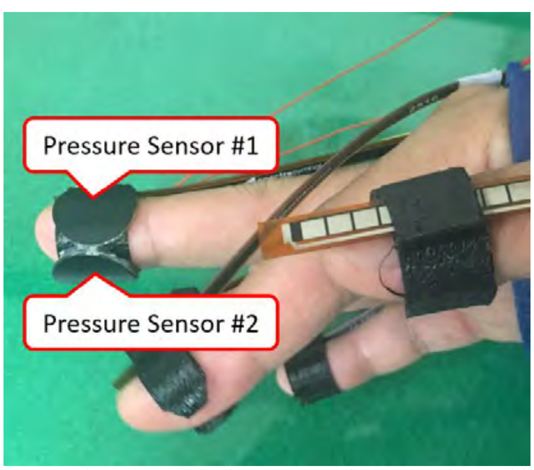


Fig. 5. Improved version (second prototype) of 3D printed wearable device (see Fig. 1) with fusion of pressure sensor added to the middle finger.

However, the letters "L" and "Z" still exhibit a similar pattern and can be differentiated through a hand motion. The letter "L" has static motion, whereas letter "Z" has dynamic motion over time, as depicted in Fig. 4(a). The same behavior is applied to discriminate the letters "I" and "J." Finally, the letters "K" and "P" can only be differentiated based on the hand orientation. To determine the hand motion, standard deviation (SD) of the angular reading from the IMU sensors is computed for each axis. The SD is a measure that is used to quantify the amount of variation or dispersion of the motion readings. For instance, the preliminary results had indicated that the mean SDs of letter "I" are approximately 0.0369 (pitch), 0.0151 (roll), and 0.0375 (yaw) while the mean SDs of letter "J" are 1.3968 (pitch), 0.5603 (roll), and 1.1548 (yaw) respectively. Thus, in this study, the computed SDs are sufficient to observe the occurrence of hand motions.

C. Sign Classifier

In this study, the signs are classified into 28 classes using a support vector machine (SVM) [40]. An SVM is a binary supervised learning classifier, that is, the class labels can only take the values of +1 and -1. The training procedure used a quadratic optimization algorithm to derive structural axes to separate the training dataset into n numbers of a hyperplane. Assume the i-th training sample using

$$(x_i, y_i), y_i \in \{-1, +1\}, i = 1,2,3, \dots, n, \tag{4}$$

where x_i is the feature vector and y_i is the training label in accordance to the feature vectors of the i-th training datasets. The decision boundary is defined as

$$f(x) = w \cdot x - b, \tag{5}$$

where the i-th feature is classified as positive (+1) if $f(x) > 0$, and negative (-1) if $f(x) < 0$. The separating hyperplane line is structured at $f(x) = 0$. The points positioned around the separating hyperplane line are known as support vectors (SVs) and their distance to the hyperplane line is known as the margin. Optimization of the SVM is calculated by finding the smallest distance among all SVs, as shown in (6), which is subject to (7).

$$\min_{w,b} \frac{\|w\|^2}{2} \tag{6}$$

$$y_i(wx_i - b) \geq 1 \tag{7}$$

The values of +1 and -1 are expressed as correctly classified alphabet and incorrectly classified alphabet respectively.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2017.2779466, IEEE Sensors Journal

However, in this study, there are more than two classes that are being classified. To obtain an *M*-class of the SVM classifiers, a set of binary classifiers need to be constructed, where each binary classifier is trained to distinguish one class from the rest. The results are then integrated to form a multi-class classification according to the maximal output of each binary classifier x_j, which is also known as the confidence value and j is referred to each alphabet binary classifier. Thus, x belongs to the class with the largest confidence value. In fact, there will be gestures that belong to none of the sign classes or the aforementioned "neutral" class in a real-world application, and thus any gesture with a confidence value of less than 0.5 (50%) is considered as an "invalid" class. In this study, a feature vector x is built using a total of ten features, which are the normalized flex sensor data for five fingers, two pressure sensor data, and finally, three computed SDs of angular readings from the IMU sensor. Sliding windows of 3 s are adopted to construct the feature vector for every 10 s sensor data in order to accommodate the changes in hand movement within a 3 s time period, particularly for the letters "J" and "Z." The values of the ten features in 3 s time period are averaged for the feature vector construction. There are total of 6,480,000 datasets (12 subjects × 20 times × 10 s × 100 Hz × 27 signs) collected in this experiments.

