

Towards Wearable A-mode Ultrasound Sensing for Real-time Finger Motion Recognition

Xingchen Yang, *Student Member, IEEE*, Xueli Sun, *Student Member, IEEE*, Dalin Zhou, *Student Member, IEEE*, YueFeng Li, *Student Member, IEEE*, and Honghai Liu*, *Senior Member, IEEE*

Abstract—It is evident that surface electromyography (sEMG) based human-machine interfaces (HMI) have inherent difficulty in predicting dexterous musculoskeletal movements such as finger motions. This paper is an attempt to investigate a plausible alternative to sEMG, ultrasound-driven HMI, for dexterous motion recognition due to its characteristic of detecting morphological changes of deep muscles and tendons. A multi-channel A-mode ultrasound lightweight device is adopted to evaluate the performance of finger motion recognition; an experiment is designed for both widely acceptable offline and online algorithms with eight able-bodied subjects employed. The experiment result presents that the offline recognition accuracy is up to $98.83\% \pm 0.79\%$. The real-time motion completion rate is $95.4\% \pm 8.7\%$ and online motion selection time is 0.243 ± 0.127 s. The outcomes confirm the feasibility of A-mode ultrasound based wearable HMI as well as its prosperous applications in prosthetic devices, virtual reality and remote manipulation.

Index Terms—Human-machine interface, A-mode ultrasound, finger motion recognition, online gesture recognition.

I. INTRODUCTION

PEOPLE suffering from transradial amputation are significantly influenced both physically and mentally because of the loss of motor functionality. To reconstruct their lost hand functions, surface electromyography (sEMG) based human-machine interface (HMI) has been applied to control prosthetic hands for decades [1]–[7]. sEMG signals, generated with muscle contraction and collected noninvasively, are naturally related with motion intentions and provide the insight into muscle activities [8]. However, sEMG signals hardly reflect the movement of muscles deep below surface tissue due to the crosstalk, attenuation, low signal to noise ratio (SNR), etc [9]–[11]. Since typical finger-related muscles such as flexor digitorum sublimis, flexor pollicis longus, and extensor pollicis longus are far away from surface tissue, in principle sEMG-based solutions are impractical for predicting dexterous finger motions [12]. Recent advanced electronics and hardware design in prostheses like the iLimb from Touch Bionics [13], challenge dexterous motion prediction algorithms for intuitively controlling a prosthetic hand [14]. Though the primary

aim has been achieved by intramuscular EMG and targeted muscle reinnervation (TMR), the invasiveness has impeded the acceptance of such control strategies in daily life activity conduction [15], [16]. A novel non-invasive HMI is urgently needed to accommodate intuitive sensing of dexterous human body movements especially finger motions.

Ultrasound could be an alternative due to its capability of penetrating several centimeters below the skin and returning the information of both superficial and deep muscles in a noninvasive manifestation [17], [18]. Recent studies present evidence in visualizing muscle activities and HMI [19]–[32]. The research group led by Zheng *et al.* conducted a few studies about ultrasound-image based HMI. They extracted muscle thickness through ultrasound imaging and established nonlinear mapping between muscle thickness and fatigue degree [23]. A linear relationship between muscle deformation and normalized torque of biceps brachii was further revealed in their subsequent research [24]. Distinct patterns in ultrasound images of different finger flexion have been reported and utilized in motion recognition with an overall accuracy of 94%. These prior studies firstly proved the feasibility of ultrasound imaging as an alternative HMI in finger motion recognition [22], [25]. In addition, finger position and fingertip force were predicted precisely by Castellini *et al.* based on ultrasound imaging using uniformly-spaced grid of interest points and the corresponding spatial first-order features [18], [26], [27]. Skidar *et al.* applied ultrasound imaging to classify 15 hand gestures offline and four motions in real time, demonstrating the feasibility of ultrasound imaging as a robust HMI [11]. Moreover, they conducted a preliminary experiment with a transradial amputee to control a virtual hand using ultrasound-image based HMI and achieved a completion rate of above 70% [28]. Classification performance of sEMG signals and ultrasound imaging for 14 finger motions was compared by Liu *et al.*, with one-way analysis of variance (ANOVA) showing a significant advantage of ultrasound-image based HMI [29].

Although the above-mentioned research indicates that ultrasound imaging is a reliable alternative to sEMG in HMI modality for finger motion and position recognition, the employed B-mode ultrasound equipment is bulky, expensive, and hardly wearable for the development of wearable prosthetic devices. In order to overcome the drawbacks of B-mode ultrasound, single-element ultrasound (A-mode ultrasound) only reflecting the depth information of a specific direction is employed. In spite of its simplicity, A-mode ultrasound enjoys notable merits in practical use. For instance, the transducer could be customized in a much smaller size and integrated

This work was supported by the National Natural Science Foundation of China (Grant Nos. 51575338, 51575407, 51475427).

Yang, Sun, Li, Liu are with the State Key Laboratory of Mechanical System and Vibration, School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai, China. (e-mail: xingchen.yang@sjtu.edu.cn; sunxueli@sjtu.edu.cn; yflee248@sjtu.edu.cn; honghai.liu@sjtu.edu.cn).

