Feasibility of Wrist-worn, Real-time Hand and Surface Gesture Recognition via sEMG and IMU Sensing

Shuo Jiang, *Student Member, IEEE*, Bo Lv, *Student Member, IEEE*, Weichao Guo, *Student Member, IEEE*, Chao Zhang, Haitao Wang, Xinjun Sheng, *Member, IEEE*, and Peter B. Shull*, *Member, IEEE*

***Abstract*—While most wearable gesture recognition approaches focus on the forearm or fingers, the wrist may be a more suitable location for practical use. We present the design and validation of a real-time gesture recognition wristband based on surface electromyography (sEMG) and inertial measurement unit (IMU) sensing fusion, which can recognize 8 air gestures and 4 surface gestures with 2 distinct force levels. Ten healthy subjects performed an initial gesture recognition experiment, followed by a second experiment 1 hour later and a third experiment 1 day later. Classification accuracies for the initial experiment were 92.6% and 88.8% for air and surface gestures, respectively, and there were no changes in accuracy results during testing 1 hour and 1 day later (p>0.05). These results demonstrate the feasibility of wrist-based gesture recognition paving the way for potential future integration in to a smart watch or other wrist-worn wearable for intuitive human computer interaction.**

***Index Terms*—Electromyography(EMG), gesture recognition, human computer interaction, wearable sensors.**

I. Introduction

DESPITE technical advances in virtual and augmented reality, the primary way people interact with computers, home electronics, and industrial equipment is by physically pressing keyboards and touchpads. Intelligent, wearable devices could enable humans to interact with electronics and industrial equipment in a safer, more natural way through automated hand gesture recognition. Recent research has shown that hand gesture recognition can enable a wide range of interactive activities such as remotely controlling a home service robot [1], interacting with smart phones [2], and even as an aid to help the deaf communicate more effectively with people who don't understand sign language [3]. Real-time hand gesture recognition could also enhance virtual and augmented reality if combined with systems like Oculus Rift or Microsoft HoloLens [4], [5].

Most wearable approaches to hand gesture recognition utilize forearm-worn sensing. Zhang *et al.* [6] proposed a novel system which utilized 4 EMG sensors and 3D accelerometer to realize 18 kinds of gestures recognition in real time virtual game control. Wolf *et al.* [7] incorporated 16 sEMG (surface Electromyography) channels and an IMU (Inertial Measurement Unit) into a sleeve covering most parts of forearm to decode 16 discrete gestures for controlling a robot arm and hand to move and grasp objects. Lu *et al.* [2] developed a forearm-worn prototype which utilized 4 sEMG sensors and an IMU. This system identified multiple distinct gestures to perform common mobile phone operations such as opening and exiting an application or dialing a phone number. Guo *et al.* [8] combined sEMG and near-infrared spectroscopy for prosthetic manipulation. Thalmic lab commercialized a gesture control armband, Myo, consisting of 8 sEMG sensors and an IMU worn on the forearm which is capable of identifying several predefined gestures [9]. Besides sEMG, Silva *et al.* [10] used forearm mechanomyography to detect muscle mechanical vibrations for classifying elbow flexion and extension. Jung *et al.* [11] showed that combining pressure sensing and an air bladder enabled robust detection of forearm muscle activity to classify 6 distinct hand gestures. Sikdar *et al.* [12] demonstrated the feasibility of using ultrasound imaging of the forearm to detect individual finger movements. A major drawback of forearm-worn devices is that they can be uncomfortable and unnatural as people typically are not accustomed to wearing a band around the forearm in daily life.

In contrast, wrist-worn devices may be a more natural and comfortable approach to hand gesture recognition. Kim *et al.* [13] developed a gesture watch with an array of infrared prox-

Manuscript received June 23, 2017; revised September 25, 2017; accepted Nov 15, 2017. Paper no. TII-17-1332.

This work was supported by the National Natural Science Foundation of China under Grant 51505283.

S. Jiang, B. Lv, W. Guo, X. Sheng and P. B. Shull are with State Key Laboratory of Mechanical System and Vibration, School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail:jiangshuo@sjtu.edu.cn; brucelv@sjtu.edu.cn; A498009151@sjtu.edu.cn; xjsheng@sjtu.edu.cn; pshull@sjtu.edu.cn).

C. Zhang is with Samsung R&D Institute of China, Beijing 100028, China (e-mail: c0502.zhang@samsung.com).

H. Wang is with AR/VR lab in Huawei, Beijing 100095, China (e-mail:wanghaitao15@huawei.com).

*To whom correspondence should be addressed. E-mail:pshull@sjtu.edu.cn

(a) (b)

Fig. 1. The wristband prototype (a) contains four sEMG sensors and one IMU module and (b) is worn just below the distal ends of the radius and ulna.

imity sensors to detect large scale hand movements.Dementyev *et al.* [14] proposed a WristFlex system which employed 15 force sensitive resistors around the wrist to distinguish pinching gestures. Kim *et al.* [15] put cameras on the wrist to directly track a series of three dimensional hand poses. Oyama *et al.* [16], [17] used 4 sEMG sensors on the wrist to classify 7 distinct wrist movements. McIntosh *et al.* [18] combined wrist sEMG and pressure sensing to classify simple hand gestures and reported that the combination of both sensing modalities resulted in higher accuracy than either individual sensing modality alone. While the above approaches to wrist-worn gesture classification are promising, several significant problems remain: infrared proximity solutions can only track large hand movements not individual fingers, the resolution of force sensitive resistors is relatively low making accurate classification difficult, and cameras are sensitive to light and background environment colors making it difficult to use cameras outside of controlled laboratory settings. In addition, there is currently no wrist-worn solutions for detecting surface pressing gestures at multiple force levels.

Existing problems mainly involve the "wearable" and "intuitive" aspects. For "wearable" problems, the current issue is that forearm-worn solutions like Myo have relatively high accuracy but are not convenient in daily life and current wrist-worn solutions like force resistive sensor do not have high resolution. This work addresses the wearable issue with a wrist-worn solution and achieved relatively high accuracy through sEMG and IMU sensor fusion. For "intuitive" problems, previous research has mainly focused on air gesture recognition. However, humans frequently interact with digital devices by touching and pressing hard surfaces in daily life and force is an important aspect of this type of interaction. Therefore, this work validated surface gestures with different force levels by sEMG and IMU sensor fusion. This surface force dimension interaction cannot be achieved by computer vision or force resistive sensing, because they are based on the geometry of hand or wrist shape change which do not have direct relationship with muscle's inner force.