The SVM is trained and tested using the "leave one subject out" (LOO) method. Here, the SVM is trained using $n - 1$ subject datasets, where n is the total number of subjects. Subsequently, the trained model is tested using the leave-out subject dataset that did not participate in the training process. This process was repeated n times, where each subject was treated as a leave-out subject once when the testing dataset. The accuracy of each trained model is computed using (8), i.e.,

$$AC = \frac{TP+TN}{TP+TN+FP+FN}, \tag{8}$$

where TP, TN, FP, and FN refer to a true positive (the number of signs correctly identified), true negative (the number of signs correctly rejected), false positive (the number of signs incorrectly identified), and false negative (the number of signs incorrectly rejected), respectively. The overall accuracy of the SVM is eventually averaged. From the cross-validation (CV) model selection point of view, Arlot *et al.* [41] concluded that LOO method is best to utilize for model training with high computational complexity which is proportional to the number of data splits. Thus, it is suitable method to estimate the risk when learning a model.

IV. Results and Discussion

A. Experiment Results

Fig. 6(a) depicted a typical example plotting of the sensor value of index finger for the alphabet "A" (red), "B" (green), and "C" (blue) respectively. It was clearly observed that the flexion region for alphabet "A" of index finger is at CR region which is verified as shown in Table 1. The similar results also applied to the alphabet "B" (in OR region) and alphabet "C" (PR). The signs of the classification results are summarized in

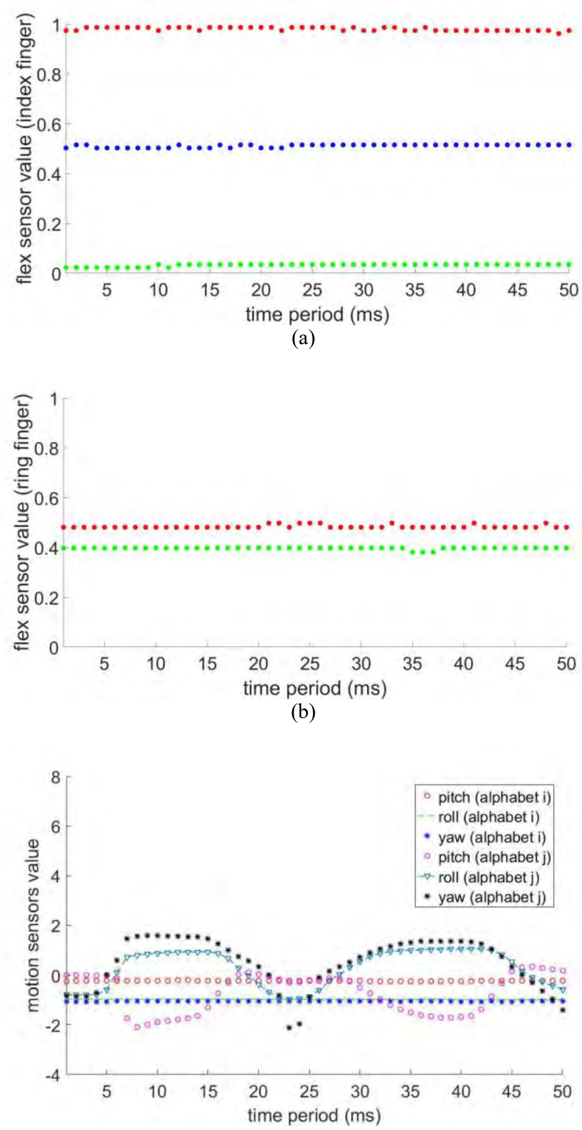


Fig. 6. The sensor values of (a) index finger for the alphabet 'A' in red, 'B' in green and 'C' in blue, and sensor values of (b) ring finger for the alphabet 'M' in green and 'N' in red. (c) The plotting of hand motion sensor data for the alphabet 'I' and alphabet 'J'.