Zhou, Liu are with the School of Computing, the University of Portsmouth, PO1 3HE, Portsmouth, UK. (e-mail: dalin.zhou@port.ac.uk; honghai.liu@port.ac.uk).

with a prosthetic device; the computation cost can be further reduced to meet the real-time requirement. A few studies had been carried out to explore its capability as an alternative for dextrous motion recognition. Zheng *et al.* extracted muscle deformation by an A-mode ultrasound transducer and derived a linear relationship between muscle deformations and wrist extension angles [20]. Sikdar *et al.* recently predicted dexterous finger motions using a mechanically scanned A-mode ultrasound transducer with good performance [30]. Hettiarachchi *et al.* conducted the finger motion recognition for transradial amputees using eight customized A-mode ultrasound transducers [31]. Li *et al.* also demonstrated that five single digit flexion could be classified using four A-mode ultrasound transducers, with a classification accuracy up to 96% [32].

Noting that all the previous work related to A-mode ultrasound are analyzed offline, the real-time performance of A-mode ultrasound remains to be verified. This paper aims to further evaluate the feasibility of A-mode ultrasound for finger motion recognition. Specifically, a 4-channel A-mode ultrasound ring sensing system is adopted to recognise more finger motions commonly used in activities of daily living (ADL). Four online evaluation metrics are adopted to assess the real-time performance. In addition, the effect of the number of exploited sensing channels on recognition accuracy is also investigated. The paper is organized as follows: Section II-A provides the information of employed subjects, device configuration and the experiment paradigm. Offline classification method and detailed experiment are described in Section II-B. Section II-C describes the real-time test and online performance indicators. Experimental results and extended discussion are presented in Section III and IV.

II. MATERIALS AND METHODS

A. Subjects and Hand Gestures

1) *Subjects*: Eight able-bodied male subjects (aged 23-26, denoted as S1 to S8) without history of neuromuscular and joint diseases participated in the experiment. The procedure was reviewed and approved by the SJTU School Ethics Committee. All subjects had received a thorough description of the experiment in both oral and written form and signed the informed consent according to declaration of Helsinki.

2) *Hand Gestures*: In order to evaluate the performance of the A-mode ultrasound for finger motion recognition, 11 hand motions including rest state were chosen without considering wrist motions. Specifically, four single digit flexion: thumb flexion (TF), index finger flexion (IF), middle finger flexion (MF), ring finger flexion (RF) and six combined finger motions: lateral grasp (LG), fine pinch (FP), tripod grasp (TG), index point (IP), fist (FS), hang loose (HL) and rest state (RS) were chosen as motion candidates. Most of these hand motions are commonly used in ADL [33], [34].

B. Offline Data Acquisition and Pattern Recognition

1) *Sensor Placement*: A commercial 4-channel A-mode transducer driver board (Zhongxu Tech., China) was chosen (length: 24 cm, width: 10 cm, height: 1.8 cm), which was used to drive four channels of A-mode ultrasound transducer

independently and receive and amplify echo signals sequentially. The driving pulse voltage, repetition frequency and amplification gain were set to -40 V, 10 Hz, and 37 dB, respectively. The repetition frequency was equal to real-time recognition frequency and set as 10 Hz. In each echo-receiving period, 8192 dots were sampled at a frequency of 100 MHz. A lower sampling rate was also feasible as long as the sampling theorem was met. The sound velocity in human tissues is around 1540 m/s [35]. With the above settings, the inspecting depth was as deep as 63 mm in this experiment. This depth is practical to detect the movements of deep muscles and tendons in the forearm.

A customized A-mode transducer with a 5 MHz piezoelectric ceramic and a matching layer was designed for this study (diameter: 14 mm, height: 18 mm). The detailed parameters can be seen in APPENDIX Table I. A correlative study evidenced that 5 MHz was the optimum frequency with a good SNR in both low and high frequency domains [31], and the matching layer was applied to reduce the reflection between the piezoelectric ceramic and skin. Standard ultrasound gel was applied between skin and surface of transducer. A customized armband was designed to secure four A-mode transducers still on the surface of forearm skin. The position of every transducer was chosen carefully ensuring the coverage of flexor digitorum sublimis (FDS), flexor digitorum profundus (FDP), flexor pollicis longus (FPL), extensor digitorum communis (EDC), and extensor pollicis longus (EPL). These five muscles play a crucial role in finger movements including digit flexion and combined finger motions. The armband was fixed on the middle portion of the forearm. The placement of the armband and covered muscles were shown in Fig. 1.

2) *Data Acquisition*: During the offline experiment, the subjects were instructed to place their elbows on an armrest and lift their forearms with palms facing upward. All the 11 hand gestures mentioned above were repeated for 10 trials. For each trial, each motion was held for 5 seconds, and only the data of the middle 3 seconds was analyzed. A customized graphical user interface (GUI) was programmed for both offline and online experiment based on C++ (Visual Studio 2010, Microsoft, USA), as shown in Fig. 2. The subjects were asked to follow cues from the GUI to perform corresponding motions at a moderate level of force. Between every two adjacent trials, there was a 10 seconds rest to avoid fatigue.

3) *Feature Extraction*: The feature extraction process was inspired by the spatial first-order feature proposed by Castellini *et al.*, which was linearly related to metacarpophalangeal joint angle and fingertip force [26], [27]. After choosing a uniformly-spaced grid of interest points in the ultrasound image, plane fitting was applied for each region of interest, and three fitting coefficients were noted as the spatial first-order feature. In our method, the plane fitting was transformed to linear fitting [29]. This transformation made the method applicable for one-dimensional A-mode ultrasound signals. The details of the feature extraction process are described as follows.