This research is innovative in that: 1) the device is located at the wrist, as opposed to the forearm, and 2) classification is for surface gestures as well as air gestures. Related to the first aspect, a wrist-worn approach could enable widespread adoption and practical use of gesture recognition as it could potentially be incorporated into smart watch type devices. For long-term, wearable use throughout the day, humans would likely be more willing to wear wrist-worn devices than forearm-worn devices, as they are usually accustomed to wearing watches or jewelry on the wrist. Related to the second aspect, previous hand gesture recognition research typically only classifies air gestures. This initial work in surface gesture recognition could enable any hard surface to potentially become an interactive touch pad.

The purpose of this work is to assess the feasibility of wrist-worn hand and surface gesture recognition through sEMG and IMU sensor fusion (Fig. 1). We hypothesize that wrist-worn sEMG and IMU sensing can be used to accurately classify 8 air gestures and 4 surface pressing gestures at 2 different force levels, and that accuracy levels will be retained one hour and one day after initial testing. An initial pilot study was first conducted to determine forearm and wrist sEMG signal strength. The wrist worn system prototype and hardware design are then presented followed by a description of the classification and calibration algorithms used for gesture recognition. Experimental validation testing and results are then presented followed by discussion and future implications.

II. PILOT STUDY TO QUANTIFY WRIST sEMG SIGNAL STRENGTH

Before the prototype design, a pilot study on the comparison of forearm sEMG with wrist sEMG was performed, because sEMG signals are generally considered small compared to the forearm. Eight total surface EMG sensors (Trigno^{TM} Wireless EMG, Delsys) were used; four sensors were placed around the wrist and the other four were placed around the forearm (flexor digitorum superficialis, beachioradialis, flexor carpi ulnaris, extensor digitorum). Raw data was recorded, and comparison indexes were calculated.

We calculated the Root Mean Square (RMS) of each channel as the index of signal intensity, and define Forearm to Wrist sEMG Signal Strength Ratio (FWR) as the comparison index of sEMG signals amplitude of wrist and forearm.

$$FWR = \frac{\sqrt{RMS^2_{ForearmSignal} - RMS^2_{ForearmNoise}}}{\sqrt{RMS^2_{WristSignal} - RMS^2_{WristNoise}}} \quad (1)$$

In addition, the signal-to-noise ratio was calculated as follows:

$$SNR = 20 \log \frac{RMS_{activated}}{RMS_{unactivated}} \quad (2)$$

Four subjects (2 male, 2 female) performed this pilot experiment. Every subject performed air gestures sets and surface gestures sets each for 10 trials. A representative trial of raw data is shown in Appendix Fig. 1. The average results of four subjects are shown in Appendix Table 1. Results showed that the FWR was close to or less than 1 for most gestures indicating that there was not a great difference in sEMG between the wrist and forearm, and there was no statistically significant difference in wrist and forearm signal amplitude ($p = 0.173$). Also, in general, wrist sEMG had more noise than forearm sEMG.

III. WRISTBAND PROTOTYPE DESIGN

A custom wristband prototype was created that combines sEMG and IMU signals to classify air and surface gestures (Fig. 1). The primary benefits of combining sEMG and IMU signals are that sEMG is well suited for distinguishing specific finger movements while IMU can precisely estimate overall hand movements [2], [19]. Two important design criteria were that the prototype should be compact enough to fit around the wrist and should be completely standalone, unwired, and untethered during operational tasks to allow natural hand and finger movements. To satisfy these criteria, we designed a prototype consisting of 4 separate modules that connect wirelessly to a computer (Fig. 2). Raw data is collected on

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/TII.2017.2779814, IEEE Transactions on Industrial Informatics

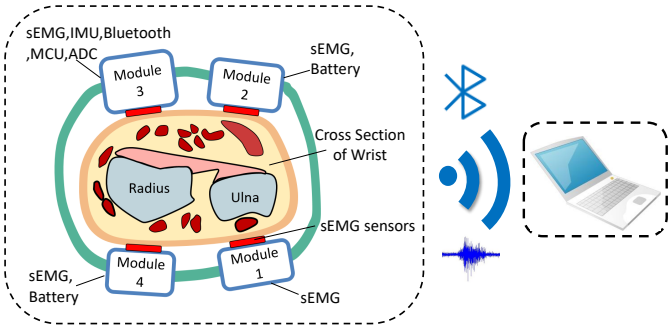


Fig. 2. sEMG/IMU wristband prototype and system framework.

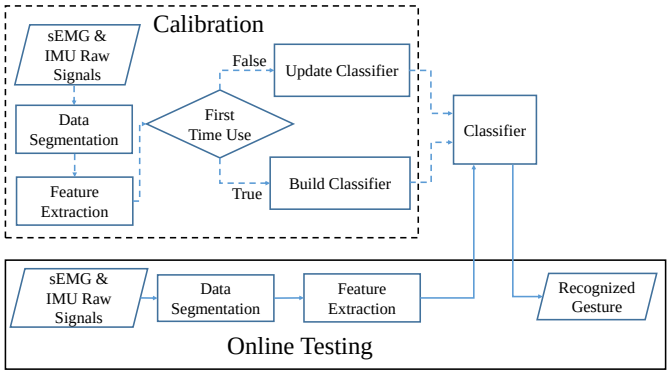


Fig. 3. Air and surface hand recognition algorithm flow chart.

the wristband prototype and then transmitted to the computer where hand gesture recognition is performed in real-time. All modules contain a sEMG sensor, while modules 2 and 4 also contain a battery, and module 3 also contains an IMU, a Bluetooth module, an MCU (Microcontroller Unit), and an ADC (Analog-to-Digital Converter). Analog signals are digitalized and then transmitted to the computer via Bluetooth. The four modules are encapsulated in 3D printed boxes, which are connected together via elastic straps to enable users of different wrist sizes to wear the device. The size of module 1 is 4.0 $\times$ 1.9 $\times$ 0.9 cm, the size of modules 2 and 4 are 4.0 $\times$ 2.3 $\times$ 1.5 cm, and the size of module 3 is 4.4 $\times$ 3.0 $\times$ 1.5 cm. The overall weight of the wristband prototype is 56 g.

sEMG signals were captured with custom-made dry electrodes. Dry electrodes were used instead of wet electrodes, because wet electrodes require electrolyte gel which is inconvenient and can cause skin irritation [20]. Each channel consists of 3 gilded copper electrodes. The middle electrode is the reference and the outer two electrodes are for differential amplification. The custom sEMG sensors were previously shown to perform comparably with other commercial devices (e.g. Delsys and Biometrics) as quantified by the commonly used metrics of input impedance, input-referred noise and Signal Noise Ratio [8]. Raw signal data from a representative trial is presented in Appendix Fig. 2.