Table 2 for the first (see Fig. 1) and second (see Fig. 5) versions of the proposed system. The results clearly signified a large difference of 32.5% between the first and second versions of the system. The accuracy in the first version was low owing to similar patterns occurring among several of the signs, which caused negative classifications. The accuracy improved significantly after the pressure sensors were added to the system. However, there are still minor misclassifications in the second version of the system. An analysis indicated that the incorrect pattern recognitions for all subjects occurred more commonly between the letters "E" and "S." This is due to the significantly lower differences in the flexion values for subjects with a smaller hand size. Thus, the system misinterpreted the thumb region as a PR instead of a CR, or vice versa. A similar issue appeared for letters "M" and "N" as well, with an incorrectly identified region for the ring finger as illustrated in Fig. 6(b). It was noticed that even though the flexion sensor

TABLE I
VECTOR FOR FLEXION VALUES SEPARATED BY REGIONS

Alphabets	TB	IN	MD	RG	PK	PS1	PS2	HM
A	OR	CR	CR	CR	CR	O	O	X
B	PR	OR	OR	OR	OR	X	O	X
C	OR	PR	PR	PR	PR	X	O	X
D	PR	OR	CR	CR	CR	O	O	X
E	CR	CR	CR	CR	CR	O	O	X
F	PR	PR	OR	OR	OR	X	O	X
G	OR	OR	CR	CR	CR	O	O	X
H	OR	OR	OR	CR	CR	X	O	X
I	PR	CR	CR	CR	OR	O	O	X
J	PR	CR	CR	CR	OR	O	O	O
K	OR	OR	OR	CR	CR	X	X	X
L	OR	OR	CR	CR	CR	X	X	X
M	OR	PR	PR	PR	CR	O	O	X
N	OR	PR	PR	CR	CR	O	O	X
O	PR	PR	PR	PR	PR	X	O	X
P	OR	OR	OR	CR	CR	X	X	X
Q	OR	OR	CR	CR	CR	X	O	X
R	OR	OR	OR	CR	CR	O	X	X
S	PR	CR	CR	CR	CR	O	O	X
T	OR	PR	CR	CR	CR	O	O	X
U	PR	OR	OR	CR	CR	X	O	X
V	PR	OR	OR	CR	CR	X	X	X
W	PR	OR	OR	OR	CR	X	X	X
X	PR	PR	CR	CR	CR	X	X	X
Y	OR	CR	CR	CR	OR	O	O	X
Z	OR	OR	CR	CR	CR	X	X	O
NT	OR	OR	OR	OR	OR	X	X	X

TB: Thumb finger
IN: Index finger
MD: Middle finger
RG: Ring finger
PK: Pinky finger
PS1: Pressure Sensor 1
PS2: Pressure Sensor 2
HM: Hand in motion
OR: Open region
PR: Partially open/close region
CR: Close region
O: Yes
X: No

value of ring finger for alphabet “N” had higher value than the value for alphabet “M”, they are both fall into same PR region which produce the false classification between both alphabets.

Table 3 illustrated the comparison of first and second versions for the alphabet “R”, “U”, and “V”. The accuracy rate of signs recognition for alphabet “U” increased significantly from mean AC of 57.25% to 97.52% when two pressure sensors data are included for the classification. Likewise, the mean AC for alphabet “R” and “V” increased dramatically from 57.78% to 97.28% and 57.22% to 97.49% respectively. Even though the AC increased, but still, there was a slight misclassification between the letters “R” and “U,” where subjects with thicker fingers tended to touch the surface of the second pressure sensor for the letter “R”, e.g., subjects 4, 5, and 12. On the contradictory, subjects with thinner fingers did not tend to touch the surface of the second pressure sensor. Nevertheless, the other signs did not incur this same issue. The inclusion of the first pressure sensor surface showed significant differences for the signs between the letters “U” and “V.” Lastly, Fig. 6(c) demonstrated the differences of alphabet “I” and “J” in term of hand motion. The result indicated that there was no motion in pitch, roll, and yaw angles when performing gesture for the alphabet “I” throughout the time. Meanwhile, the action of performing the sign gesture for alphabet “J” showed dynamic changes of angular which are distinguishable from the alphabet “I”.