Since the raw echo signals received from the A-mode ultrasound transducer are always corrupted with noise, processing was done to get rid of the noise. Fig. 3 shows the

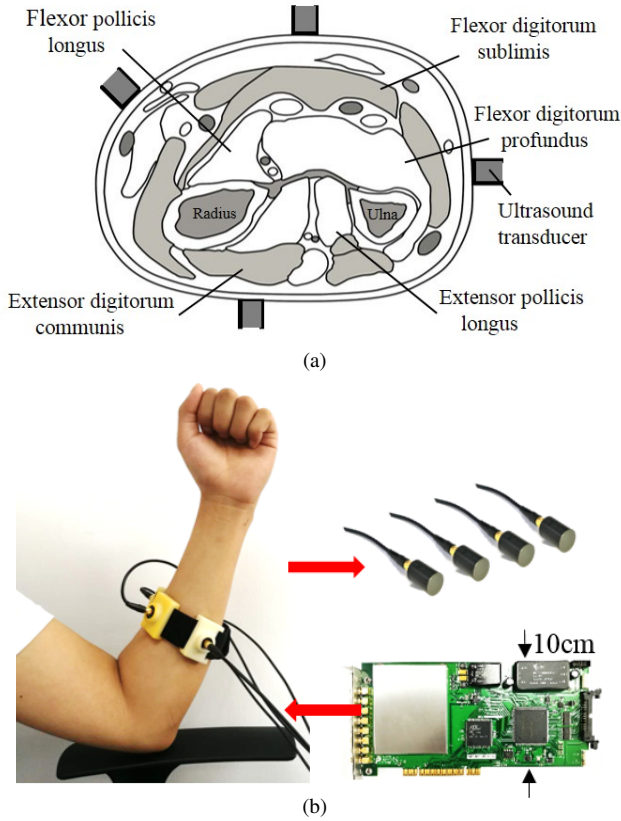


Fig. 1. Placement of the A-mode ultrasound transducers. (a) Section view of the midposition of the forearm and the distribution of ultrasound transducers (b) Placement of the armband and overview of the hardware system.

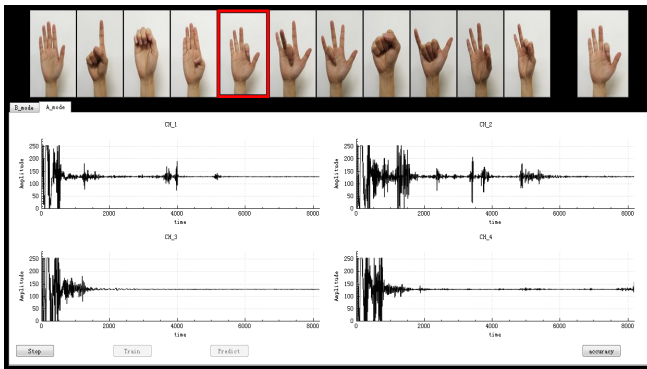


Fig. 2. A customized graphical user interface (GUI) for finger motion recognition. From left to right, the 11 gestures were rest state (RS), index point (IP), lateral grasp (LG), thumb flexion (TF), index finger flexion (IF), middle finger flexion (MF), ring finger flexion (RF), fist (FS), hang loose (HL), fine pinch (FP) and tripod grasp (TG). During real-time test, the red frame was used to prompt current motion and the predicted result was shown on the upper-right corner.

schematic flow of the preprocessing. It consists of time gain compensation (TGC), Gaussian filtering, Hilbert transform, and log compression [35].

Segmentation and linear fitting were then adopted for feature extraction. Firstly, the preprocessed echo signals were segmented into a series of windows with a fixed length without overlap along the propagation direction of the ultrasound. The first 20 sampling dots and the last 20 sampling dots for each

echo signal were removed before segmentation, since they carried little meaningful information. The segment length was selected as 280 sampling dots empirically. As the sampling frequency was 100 MHz, this segment length represented a time interval of 2.8 us. All the subjects and channels of ultrasound transducers shared the same segment parameter. Secondly, linear fitting was applied for each window, by which morphological information of muscles contained in ultrasound signals could be preserved. The two fitting coefficients were noted as features of corresponding window. The features of all the windows from four channels were combined as a feature vector. Hence, the dimensionality of the feature vector is 232 ($4 \times 29 \times 2$).

4) *Hand Gesture Recognition*: Two classifiers were employed for recognizing the hand gestures in this experiment including linear discriminant analysis (LDA) classifier (Bayesian method) and support vector machine (SVM) classifier (Linear kernel) [36], [37]. The adopted classification strategies were commonly applied in sEMG-based pattern recognition due to their simplicity and robust performance.

Five-fold cross-validation was adopted. That is to say, the database was randomly and equally divided into five parts, one of the five parts were notified as testing set, and the others were regarded as training set. The five-fold cross-validation was repeated for ten times, yielding 50 classification accuracies. The classification accuracy (CA) was defined as Equation 1:

$$CA = \frac{\text{Number of correctly recognized motions}}{\text{Total number of testing motions}} * 100\% \quad (1)$$

C. Real-time Test and Performance Metrics

1) *Real-time Test*: Since the performance of the LDA classifier and the SVM classifier were not significantly different according to one-way ANOVA during the offline classification, only the LDA classifier was adopted for the online test for its lower computation cost. With the aforementioned feature and LDA classifier, the feature extraction and classification time was less than 11 ms for each prediction (3.1-GHz Inter Core i5-2400 computer).

The aforementioned customized GUI in Fig. 2 was also applied in the real-time experiment, and the subjects were asked to maintain the same posture as in the offline experiment. Different from the offline experiment, 3 trials training and 6 trials testing were conducted. For each trial in the training stage, each motion was held for 5 seconds, and only the data from the middle 3 seconds were adopted for training. In the testing stage, the definition of a trial was as follows. The GUI randomly prompted a targeted motion with a red frame, and the subjects were instructed to perform the corresponding motion. The prompt for each targeted motion would last for 5 seconds. During this period, the real-time predicted motion would be displayed on the upper-right corner in the GUI to provide a visual feedback. In order to calculate the online performance metrics defined in the following subsection, there was a 5 seconds rest state before each targeted motion [38]. All the real-time predicted labels and real labels were recorded for further analyses.