IV. AIR AND SURFACE HAND GESTURE RECOGNITION ALGORITHM

The air and surface hand gesture recognition algorithm is comprised of Calibration and Online Testing (Fig. 3). The purpose of Calibration is to build and update the classifier based on labeled offline data. The classifier is initially built during the First Time Use and updated in all subsequent uses. Both Calibration and Online Testing begin with the identical sequence of acquiring sEMG and IMU Raw Signals, segmenting data into windows, and then using the segmented data to calculate and extract features. Online Testing interprets unlabeled real-time data based on the current Classifier to determine the Recognized Gesture.

A. Data Segmentation

Data segmentation is used to divide raw data into windows to calculate features over a series of points. Segment length is important in real life applications because segments that are too short can result in large variance within individual features, and segments that are too long can hinder real-time operation [21]. In general, delays of 300ms or less are acceptable for real-time control [22], and we thus choose the segment length of 200ms based on previous real-time control applications [23]. Data segmentation is updated every 100ms, a time length chosen to be shorter than the window length and longer than the processing time and thus the decision time was every 100 ms.

B. Feature Extraction

Given that raw data can be redundant and may not have direct physical meaning at a single point, features are designed to represent the relevant structure of raw data. Based on previous research showing the effectiveness and low computational expense of time domain features which do not require matrix transformations like frequency domain features [21], [24], we chose the following time domain features: mean absolute value, zero crossings, slope sign changes, waveform length for sEMG signals and mean absolute value and wavelength for IMU signals. Mean absolute values contain information about signal strength and amplitude, and zero crossings and slope sign change are simple features representing frequency information. The wavelength feature reflects waveform complexity and contains information about waveform duration, amplitude and frequency. Mean absolute value, zero crossings, slope sign changes, and waveform lengths were calculated as follows:

$$Mean\ absolute\ value: \bar{x}_i = \frac{1}{L}\sum_{k=l}^{L}|x_k|, for\ i = 1, ..., I \quad (3)$$

$$Zero\ crossing: x_k x_{k+1} < 0 \quad (4)$$

$$Slope\ sign\ changes: (x_k - x_{k-1})(x_{k+1} - x_k) < 0 \quad (5)$$

$$Waveform\ length: l_0 = \sum_{k=2}^{L}|x_k - x_{k-1}| \quad (6)$$

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/TII.2017.2779814, IEEE Transactions on Industrial Informatics

where L is the length of samples (in this paper, L is 200), x_k is the k th sample in segment i, and I is the total number of segments.

C. Classifier

After extracting features, a classifier translates features to recognized gestures. Our system is designed for real-time application and thus computational efficiency is very important. In previous research, LDA did not sacrifice accuracy [25]–[27] with fast speed and simplicity [21], [28] for training. For this real-time classification application, it is important that the classifier should be both effective and efficient. We thus chose to use LDA as it has been previously shown to be computationally efficient without compromising accuracy [21] and often more robust than other approaches [29]. To illustrate this, Principle Component Analysis (PCA) was used to determine three dominant features resulting from LDA which were then depicted to visually demonstrate separation among features (Fig. 4). LDA is based on the Bayesian decision rule and Gaussian assumption [30]. The discriminant function of LDA is:

$$g_c(x) = \mu_c^T \Sigma^{-1} x - \frac{1}{2} \mu_c^T \Sigma^{-1} \mu_c \tag{7}$$

where μ_c is the mean vector of training samples of class c, Σ is the pooled sample covariance matrix, and x is the current feature vector.

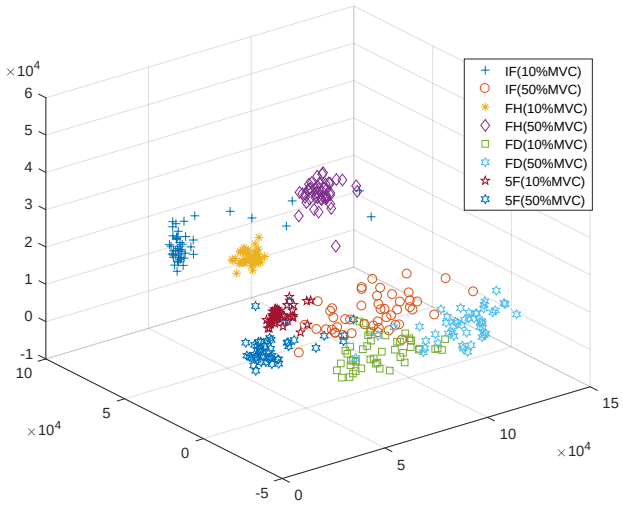


Fig. 4. An example of clustering of three dominant features along each axis during gesture classification.

D. Calibration

Unlike offline data processing methods such as cross-validation that employs half or more data sets to train a classifier, online recognition requires first training a classifier using limited datasets and then performing online classification. Therefore, for first time use, we employed 3 trials of data sets to build a classifier for the following online recognition. However, a primary problem for hand gesture classification is that sEMG signals may differ each time the wristband is taken off and put back on primarily because of electrode shift and variations in electrode conductivity [31]. Calibration is needed to address this problem. Additionally, prolonged periods of calibration are impractical for daily use while an absence of any calibration generally leads to deteriorating levels of classification accuracy. We therefore seek a calibration method requiring acceptable time that can maintain high classification accuracy.

In spite of nonstationary characteristics of sEMG signals, the first-time training data can be still of great significance for next-time use because the muscles used for certain gestures are the same. This idea led us to make full use of long time training history data and combine it with short time current training data to design a relative robust classifier. According to the discriminant function of LDA algorithm, the results are determined by mean vector and covariance matrix. Therefore we can update the mean vector and covariance matrix based on historical data and new current data via the following:

$$\mu_c = (1-r)\bar{\mu}_c + r \sum_{k=1}^{p} \hat{\mu}_c^k \tag{8}$$

$$\Sigma_c = (1-r)\bar{\Sigma}_c + r \sum_{k=1}^{p} \hat{\Sigma}_c^k \tag{9}$$

where $\bar{\mu}_c$ and $\bar{\Sigma}_c$ are the current mean vector and covariance matrix. p is the number of history training model. We choose the weighting factor, r, to be 0.5 based on previous work [32], [33].

V. EXPERIMENTAL VALIDATION

A. Experimental Protocol

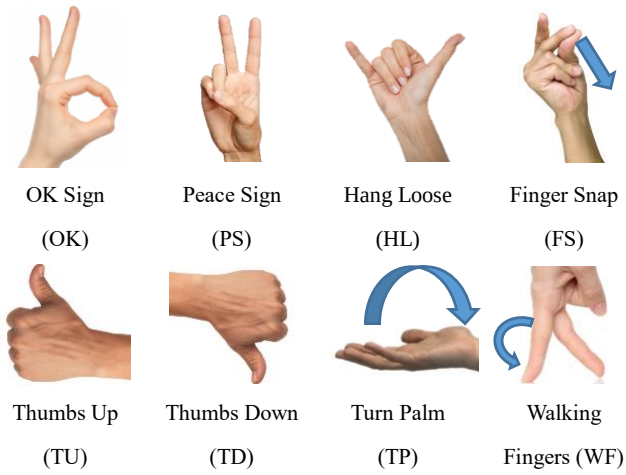


Fig. 5. Air Gestures. Blue arrows indicate the direction of motion. Gestures without arrow are stationary poses.