TABLE II
SUMMARY CLASSIFICATION RESULTS FOR SIGN RECOGNITION

Subject	Sample Size ((TP + TN) / Total)	AC (%)
1st version	425,736 / 648,000	65.7
2nd version	636,336 / 648,000	98.2

AC: Accuracy

TABLE III
COMPARISON OF CLASSIFICATION RESULTS FOR SIGN RECOGNITION FOR ALPHABET ‘R’, ‘U’ AND ‘V’

Alphabet \ Subject	R	U	V
1	65.1 / 97.2	66.2 / 97.5	65.8 / 98.5
2	51.2 / 96.4	52.1 / 97.1	50.5 / 98.2
3	53.4 / 97.4	52.8 / 98.6	53.5 / 97.2
4	56.1 / 95.5	55.2 / 96.4	55.3 / 97.5
5	61.5 / 95.8	60.8 / 97.5	61.1 / 96.7
6	62.8 / 97.9	61.8 / 97.8	60.5 / 96.9
7	60.8 / 96.4	61.1 / 98.4	61.0 / 97.5
8	69.5 / 98.7	68.4 / 96.7	67.5 / 98.5
9	51.6 / 97.5	50.8 / 97.2	51.1 / 96.9
10	52.7 / 98.6	51.8 / 97.8	50.8 / 97.6
11	53.8 / 99.5	52.8 / 99.4	53.1 / 98.2
12	54.8 / 96.4	53.2 / 95.8	54.2 / 96.1

/ – Accuracy rate for *1st version vs. (2nd) version in %

On the other hand, Table 4 shows a comparison of the proposed system with previous existing methods. In fact, there have been a number of studies over the past decades focusing on gesture recognition using image processing techniques. Many sophisticated algorithms have been proposed to distinguish gestures patterns, e.g., neural networks and decision trees. As semiconductor and electronic component technologies advance, hardware components of a smaller size, higher performance, and lower power consumption have created alternate methods for gesture recognition. Consequently, a data glove with small-scale sensors has been introduced. Recent studies have shifted focus from using image techniques to using a flex sensor, motion sensor, tilt sensor, or optical sensor for gesture recognition. Moreover, current studies are literally focusing solely on alphabets, numbers, and simple gestures. The average accuracy of sign or gesture recognition observed from these studies is over 90% for letters and numbers at the language level, but is slightly lower for more complicated gesture patterns. Conclusively, although the wearable device proposed in this study is currently only targeted at the alphabet level, its accuracy is significantly higher and it outperforms existing studies that have been proposed in recent decades.

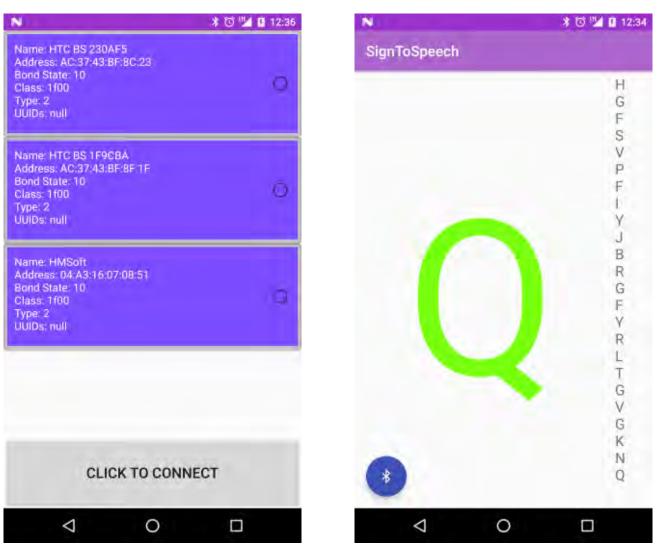
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2017.2779466, IEEE Sensors Journal

TABLE IV
COMPARISON OF SIGN LANGUAGE RECOGNITION SYSTEMS

Author, Year	Interfaces	Methods	Language Level	AC (%)
Shukor, 2014 [2]	Glove, tilt sensors	Template matching	Alphabets Numbers Gestures	95.0 93.3 78.3
Sriram, 2013 [3]	Glove, accelerometer sensors	Template matching	Alphabets	Not mentioned
Matiwade, 2016 [4]	Glove, flex sensors	Logics level	Alphabets	Not mentioned
Goval, 2013 [5]	Images	Keypoint Localization	Alphabets	95.0
Murakami, 1991 [7]	Data glove, images	Neural networks	Alphabets Words	98.0
Patil, 2014 [15]	Glove, flex sensors	Template matching	Alphabets	Not mentioned
Lu, 2016 [17]	Glove, motion sensors	Extreme learning machine	Numbers	91.2 (highest)
Xie, 2015 [19]	Ring, motion sensors	Segmentation algorithm	Basic Gestures Complex Gestures	98.9 97.2
Hsu, 2015 [20]	Pen, inertial sensors	Min-max template	2D digits 3D digits 2D English character 3D gestures	99.4 94.6 94.3 93.0 99.8
Wu, 2015 [29]	Wrist-worn, sEMG, inertial sensor	LibSVM	Words	95.9