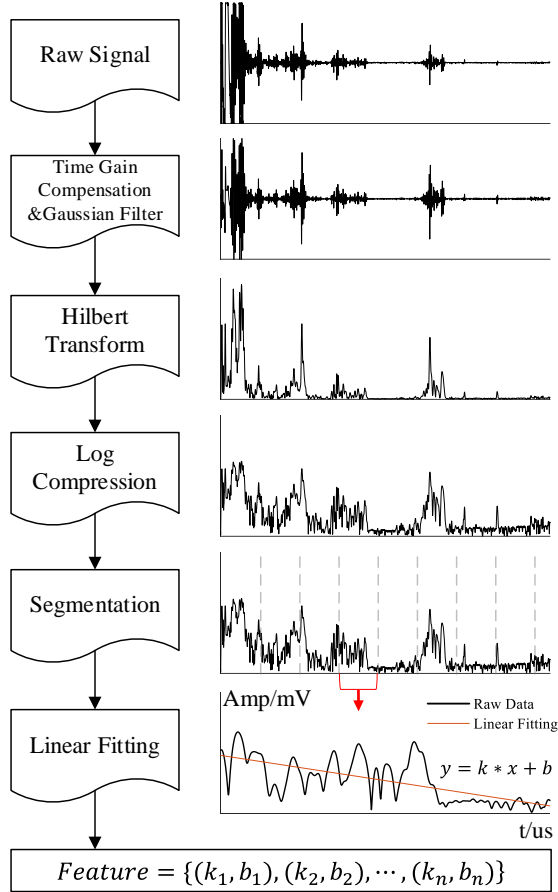


Fig. 3. Flow chart of preprocessing and feature extraction. t = time, denoting propagation time of ultrasound; Amp = Amplitude, denoting amplitude of ultrasound.

2) *Performance Metrics*: To evaluate the online performance of HMI-A (Human-machine interface based on A-mode ultrasound) for finger motion recognition, four metrics applied in sEMG-based HMI were chosen, including motion selection time (ST), motion completion time (CT), motion completion rate (CR) and real-time accuracy (RA) [39], [40].

ST is the time interval between motion onset to the first right prediction of the targeted motion, where the motion onset is defined as the time of the last prediction of the rest state [33]. ST is able to describe the responsiveness of the HMI. CT is the time consumed to predict the current motion rightly for 10 times, which represents the control stability of the HMI. If a specific motion is correctly predicted 10 times in 5 seconds, it is regarded as an instance of recognition completion. Otherwise, it is regarded as recognition failure. This time limitation is selected based on clinical experience [39]. Since any prosthetic hand operations exceeding 5 seconds is too time-consuming to be tolerated by users. CR is defined as the percentage of completed motions within 5 seconds, and it reflects the usability of the HMI for users. RA is defined as the classification accuracy from the first correct prediction to the end of the 5 seconds prediction time, which can represent

the prediction stability. The ST, CT and RA are counted only when the current motion is noted as recognition completion. Fig. 4 shows the detailed explanation of ST and CT.

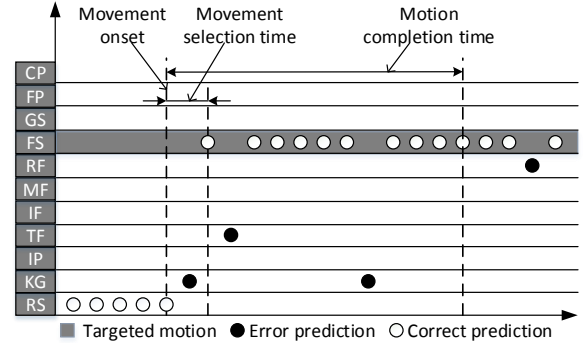


Fig. 4. Illustration of real-time performance metrics. Each targeted motion starts from the rest state. A new prediction is carried out for every 100 ms.

D. Statistical Analysis

Average offline classification accuracy and real-time performance metrics are reported with mean and standard deviation (SD). One-way ANOVA was applied to assess the statistic difference between the means of compared data based on MATLAB (MathWorks Inc. USA). Statistical significance level was set as $p < 0.05$ for all the comparisons.

III. EXPERIMENT RESULTS

A. Offline Recognition Accuracy

Fig. 5 presents the average classification error rate of 11 finger motions for each subject. The average classification accuracy of the eight able-bodied subjects is $98.83\% \pm 0.79\%$ for the LDA classifier and $98.41\% \pm 0.78\%$ for the SVM classifier, respectively. One-way ANOVA demonstrates that the difference between the performance of the LDA classifier and SVM classifier is not significant ($p > 0.05$).

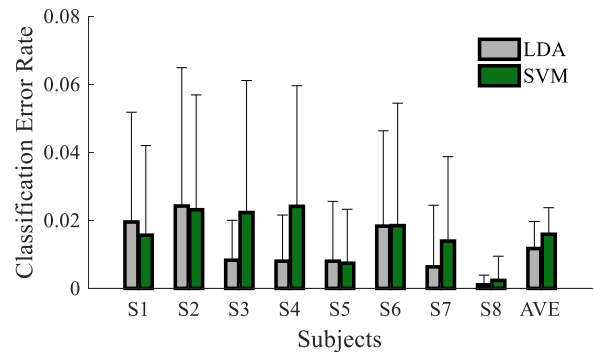


Fig. 5. Classification error rate of HMI-A for the 8 different subjects. Error bars represent the standard deviations of the error rate across the 11 motions. Last bar represents the average error rate and standard deviation across all the motions and subjects.