To assess the feasibility of the wristband prototype, gesture recognition testing was performed with air gestures and surface gestures. Air gestures refer hand poses and movements

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/TII.2017.2779814, IEEE Transactions on Industrial Informatics

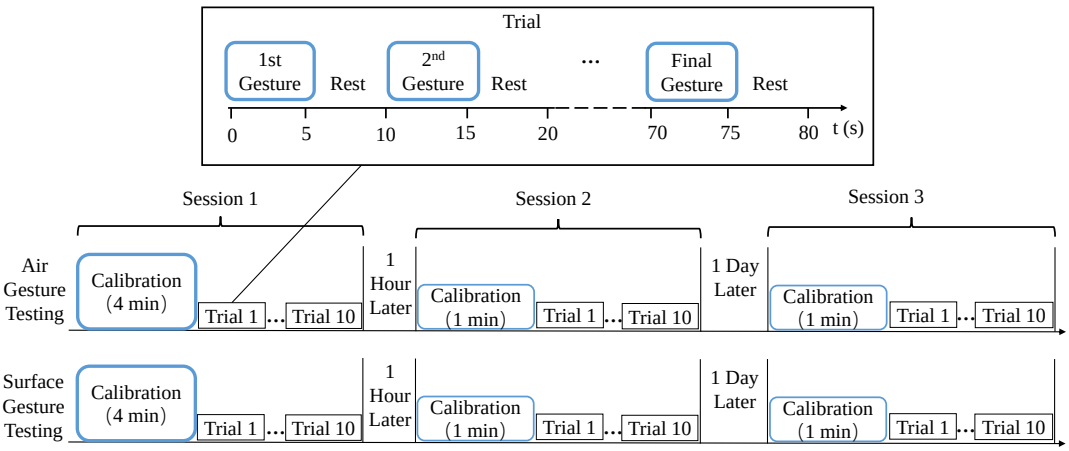


Fig. 7. Experimental protocol.

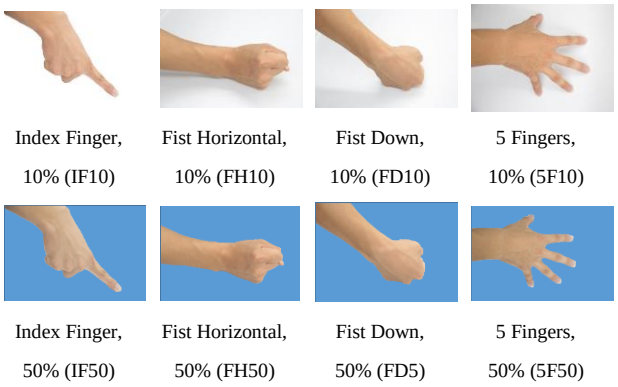


Fig. 6. Surface gestures. Percentages indicate the level of maximum voluntary contraction (MVC).

performed in the air without contacting any objects, similar to sign language. Humans often use air gestures to communicate [34] and thus may also be a natural way to interact with machines and intelligent devices. In this study, we choose the following air gestures because they are intuitive and commonly used in daily life: OK Sign, Peace Sign, Hand Loose, Finger Snap, Thumbs Up, Thumbs Down, Turn Palm, Walking Fingers (Fig. 5). Surface gestures involve finger or hand contact with a surface such as a table or tablet screen. Surface gestures through a wrist-worn device can enable users to interact through any existing hard surface, and the contact force magnitude against the surface can add an additional dimension of interaction. For example, when controlling a rolling robot, finger pressing force magnitude could be directly mapped to robot speed. In our design, we thus classify different surface gestures at distinct force levels. In this study, we choose the following surface gestures: Index Finger contacting the surface, Fist Horizontal contacting the surface, Fist Down contacting the surface, and 5 Fingers contacting the surface (Fig. 6). Also, each of these four surface gestures was classified at either 10% or 50% of the maximum voluntary contraction level.

Ten healthy subjects (8 male/2 female, 23.6 $\pm$ 1.0 years) participated in this study in accordance with the Declaration of Helsinki. The experiment involved 3 testing sessions spaced 1 hour and 1 day apart, and each session included air gesture testing and surface gesture testing (Fig. 7). Prior to testing, alcohol cotton was lightly rubbed on subjects skin to enhance electrical conductivity. All force levels were determined subjectively by each individual. Prior to collecting data for surface gestures, subjects were asked to perform 100%, 50%, and 10% pressing force levels to get familiar with the feeling, and then in the experiment, they performed these corresponding force levels according to their own feeling. During Session 1, subjects performed an initial calibration consisting of 3 trials to build the classifier (Fig. 3). During training trials, each gesture image was shown on the screen for 5 seconds to prompt subjects to perform the corresponding gesture, and this was followed by 5 seconds of rest (Fig. 7). Subjects then performed 10 testing trials. During testing trials, subjects were instructed to perform the desired target gesture as shown on the screen. After Session 1, subjects took off the wristband and rested for 1 hour. During Session 2, subjects first performed a short 1 min calibration trial, and data were combined with previous training data via equations (8) and (9) to update the classifier. Then ten testing trials were performed as in Session 1 and the results and online recognition rates were stored. One day later, subjects performed Session 3, which was identical to Session 2.

B. Data Analysis

To further analyze sEMG and IMU contributions to classification accuracy and the effect of classifying all surface and air gestures simultaneously, 10-fold cross validation was applied to session 1's data to determine classification accuracy of sEMG only, IMU only, and then the combination of both air and surface gestures classified simultaneously. One way analysis of variance (ANOVA) was performed to assess whether there were differences in classification accuracy between sessions, and to determine if there was a difference between overall air gestures classification accuracy and overall surface gestures classification accuracy. In the case when there was a difference, Tukeys procedure was used for post hoc analysis. Classification accuracy was defined as the number

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/TII.2017.2779814, IEEE Transactions on Industrial Informatics

of correctly identified gestures divided by total target gestures during each trial or session. Statistical significance was set to p=0.05. Also, confusion matrices were compiled to detail misclassified air and surface gestures.