B. Sign Interpretation Application

The classified sign gestures from the proposed smart wearable device is transmitted to the sign interpretation system. The sign interpretation system was built on an Android-based mobile device, a Google Nexus 6p [42]. First, the application searches for available Bluetooth devices and displays the search results through a card list, as shown in Fig. 7(a). Next, the selected Bluetooth device (HMSoft) initiates a wireless connection with the proposed wearable device. Once the connection is successfully established, the application will start to receive data in a text format, ranging from "A" to "Z," or in a neutral state, i.e., "NT," or as an invalid sign, i.e., "INV," as illustrated in Fig. 7(b). The received text is displayed at the center of the screen, whereas the past history of received texts is displayed on the right side of the screen with the newest text shown at the bottom of the list. Moreover, a text-to-speech service is also implemented in the application. The service converts the received text into an output, which is played back concurrently by the mobile device speaker. In this study, the Android-based sign interpretation application is merely utilized for receiving classified sign gestures from the smart wearable device and further displaying the results on the screen.



(a) (b)

Fig. 7. Sign interpretation system showing (a) Bluetooth search screen and (b) text received from the proposed wearable device through a Bluetooth connection, as well as a text-to-speech service converting the text into audible output played back simultaneously using the built-in mobile speaker.

V. CONCLUSION

In this study, we successfully designed and implemented a novel and smart wearable hand device as a sign interpretation system using a built-in SVM classifier. An Android-based mobile application was developed to demonstrate the usability of the proposed smart wearable device with an available text-to-speech service. The participating subjects gave a high rating to the proposed smart wearable sign interpretation system in terms of its comfort, flexibility, and portability. The device holders were 3D-printed using a flexible filament, and the same holders are able to fit different hand and finger sizes, thus eliminating the necessity of custom-made devices. Future work on the proposed smart wearable hand device will consider the design of a smaller sized printed circuit board, the inclusion of words and sentences at the sign language level, and instantly audible voice output components.

REFERENCES

[1] J. Wang and T. Zhang, "An ARM-based embedded gesture recognition system using a data glove," presented at the *26th Chinese Control and Decision Conf.*, Changsa, China, May 31 - June2, 2014.

[2] A. Z. Shukor, M. F. Miskon, M. H. Jamaluddin, F. A. Ibrahim, M. F. Asyraf and M. B. Bahar, "A new data glove approach for Malaysian sign language detection," *Procedia Comput. Sci.*, vol. 76, no. 1, pp. 60-67, Dec. 2015.

[3] N. Sriram and M. Nithiyanandham, "A hand gesture recognition based communication system for silent speakers," presented at the *Int. Conf. Human Comput. Interact.*, Chennai, India, Aug. 23-24, 2013.

[4] S. V. Matiwade and M. R. Dixit, "Electronic support system for deaf and dumb to interpret sign language of communication," *Int. J. Innov. Research Sci. Eng. Technol.*, vol. 5, no. 5, pp. 8683-8689, May 2016.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2017.2779466, IEEE Sensors Journal