Fig. 6 presents the confusion matrix of motion recognition results based on the LDA classifier which is similar to that of

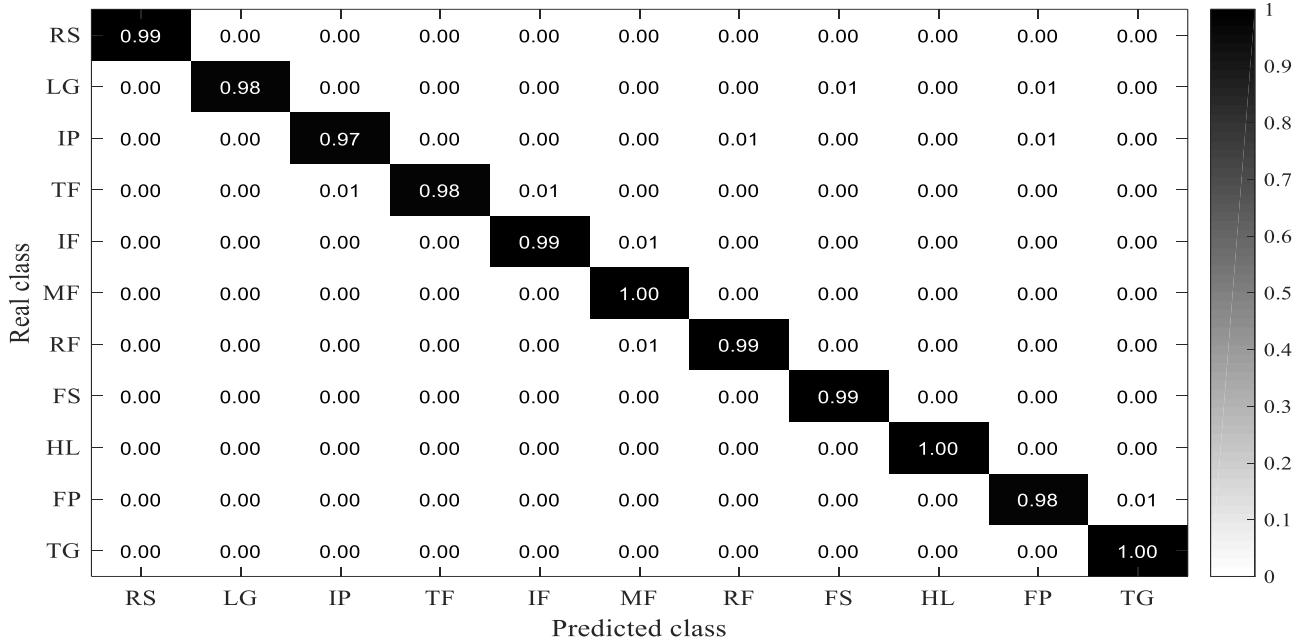


Fig. 6. Confusion matrix of HMI-A for 11 different finger motions based on LDA classifier. The motions are corresponding to the experimental paradigm in Fig. 2.

the SVM classifier ($p > 0.05$). For each motion, the average classification accuracy is above 97%.

Fig. 7 shows the relationship between classification accuracy and the adopted channel number of A-mode ultrasound transducer. All the channel combinations were analyzed to provide a preliminary insight due to the limited detection sites. The results show that the classification accuracy drops with the decrease of channel number for both the LDA classifier and the SVM classifier, but for the LDA classifier, the average recognition accuracy is nearly 90% even if only one channel is applied and nearly 95% if two channels are adopted. The performance of the SVM classifier is slightly inferior to LDA classifier when only one or two channels are adopted. When more channels are incorporated, the performance of the SVM classifier is close to the LDA classifier.

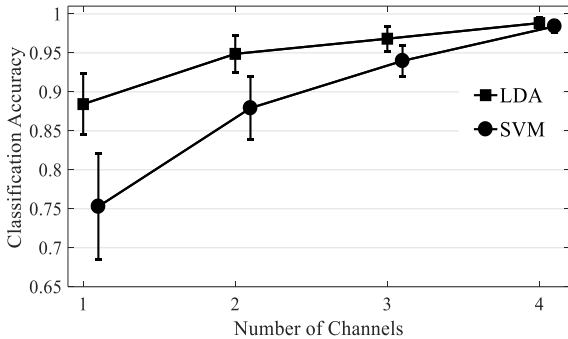


Fig. 7. Classification accuracy versus number of A-mode ultrasound transducer channels across 11 different finger motions.

B. Online Evaluation Performance

The online performance of the HMI-A for finger motion recognition is summarized in Table I. All the values in this table are mean values for 10 motions except for the rest state, which is not analyzed in the real-time test. The average ST, CT, CR and RA for the HMI-A are 0.243 ± 0.127 s, 1.231 ± 0.239 s, $95.4\% \pm 8.7\%$, and $92.4\% \pm 8.7\%$, respectively.