VI. Results

To validate the feasibility of our proposed system and algorithm, 10-fold cross validation were first conducted and showed 94.6% accuracy for air gestures and 87.3% for surface gestures, respectively. Also, 10-fold cross validation results showed for that for air gestures, IMU only accuracy was 93.8%, wrist sEMG only was 75.7% and combined was 94.6%. For surface gestures, IMU only accuracy was 78.8%, wrist sEMG only was 74.3% and the combined was 87.4%. 10-fold offline cross validation of LDA showed 88.8% for all 16 gestures including air and surface gestures with 2 different force levels. Real-time classification accuracy for air gestures was 92.6%(4.2%), 91.7%(3.5%), 91.6%(3.4%) for Session 1, 2, and 3, respectively and for surface gestures was 88.8%(3.0%), 86.1%(4.6%), 86.4%(6.0%) for Session 1, 2, and 3, respectively (Fig. 8), where errors are listed as mean (one standard deviation). There was no difference in classification accuracy among testing sessions for air gestures (p=0.84) or for surface gestures (p=0.33). Overall, classification accuracy was higher for air gestures than for surface gestures (p<0.001). The most accurately classified air gesture was WF, and the least accurately classified air gesture was OK (Table 1). The most accurately classified surface gesture was FH10, and the least accurately classified surface gesture was IF50 (Table 1). For air gestures, among all gestures and all sessions the lowest classification accuracy was 80.1% (Session 2, OK) and the highest was 98.4% (Session 2, TP), and for surface gestures, among all gestures and all sessions the lowest classification accuracy was 78.2% (Session 3, IF50) and the highest was 98.2% (Session 2, FH1) (Table 1).

Air gestures were most commonly misclassified as Thumbs Up and least commonly misclassified as Thumbs Down (Table 2), and surface gestures were most commonly misclassified as Fist Down 10% and least commonly misclassified as Index Finger 50% (Table 3). For air gestures, the most frequent misclassification was OK Sign misclassified as Hang Loose (error rate = 6.9%), and for surface gestures the most frequent misclassification was FD50 misclassified as FD10 (error rate =10.0%). The Thumbs Down air gesture was never confused with the Peace Sign or Turn Palm air gestures, and the 5F10 surface gesture was never confused with the FH50 surface gestures.

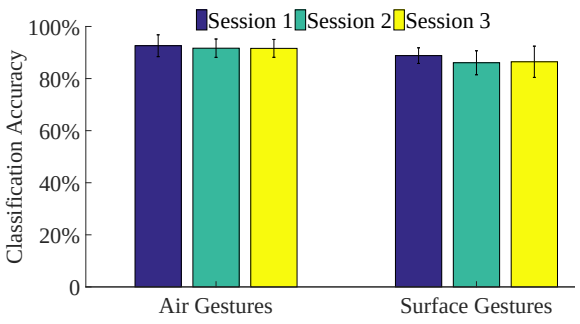


Fig. 8. Classification accuracy over time. There were no differences in classification accuracy between for either air or surface gestures. The time between Session 1 and Session 2 was 1 hour, and the time between Session 2 and Session 3 was 1 day.

TABLE I
Classification accuracy for each gesture and each session.

Air Gesture	OK	PS	HL	FS	TU	TD	TP	WF	AVG
Session 1	83.1	92.5	86.8	89.0	97.0	96.8	98.0	97.3	92.6
Session 2	80.1	91.4	88.3	89.3	91.5	97.1	98.4	97.1	91.7
Session 3	88.8	92.3	85.1	81.6	95.5	96.6	95.2	97.7	91.6
AVG	84.0	92.1	86.7	86.6	94.7	96.8	97.2	97.4	92.0
Surface Gesture	IF10	IF50	FH10	FH50	FD10	FD50	5F10	5F50	AVG
Session 1	85.7	87.1	95.4	87.0	91.3	85.7	88.8	89.4	88.8
Session 2	90.6	81.2	98.2	83.4	79.0	83.7	87.1	85.3	86.1
Session 3	86.5	78.2	97.8	87.7	80.5	82.9	91.8	86.2	86.4
AVG	87.6	82.2	97.1	86.0	83.6	84.1	89.2	87.0	87.1

VII. Discussion

This study presented design and feasibility testing of a real-time wireless wristband for gesture recognition. Results demonstrated the feasibility of wrist-worn sEMG and IMU sensing for eight air and four surface gestures at two distinct force levels.

Detailed comparison of our system with previous study and specifications of our custom made sensors are also compared with commercial products in Table 4. Wrist-worn classification results via sEMG and IMU sensing in this paper were in general comparable or slightly higher than other wrist-worn approaches. Dementyev *et al.* [14] employed force sensitive resistors around wrist to detect tendon movements and achieved 80.5% real-time accuracy for 5 finger pinch gestures. Kato *et al.* [35] used a vibration actuator to generate sweep signals and microphones to receive bone-conducted sound and classify gesture via different power spectral density which realized 88% accuracy for 7 gestures. Rekimoto *et al.* [36] measured wrist shape change to recognize 2 finger gestures via capacitive sensor. The reason for the better performance of our approach in some cases may be because sEMG measures intrinsic signals of muscle activity directly related to gestures while wrist shape change or external force against wristband does not have significant and obvious relationships with different gestures.

Despite common knowledge that sEMG signal amplitudes are significantly higher at the muscle belly than in the tendon areas [37], classification results from this study, involving sEMG sensing at the muscle tendons of the wrist, were comparable to other research involving sEMG sensing at the muscle belly of the forearm [6]. This may be because sEMG sensing at specific points on the wrist typically involves sEMG contributions of multiple muscles due to the high density

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/TII.2017.2779814, IEEE Transactions on Industrial Informatics

TABLE II
CONFUSION MATRIX FOR AIR GESTURES FOR SESSION 1.

		Classified Gesture							
		OK Sign	Peace Sign	Hang Loose	Finger Snap	Thumbs Up	Thumbs Down	Turn Palm	Walking Fingers
Desired Gesture	OK Sign	83.1	3.8	6.9	3.6	2.3			0.2
	Peace Sign	2.0	92.5	1.9	0.6	1.2		1.7	0.1
	Hang Loose	2.1	2.3	86.8	4.2	3.3		1.0	0.4
	Finger Snap	2.2	1.6	0.5	89.0	6.2		0.3	0.2
	Thumbs Up	0.1	0.4	1.7	0.3	97.0		0.4	
	Thumbs Down	0.3			0.1	0.6	96.8		2.2
	Turn Palm	0.2	0.1	0.3	0.3	1.0		98.0	0.3
	Walking Fingers	0.3	1.1		0.5	0.6	0.1		97.3

TABLE III
CONFUSION MATRIX FOR SURFACE GESTURES FOR SESSION 1.