[5] S. Goyal, I. Sharma and S. Sharma, "Sign language recognition system for deaf and dumb people," *Int. J. Eng. R. Technol.*, vol. 2, no. 4, pp. 382-387, Apr. 2013.
[6] S. P. More and A. Sattar, "Hand gesture recognition system using image processing," presented at *Int. C onf. Electrical, Electronics, Opt. Techniq.*, Chennai, India, Mar. 2016.
[7] K. Murakami and H. Taguchi, "Gesture recognition using recurrent neural network," in *Proc. SIGCHI Conf. Human Factors Comput. Syst.*, New York, USA, 1991.
[8] P. R. V. Chowdary, M. N. Babu, T. V. Subbareddy, B. M. Reddy and V. Elamaran, "Image processing algorithms for gesture recognition using matlab," presented at *Int. Conf. Adv. Comm. Control Comput. Technol.*, Ramanathapurum, India, Jan. 2015.
[9] T. Khan and A. H. Pathan, "Hand gesture recognition based on digital image processing using matlab," *Int. J. Sci. Eng. R.*, vol. 6, no. 9, pp. 338-346, Sep. 2015.
[10] J. Siby, H. Kader and J. Jose, "Hand gesture recognition," *Int. J. Innov. Technol. R.*, vol. 3, no. 2, pp. 1946-1949, March 2015.
[11] L. Lamberti and F. Camastra, "Real-time hand gesture recognition using a color glove," presented at *Int. Conf. Image Analy. Process.*, Ravenna, Italy, Sep. 2011.
[12] Y. Iwai, K. Watanabe, Y. Yagi and M. Yachida, "Gesture recognition using colored gloves," in *Proc. 13th Int. Conf. Pattern Recog.*, Vienna, Austria, Aug. 25-29, 1996.
[13] S. P. Dawane and H. G. A. Sayyed, "Hand gesture recognition for deaf and dumb people using gsm module," *Int. J. Sci. R.*, vol. 6, no. 5, pp. 2226-2230, May 2017.
[14] C. Preetham, G. Ramakrishnan, S. Kumar and A. Tamse, "Hand talk-implementation of a gesture recognizing glove," presented at *2013 Texas Instruments India Educators' Conf.*, Bangalore, India, Apr. 4-6, 2013.
[15] K. Patil, G. Pendharkar and G. N. Gaikwad, "American sign language detection," *Int. J. Sci. R. Pub.*, vol. 4, no. 11, pp. 1-6, Nov. 2014.
[16] J. Kim, N. D. Thang and T. Kim, "3-D hand motion tracking and gesture recognition using a data glove," presented at *IEEE Int. Symp. Industrial Elec. 2009*, Seoul, South Korea, Jul. 5-8, 2009.
[17] D. Lu, Y. Yu and H. Liu, "Gesture recognition using data glove: an extreme learning machine method," in *Proc. 2016 IEEE Int. Conf. Robotics and Biomimetics*, Qingdao, China, Dec. 3-7, 2016.
[18] J. Lm, D. Lee, B. Kim, I. Cho and J. Ryou, "Recognizing hand gestures using wrist shapes," presented at *2010 Digest Technical Papers Int. Conf. Consumer Elec.*, Las Vegas, USA, Jan. 9-13, 2010.
[19] R. Xie, X. Sun, X. Xia and J. Cao, "Similarity matching-based extensible hand gesture recognition," *IEEE Sensors J.*, vol. 15, no. 6, pp. 3475-3483, Jun. 2015.
[20] Y. L. Hsu, C. L. Chu, Y. J. Tsai and J. S. Wang, "An inertial pen with dynamic time warping recognizer for handwriting and gesture recognition," *IEEE Sensors J.*, vol. 15, no. 1, pp. 154-163, Jan. 2015.
[21] L. Yin, M. Dong, Y. Duan, W. Deng, K. Zhao and J. Guo, "A high-performance training-free approach for hand gesture recognition with accelerometer," *Mult. Tools App.*, vol. 72, no. 1, pp. 843-864, Sep. 2014.
[22] J. Galka, M. Masior, M. Zaborski, K. Barczewska, "Inertial motion sensing glove for sign language gesture acquisition and recognition," *IEEE Sensors J.*, vol. 16, no. 16, pp. 6310-6316, Aug. 2016.
[23] X. Cai, T. Guo, X. Wu and H. Sun, "Gesture recognition method based on wireless data glove with sensors," *Sensor Letters*, vol. 13, no. 2, pp. 134-137, Feb. 2015.
[24] K. Liu, C. Chen, R. Jafari and N. Kehtarnavaz, "Fusion of inertial and depth sensor data for robust hand gesture recognition," *IEEE Sensors J.*, vol. 14, no. 6, pp. 1898-1903, Jun. 2014.
[25] K. W. Kim, M. S. Lee, B. R. Soon, M. H. Ryu and J. N. Kim, "Recognition of sign language with an inertial sensor-based data glove," *Technol. Health Care*, vol. 24, no. 1, pp. 223-230, 2016.
[26] L. Sousa, J. M. F. Rodrigues, J. Monteiro, P. J. S. Cardoso and R. Lam, "GyGSLA: a portable glove system for learning sign language alphabet," presented at *Int. Conf. Universal Access in Human-Comput. Interact.*, Toronto, ON, Canada, Jul. 17-22, 2016.