Fig. 8 shows the online performance of the eight able-bodied subjects for 10 finger motions. As shown in Fig. 8(a), except for IP, RF, and FS, the ST of the other seven motions is less than 0.3 s. It means that the delay of the motion selection is not perceivable by users, which is very crucial in intuitive control [36]. Fig. 8(b) presents that the motions other than RF and FP are completed within 1.3 s. Because the recognition frequency is 10 Hz, only three or fewer recognition faults occur before motion completion. According to Fig. 8(c), the CRs of all the 10 motions are nearly 90% except for MF, and the average CR is as high as 95%. In the realistic application of prosthetic hand control, the CR is the most telling metric which reflects whether intended motions of users can be achieved [39]. Fig. 8(d) shows that the RA for every motion is above 85%. Except MF, RF and FP, the CAs of the other seven motions are above 90%. Apparently, the muscle movements of MF and RF are much similar, which leads to a significant confusion. Another notable confusion is between TG and FP.

The comparison for the online performance of single finger flexion and combined finger motions is shown in Fig. 9. The RA and CR of combined finger motions are higher than single finger flexion in terms of recognition accuracy, but the CT and ST are longer compared with single finger flexion. One-way ANOVA shows that there is no significant difference between

TABLE I
PERFORMANCE METRICS FOR EIGHT HEALTHY SUBJECTS

Metrics	Subjects								Mean \pm SD
	S1	S2	S3	S4	S5	S6	S7	S8	
ST (s)	0.246	0.170	0.214	0.317	0.245	0.273	0.165	0.31	0.243 \pm 0.127
CT (s)	1.233	1.087	1.286	1.407	1.243	1.240	1.097	1.256	1.231 \pm 0.239
CR	0.933	0.983	0.933	0.867	1	0.983	1	0.933	0.954 \pm 0.087
RA	0.899	0.984	0.900	0.856	0.894	0.927	0.987	0.926	0.924 \pm 0.087

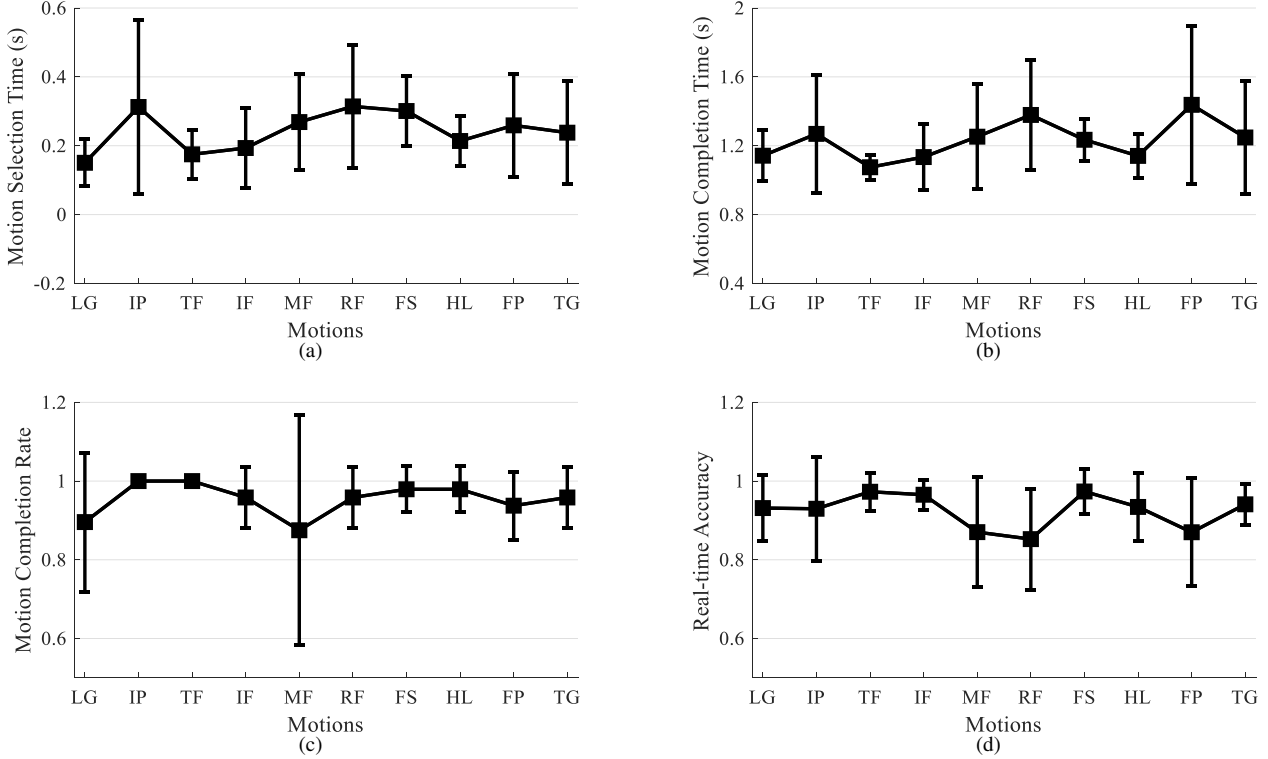


Fig. 8. Average online performance of eight able-bodied subjects for 10 different finger motions. (a) Motion selection time (ST). (b) Motion completion time (CT). (c) Motion completion rate (CR). (d) Real-time accuracy (RA).

the performance of single finger flexion and combined finger motions.

As shown in Fig. 10(a), in cases less than 200 ms, about 44% single finger flexion motions and 40% combined finger motions are successfully selected. From 200 ms to 300 ms, more combined finger motions are selected than single finger flexion. Within 300 ms, about 80% motions are selected both for single finger flexion and combined finger motions. According to Fig. 10(b), up to 75% single finger flexion motions are recognized for 10 times in 1.2 s, but only 60% combined finger motions are completed in 1.2 s. Within 1.4 s, about 80% motions are completed both for single finger flexion and combined finger motions.