		Classified Gesture							
		IF10	IF50	FH10	FH50	FD10	FD50	5F10	5F50
Desired Gesture	IF10	85.7	2.8	0.6	0.2	1.4	0.2	6.5	2.7
	IF50	5.3	87.1	0.6	0.4	0.9	0.5	1.0	4.3
	FH10	0.2		95.4	2.2	1.0	0.7	0.1	0.3
	FH50		0.3	6.3	87.0	1.3	4.1		0.9
	FD10			3.4	1.9	91.3	3.2		0.1
	FD50	0.1	0.1	0.9	3.0	10.0	85.7		0.2
	5F10	6.7	0.1	0.6		0.7	0.1	88.8	3.1
	5F50	1.0	2.9	0.6		0.4	0.3	5.4	89.4

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/TII.2017.2779814, IEEE Transactions on Industrial Informatics

TABLE IV
COMPARISON OF OUR SYSTEM WITH PREVIOUS RESEARCH AND COMMERCIAL PRODUCTS.

		This work	Paper[14]	Paper[35]	Paper[36]	Paper[6]
System	Placement	Wrist	Wrist	Wrist	Wrist	Forearm
	Online or offline	Online	Both	Offline	Offline	Online
	Includes surface gestures	√	×	×	×	×
	sensors	EMG+ IMU	FSR	Microphone	Capacitive sensor	EMG+ IMU
	Numbers of gestures	8	5	8	2	18
	Accuracy	92.8%	Online: 80.5% Offline: 96.3%	88%	-	91.7%

		This work	Delsys	Biometrics	Mega
EMG sensor spec	Electrodes	dry	dry	dry	wet
	SNR(dB)	43.8	38.6	44.4	36.4
	Input impedance	>100MΩ	>100MΩ	>100MΩ	>70MΩ
	Input referred noise (RMS)	0.86mV	-	0.86mV	9.6uV

of tendons, whereas sEMG sensing at the forearm generally only involves sEMG contributions from a single muscle. It could also be that accelerometer and gyroscope data helped to improve performance, especially for air gestures involving movement and disparate overall hand postures.

To achieve real-time performance, this study chose computationally efficient features and classifiers and generated classification decisions every 100 ms. Though this study involved real-time hand gesture classification, studies utilizing offline methods have been able to show higher classification accuracies. For example Dementyev *et al.* [14] were able to accurately classify 96.3% of hand gestures through offline methods but only 80.5% during real-time. Similarly, Catalan *et al.* [38] correctly classified 92.1% of hand gestures via offline methods but only 67.4% in real-time. One possible reason for online and offline differences in accuracy may be due to different computation methods. Offline methods typically use cross validation which is near the upper bound of highest possible accuracy [14]. For example, K-fold cross validation uses K-1 subsamples to train the classifier, however it is unpractical to use such a long-time training data in most real-time applications. Also, online experiments deal with dynamic and real time movements while offline methods reserve and count steady-state movements [24].

sEMG is generally considered not robust enough for practical use over time because of its nonstationary nature [39] and physical electrodes shifting. For example, without recalibration, errors can increase by 15% by the second day of use [39]. In this paper, we integrated previous historical data with current data to improve the robustness of the whole system and there were no significant decreases in accuracy 1 hour and 1 day later (Fig.8 and Table 1).

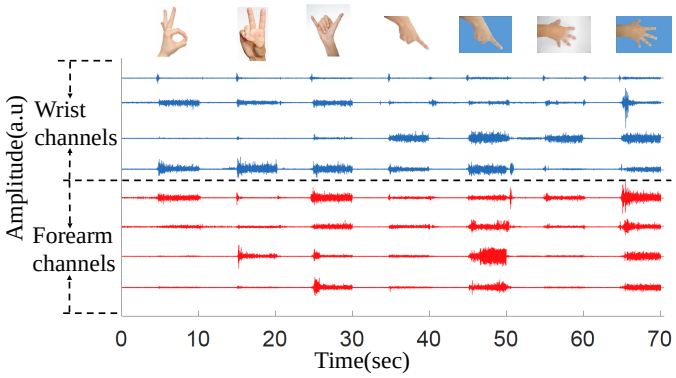
Although, classification accuracies were relatively high in general, some gestures were more susceptible to misclassification. Within air gestures, OK Sign, Peace Sign, and Hang Loose were more frequently misclassified amongst each other than other groups of gestures (Table 2). This is likely because all three gestures were static, requiring vertical wrist poses and thus the acceleration and gyroscope signals were nearly identical and not useful contributors for classification. For surface gestures, 50% MVC movements were frequently misclassified as 10% MVC, with error rates above 5% (Table 3). One possible reason may be that muscle fatigue over time causes lower muscle activations and results in misclassification as 10% MVC. In addition, IF10 is relatively frequently confused with 5F10, which is likely because the muscles used for these two gestures are similar.

One potential limitation of this work is that, as with most other approaches, we do not yet have a calibration-free, universal model that can fit all people. Users still need to perform 1 minute calibration trial before each use. Future work could focus on creating a large database of hand gestures for creating a universal model, without the need for calibration. Also, it is possible that by using more computing resources (to maintain real-time functionality) other clustering algorithms could result in higher classification accuracies and this should be a focus of future research. Besides wrist area is more limited which makes it difficult to employ more channels of the same size sensors as compared to forearm based solutions, and thus we plan to make the sensors smaller and the system more compact with more channels to further improve accuracy.

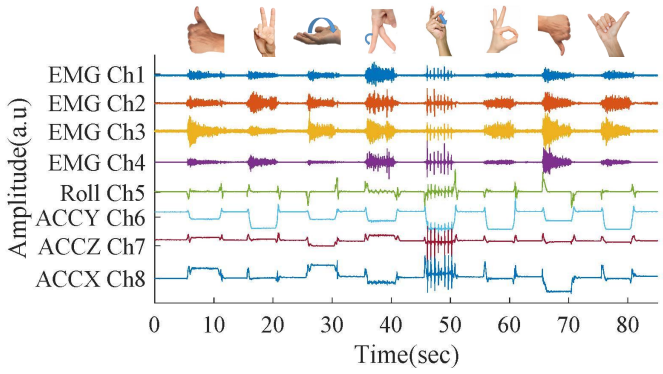
VIII. CONCLUSION

In this paper, we presented a wrist-worn approach via sEMG and IMU sensing for hand gesture recognition, and validated it for 8 air gestures and 4 surface gestures with 2 distinct force levels. To maintain performance when testing 1 hour and 1 day after initial testing, a 1-minute calibration method was proposed. Results showed that the fusion of new training data with historical training data enabled similar accuracy results as the initial use. This paper demonstrated the feasibility of gesture recognition on the wrist which may be a more practical area for wearable devices than the forearm, and thus holds potential to be integrated in to a smart watch or other wrist-worn wearable for intuitive human computer interaction.

APPENDIX



Appendix Fig 1. Representative trial showing raw sEMG signals from the wrist and forearm.