[27] N. Caporusso, L. Biasi, G. Cinquepalmi, G. F. Trotta and A. Brunetti, "A wearable device supporting multiple touch- and gesture-based languages for the deaf-blind," presented at *Int. Conf. Appl. Human Factors and Ergonomics*, LA, CA, USA, Jul. 17-21 2017.
[28] Z. Yu, X. Chen, Q. Li, X. Zhang and P. Zhou, "A hand gesture recognition framework and wearable gesture-based interaction prototype for mobile devices," *IEEE Trans. Human-Mach. Syst.*, vol. 44, no. 2, pp. 293-299, Apr. 2014.
[29] J. Wu, Z. Tian, L. Sun, L. Estevez and R. Jafari, "Real-time American sign language recognition using wrist-worn motion and surface emg sensors," presented at *IEEE 12th Int. Conf. Wearable Implantable Body Sensor Net.*, Cambridge, USA, Jun. 9-12, 2015.
[30] J. Wu, L. Sun and R. Jafari, "A wearable system for recognizing American sign language in real-time using imu and surface emg sensors," *IEEE J. Biomedic. Health Info.*, vol. 20, no. 5, pp. 1281-1290, Sept. 2016.
[31] Flex sensor 4.5 inches. Available: https://www.sparkfun.com/products/8606, Accessed on: Aug. 7, 2017.
[32] Flex sensor 2.2 inches. Available: https://www.sparkfun.com/products/10264, Accessed on: Aug. 7, 2017.
[33] Adafruit BNO055 absolute orientation sensor. Available: https://learn.adafruit.com/adafruit-bno055-absolute-orientation-sensor/overview, Accessed on: Aug. 7, 2017.
[34] Arduino Pro Mini 328 – 5V/16 MHz. Available: http://www.devicemart.co.kr/1287163, Accessed on: Aug. 7, 2017.
[35] Bluetooth 4.0 HM-10 BLE module, Available: http://mechasolution.com/shop/goods/goods_view.php?goodsno=330423&category=132024, Accessed on: Aug. 7, 2017.
[36] Lithium ion polymer battery – 3.7v 500mAh. Available: https://www.adafruit.com/product/1578, Accessed on: Aug. 7, 2017.
[37] B. G. Lee and W. Y. Chung, "Wearable glove-type driver stress detection using a motion sensor," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 7, pp. 1835-1844, Jul. 2017.
[38] C. C. Lip, D. A. Ramli (2012), "Comparative study on feature, score and decision level fusion schemes for robust multibiometric systems," *Frontiers in Computer Education. Advances in Intelligent and Soft Computing*, vol. 133, pp. 941-948, doi: 10.1007/978-3-642-27552-4_123.
[39] Flexiforce pressure sensor – 25lbs. Available: http://www.devicemart.co.kr/1289308, Accessed on: Aug. 9, 2017.
[40] C. L. Lim, C. Rennie, R. J. Barry, H. Bahramali, I. Lazzaro, B. Manor and E. Gordon, "Decomposing skin conductance into tonic and phasic components," *Int. J. Psychophys.*, vol. 25, no. 2, pp. 97-109, Feb. 1997.
[41] S. Arlot and A. Celisse, "A survey of cross-validation procedures for model selection," *Statistics Theory*, vol. 4, no. 1, pp. 40-79, Jul. 2009.
[42] Google Nexus 6p. Available: https://www.google.com/nexus/6p/, Accessed on: Aug. 9, 2017.

Boon Giin Lee (M'14) received his B.I.T. degree in information technology from Multimedia University, Melaka, Malaysia, in 2007, his M.S. degree in electronic engineering from Dongseo University, Busan, South Korea, and his Ph.D. degree in electronic engineering from Pukyong National University, Bu san, South Korea.

From 2012 to 2013, he was a researcher with the Ubiquitous Sensor Network Laboratory in Pukyong National University, Busan, South Korea. He has been an assistant professor with the Electronic Engineering Department, Keimyung University, Daegu, South Korea. His research interests include the ubiquitous healthcare, machine learning, signal processing, mobile application development, human-computer interaction, sensors network and multimedia design.

Su Min Lee is currently study her bachelor degree in Electronic Engineering from Keimyung University, Daegu, Republic of Korea. She had been receiving prizes from competitions such as Wearable Computing Contest organized by KAIST. Her research interests include AR design, human-computer interaction and circuit design.