IV. DISCUSSION AND CONCLUSION

The offline and online performance of HMI-A for finger motion recognition has been evaluated in this paper. Experi-

ment results conclude that $98.83\% \pm 0.79\%$ of hand motions in test set are correctly predicted offline, and $95.4\% \pm 8.7\%$ of the intended motions are completed with a selection time of 0.243 ± 0.127 s and an accuracy of $92.4\% \pm 8.7\%$ in real time. The delay of the control system is less than 300 ms and will not be perceivable [36]. Further study presents that 11 finger motions are recognized precisely offline with only one or two A-mode ultrasound transducers. Besides, there is no significant difference for the real-time performance of single finger flexion and combined finger motions. These outcomes elucidate the feasibility of HMI-A as an ideal alternative to the sEMG-based HMI especially for scenarios requiring dexterous finger motion recognition.

A. Feature validation

Fig. 11 represents the projections of different features on two-dimensional space by principal component analysis (P-

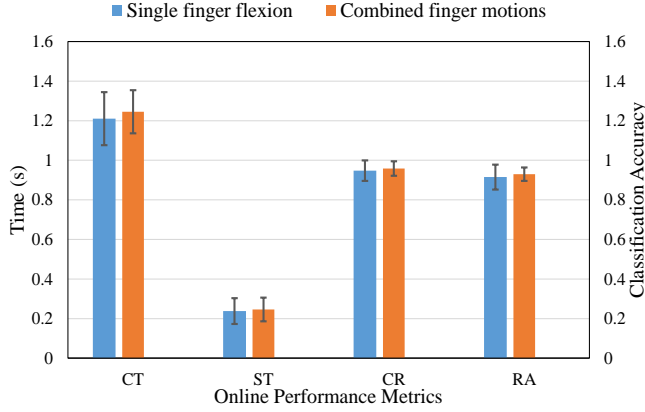


Fig. 9. Comparison of online performance for single finger flexion and combined finger motions.

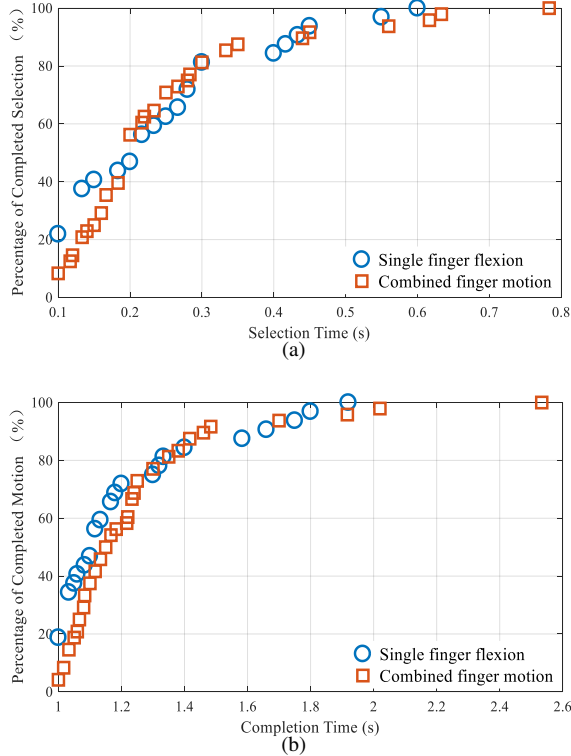


Fig. 10. Online performance of HMI-A. (a) Cumulative percentage of completed selection motions versus selection time. (b) Cumulative percentage of completed motions versus completion time.

CA). The features are projected to the two most discriminatory dimensions for the visualization. It is apparent that features for different motions can be distinguished clearly even though only two-dimensional features (232 dimension in total) are considered. This validates the effectiveness of selected feature for finger motion classification.

The relationship between segment length of signals during feature extraction and classification accuracy is shown Fig. 12. Overall, with the increase in the segment length, the computation cost decreases, but the classification accuracy drops. When the segment length increases to 600, the classification

accuracy is still above 98%. This result demonstrates that the performance of selected feature is not sensitive to the segment length of the signals.

B. Ultrasound transducer shift

Sensor shift is a critical problem for both EMG-based HMI and ultrasound-based HMI, which always occurs as the user dons the device and destroys the performance of human-machine interface severely [27], [41], [42]. In order to evaluate the sensitivity of HMI-A to ultrasound transducer shift, a preliminary study for this problem has been conducted. Six commonly used hand motions were selected in this preliminary experiment, including RS, IP, LG, FS, FP, and TG. Results showed that a classification accuracy of above 83% was achieved with ± 5 mm circumferential transducer shift during donning and doffing. Besides, a classification accuracy of 93% was achieved with negligible transducer shift during donning and doffing.

These results demonstrate that transducer shift does influence the classification performance of HMI-A, but the performance is still acceptable with a minimal transducer shift during donning and doffing. To mitigate the effect, a strategy of training all or part of expected displacement positions can be adopted [43]. In addition, some robust features and transducer configuration methods can be further studied [42], [44].

C. Advantages of the HMI-A

1) *Advantages over EMG-based HMI*: The A-mode ultrasound can achieve sub-millimeter spatial resolution and millisecond temporal resolution, penetrating several centimeters below the skin; This is similar to its counterpart of the B-mode ultrasound [27]. The HMI-A can overcome some inherent limitations of sEMG-based HMI such as inability to differentiate contraction of deep muscles and lack of robust graded signals [45]–[47]. Moreover, the sEMG-based HMI commonly utilizes temporal signal features, thus requiring a long time window to extract meaningful features which causes a latency between signal acquisition and classification. By contrast, a spatial feature is applied in the HMI-A and the real-time performance could be improved easily. The maximum frame rate of the ultrasound equipment used in this experiment is 100 Hz, whose corresponding temporal resolution could be 10 ms. Namely, a new prediction can be given for every 10 ms without decreasing the classification accuracy.