Appendix Fig 2. Raw signals of a representative trial

APPENDIX TABLE I
COMPARISON OF FOREARM AND WRIST SURFACE SEMG SIGNAL STRENGTH.

Air Gestures	OK	PS	HL	FS	TU	TD	TP	WF	AVG
FWR	0.92	1.23	0.96	0.85	1.10	0.79	0.86	0.99	0.96
SNR(Forearm)	6.48	8.00	7.29	12.28	6.70	6.03	4.24	9.36	7.55
SNR(Wrist)	5.25	4.78	5.80	11.35	5.48	6.00	3.48	7.37	6.19
Surface Gestures	**IF10**	**IF50**	**FH10**	**FH50**	**FD10**	**FD50**	**5F10**	**5F50**	**AVG**
FWR	1.03	1.70	1.21	1.91	1.95	1.97	1.09	1.69	1.57
SNR(Forearm)	5.21	12.79	2.82	11.41	7.38	11.07	4.67	12.56	8.49
SNR(Wrist)	3.50	8.56	1.31	5.88	3.05	6.13	2.56	7.80	4.85

REFERENCES

[1] N. Yu, C. Xu, K. Wang, Z. Yang, and J. Liu, "Gesture-based telemanipulation of a humanoid robot for home service tasks," in *Cyber Technology in Automation, Control, and Intelligent Systems (CYBER), 2015 IEEE International Conference on*. IEEE, 2015, pp. 1923–1927.

[2] Z. Lu, X. Chen, Q. Li, X. Zhang, and P. Zhou, "A hand gesture recognition framework and wearable gesture-based interaction prototype for mobile devices," *IEEE transactions on human-machine systems*, vol. 44, no. 2, pp. 293–299, 2014.

[3] M. E. Al-Ahdal and M. T. Nooritawati, "Review in sign language recognition systems," in *Computers & Informatics (ISCI), 2012 IEEE Symposium on*. IEEE, 2012, pp. 52–57.

[4] J. Looker and T. Garvey, "Reaching for holograms : Assessing the ergonomics of the microsoft hololen 3d gesture known as the "air tap"," in *Eeum Design Connects*, 2015.

[5] H. Liang, J. Yuan, D. Thalmann, and N. M. Thalmann, "Ar in hand: Egocentric palm pose tracking and gesture recognition for augmented reality applications," in *Proceedings of the 23rd ACM international conference on Multimedia*. ACM, 2015, pp. 743–744.

[6] X. Zhang, X. Chen, W.-h. Wang, J.-h. Yang, V. Lantz, and K.-q. Wang, "Hand gesture recognition and virtual game control based on 3d accelerometer and emg sensors," in *Proceedings of the 14th international conference on Intelligent user interfaces*. ACM, 2009, pp. 401–406.

[7] M. T. Wolf, C. Assad, M. T. Vernacchia, J. Fromm, and H. L. Jethani, "Gesture-based robot control with variable autonomy from the jpl biosleeve," in *Robotics and Automation (ICRA), 2013 IEEE International Conference on*. IEEE, 2013, pp. 1160–1165.

[8] W. Guo, X. Sheng, H. Liu, and X. Zhu, "Development of a multi-channel compact-size wireless hybrid semg/nirs sensor system for prosthetic manipulation," *IEEE Sensors Journal*, vol. 16, no. 2, pp. 447–456, 2016.

[9] E. Kutafina, D. Laukamp, R. Bettermann, U. Schroeder, and S. M. Jonas, "Wearable sensors for elearning of manual tasks: Using forearm emg in hand hygiene training," *Sensors*, vol. 16, no. 8, p. 1221, 2016.

[10] A. Courteville, T. Gharbi, and J.-Y. Cornu, "Mmg measurement: A high-sensitivity microphone-based sensor for clinical use," *IEEE transactions on biomedical engineering*, vol. 45, no. 2, pp. 145–150, 1998.

[11] P.-G. Jung, G. Lim, S. Kim, and K. Kong, "A wearable gesture recognition device for detecting muscular activities based on air-pressure sensors," *IEEE Transactions on Industrial Informatics*, vol. 11, no. 2, pp. 485–494, 2015.

[12] S. Sikdar, H. Rangwala, E. B. Eastlake, I. A. Hunt, A. J. Nelson, J. Devanathan, A. Shin, and J. J. Pancrazio, "Novel method for predicting dexterous individual finger movements by imaging muscle activity using a wearable ultrasonic system," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 1, pp. 69–76, 2014.

[13] J. Kim, J. He, K. Lyons, and T. Starner, "The gesture watch: A wireless contact-free gesture based wrist interface," in *Wearable Computers, 2007 11th IEEE International Symposium on*. IEEE, 2007, pp. 15–22.

[14] A. Dementyev and J. A. Paradiso, "Wristflex: low-power gesture input with wrist-worn pressure sensors," in *Proceedings of the 27th annual ACM symposium on User interface software and technology*. ACM, 2014, pp. 161–166.

[15] D. Kim, O. Hilliges, S. Izadi, A. D. Butler, J. Chen, I. Oikonomidis, and P. Olivier, "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*. ACM, 2012, pp. 167–176.

[16] T. Oyama, Y. Mitsukura, S. G. Karungaru, S. Tsuge, and M. Fukumi, "Wrist emg signals identification using neural network," in *Industrial Electronics, 2009. IECON'09. 35th Annual Conference of IEEE*. IEEE, 2009, pp. 4286–4290.

[17] T. Oyama, H. Choge, S. Karungaru, S. Tsuge, Y. Mitsukura, and M. Fukumi, "Identification of wrist emg signals using dry type electrodes," in *ICCAS-SICE, 2009*. IEEE, 2009, pp. 4433–4436.

[18] J. McIntosh, C. McNeill, M. Fraser, F. Kerber, M. Löchtefeld, and A. Krüger, "Empress: Practical hand gesture classification with wrist-mounted emg and pressure sensing," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 2016, pp. 2332–2342.

[19] C. Zhu and W. Sheng, "Wearable sensor-based hand gesture and daily activity recognition for robot-assisted living," *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol. 41, no. 3, pp. 569–573, 2011.

[20] P. Laferriere, E. D. Lemaire, and A. D. Chan, "Surface electromyographic signals using dry electrodes," *IEEE Transactions on Instrumentation and Measurement*, vol. 60, no. 10, pp. 3259–3268, 2011.

[21] K. Englehart and B. Hudgins, "A robust, real-time control scheme for multifunction myoelectric control," *IEEE transactions on biomedical engineering*, vol. 50, no. 7, pp. 848–854, 2003.

[22] B. Hudgins, P. Parker, and R. N. Scott, "A new strategy for multifunction myoelectric control," *IEEE Transactions on Biomedical Engineering*, vol. 40, no. 1, pp. 82–94, 1993.

[23] M. A. Oskoei and H. Hu, "Support vector machine-based classification scheme for myoelectric control applied to upper limb," *IEEE transactions on biomedical engineering*, vol. 55, no. 8, pp. 1956–1965, 2008.

[24] W. Guo, X. Sheng, H. Liu, and X. Zhu, "Toward an enhanced human–machine interface for upper-limb prosthesis control with combined emg and nirs signals," *IEEE Transactions on Human-Machine Systems*, 2017.

[25] L. J. Hargrove, K. Englehart, and B. Hudgins, "A comparison of surface and intramuscular myoelectric signal classification," *IEEE transactions on biomedical engineering*, vol. 54, no. 5, pp. 847–853, 2007.

[26] E. J. Scheme, K. B. Englehart, and B. S. Hudgins, "Selective classification for improved robustness of myoelectric control under nonideal conditions," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 6, pp. 1698–1705, 2011.

[27] N. Wang, Y. Chen, and X. Zhang, "The recognition of multi-finger prehensile postures using lda," *Biomedical Signal Processing and Control*, vol. 8, no. 6, pp. 706–712, 2013.

[28] A. D. Bellingegni, E. Gruppioni, G. Colazzo, A. Davalli, R. Sacchetti, E. Guglielmelli, and L. Zollo, "Nlr, mlp, svm, and lda: a comparative analysis on emg data from people with trans-radial amputation," *Journal of neuroengineering and rehabilitation*, vol. 14, no. 1, p. 82, 2017.

[29] P. Kaufmann, K. Englehart, and M. Platzner, "Fluctuating emg signals: Investigating long-term effects of pattern matching algorithms," in *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE*. IEEE, 2010, pp. 6357–6360.

[30] H. Liu, X. Yuan, Q. Tang, and R. Kustra, "An efficient method to estimate labelled sample size for transductive lda (qda/mda) based on bayes risk," *Lecture notes in computer science*, pp. 274–285, 2004.
[31] J. W. Sensinger, B. A. Lock, and T. A. Kuiken, "Adaptive pattern recognition of myoelectric signals: exploration of conceptual framework and practical algorithms," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 17, no. 3, pp. 270–278, 2009.
[32] J. Liu, X. Sheng, D. Zhang, J. He, and X. Zhu, "Reduced daily recalibration of myoelectric prosthesis classifiers based on domain adaptation," *IEEE journal of biomedical and health informatics*, vol. 20, no. 1, pp. 166–176, 2016.
[33] J. Liu, X. Sheng, D. Zhang, and X. Zhu, "Boosting training for myoelectric pattern recognition using mixed-lda," in *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*. IEEE, 2014, pp. 14–17.
[34] N. Fay, C. J. Lister, T. M. Ellison, and S. Goldin-Meadow, "Creating a communication system from scratch: gesture beats vocalization hands down," *Frontiers in psychology*, vol. 5, 2014.
[35] H. Kato and K. Takemura, "Hand pose estimation based on active bone-conducted sound sensing," in *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*. ACM, 2016, pp. 109–112.
[36] J. Rekimoto, "Gesturewrist and gesturepad: Unobtrusive wearable interaction devices," in *Wearable Computers, 2001. Proceedings. Fifth International Symposium on*. IEEE, 2001, pp. 21–27.
[37] C. J. De Luca, "The use of surface electromyography in biomechanics," *Journal of applied biomechanics*, vol. 13, no. 2, pp. 135–163, 1997.
[38] M. Ortiz-Catalan, R. Brånemark, and B. Håkansson, "Biopatrec: A modular research platform for the control of artificial limbs based on pattern recognition algorithms," *Source code for biology and medicine*, vol. 8, no. 1, p. 11, 2013.
[39] J. He, D. Zhang, X. Sheng, and X. Zhu, "Effects of long-term myoelectric signals on pattern recognition," in *International Conference on Intelligent Robotics and Applications*. Springer, 2013, pp. 396–404.

Shuo Jiang received B.S. in Mechatronic Engineering from Zhejiang University, Hangzhou, China, in 2015. He is currently pursuing the Ph.D. degree in the School of Mechanical Engineering at Shanghai Jiao Tong University, Shanghai, China.

His research interests include multi-sensor fusion for intuitive human computer interaction.

Bo Lv received the bachelor's degree from the School of Mechanical Engineering & Automation at Xi'an Jiao Tong University, Shaanxi, China, in 2015. He is currently working toward the Ph.D. degree in the School of Mechanical Engineering at Shanghai Jiao Tong University, Shanghai, China.

His research interests include machine learning, biological signal processing with application to upper limb prostheses control.

Weichao Guo received the bachelor's degree from the School of Power and Mechanical Engineering at Wuhan University, Wuhan, China, in 2012. He is currently working toward the Ph.D. degree in the School of Mechanical Engineering at Shanghai Jiao Tong University, Shanghai, China.

His research interests include fusion of multi-sensor information and enhanced human-machine interface.

Chao Zhang received the B.S. degree in electrical engineering from Tianjin University, Tianjin, China, in 2009, and the Ph.D. degree in computer science from the University of New South Wales, NSW, Australia in 2014. He is currently a Senior Research Scientist at Samsung R&D, China.

His research interests include image processing, bio-signal processing, visual relationship understanding, and image Q&A.

Haitao Wang received the M.S. degree in engineering from the University of Science and Technology Beijing, in 2001, and the Ph.D. degree in computer science from Institute of Automation, Chinese Academy of Sciences, Beijing, China, in 2004.

From 2004 to 2016, he was a Researcher of SAIT Beijing Lab, Samsung Advanced Institute of Technology (SAIT). Currently he is Technical Expert in AR/VR lab, Huawei. His research interests include HCI, 3D capturing and 3D display, machine learning.

Xinjun Sheng received the B.Sc., M.Sc. and Ph.D. degrees in mechanical engineering from Shanghai Jiao Tong University, Shanghai, China, in 2000, 2003 and 2014. In 2012, he was a visiting scientist in Concordia University, Canada. He is currently an Associate Professor in the School of Mechanical Engineering at Shanghai Jiao Tong University.

His current research interests include robotics and biomechatronics. Dr. Sheng is a member of IEEE, RAS, EMBS, and IES.

Peter B. Shull received the B.S. in Mechanical Engineering and Computer Engineering from LeTourneau University in 2005 and M.S. and Ph.D. degrees in Mechanical Engineering from Stanford University in 2008 and 2012, respectively. From 2012 to 2013, he was a postdoctoral fellow in the Bioengineering Department of Stanford University. He is an Associate Professor in Mechanical Engineering at Shanghai Jiao Tong University.

His research interests include wearable systems, real-time movement sensing and feedback, gait retraining, and biomechanics.