Extensive research has been conducted on sEMG-based HMI to address the real-time performance. Li *et al.* carried out an experiment on five lateral transradial amputated subjects to perform 10 wrist and hand motions, using both amputated and intact arms. As to their intact arms, the average ST was 0.18s for wrist and hand motions and 0.33 s for hand motions only. The CT was 1.14 s for wrist and hand motions and 1.54 s for hand motions only. The CRs were 81.2% for wrist and hand motions and 69.4% for hand motions only [39]. Recently, Guo *et al.* evaluated four online metrics for sEMG only, near infrared spectroscopy (NIR) only and combined modality. The conclusion pointed out that the performance of

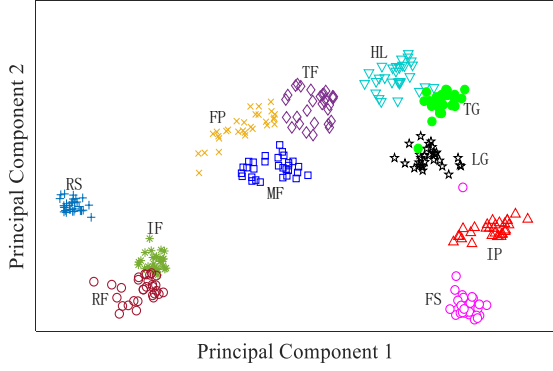


Fig. 11. Projections of different features on two-dimensional space by principal component analysis (PCA).

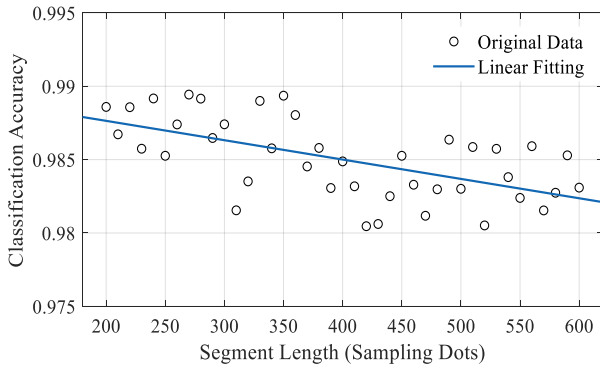


Fig. 12. Relationship between classification accuracy and segment length of signals during feature extraction. The unit for the segment length is sampling dots. Each sampling dot represents a time interval of 0.01 us.

combined modality was optimal with ST of 0.27 s, CT of 1.29 s, CR of 94% and RA of 90% [40].

It is not fair to compare their experimental results with ours directly because of the difference of gesture sets, number and type of sensors used, anatomical structures, etc. However, by referring to their results, the online performance of HMI-A is promising, with shorter ST (0.243 s), CT (1.231 s) and higher CR (95.4%) and RA (92.4%) on finger motion recognition.

As for intramuscular EMG, it can provide independent signals free of crosstalk from deep muscles compared with sEMG. However, the gesture recognition performance of sEMG and intramuscular EMG is similar [48], [49]. Furthermore, the intramuscular EMG is not preferable as compared with non-invasive ultrasound sensing because of its harmful nature for muscles.

2) *Advantages over ultrasound imaging based HMI*: Compared with B-mode ultrasound, both the equipment and the transducer of A-mode ultrasound are much cheaper, smaller, more wearable, and easier to compute. Generally, the probe of B-mode ultrasound equipment is made up of an array of piezoelectric transducers, including 128 or more single-element ultrasound transducers. Each small transducer of the B-mode ultrasound probe can be regarded as an A-mode transducer. The volume and weight of B-mode ultrasound

probe hinders its fixation on users' skin in real-life application, restricting its wearability. Besides, the finger-related muscles are distributed around the circumference of the forearm. With the distribution of multi-channel transducers, A-mode ultrasound can monitor the movements of all these muscles. While the B-mode ultrasound probe is too bulky to cover all the targeted areas. Moreover, the echo signal of A-mode ultrasound is not two-dimensional image but one-dimensional signal, contributing to a shorter computation latency of HMI-A in real-time classification with the advantage of less data being processed. An offline classification accuracy ranging from 92% to 96% has been reported for different finger motions of able-bodied participants based on ultrasound imaging [25], [28], [29]. Compared with the previous recognition accuracy based on ultrasound imaging, the classification accuracy of our approach based on HMI-A is desirable both online (92.4%) and offline (98.83%).

D. Limitations of the HMI-A

Although the real-time performance of HMI-A is superior, there are still some limitations. It is found that the standard deviations for some online metrics are relative large, for instance the SD of MF shown in Fig. 8(c). It is possibly due to the signal similarity of the MF and RF. For some inexperienced subjects such as the S4 in Table I, the MF is frequently confused with RF. But for experienced ones, the confusion between the MF and RF hardly occurs.

As a kind of one-dimensional ultrasound, A-mode ultrasound cannot provide pictorial representation of anatomical structures of the forearm like B-mode ultrasound imaging. Only the muscle information on a predefined direction can be observed by A-mode ultrasound while the contraction can be fully detected by B-mode ultrasound. A slight shift of the A-mode ultrasound transducer will result in a significant variance of the signals, while the B-mode ultrasound will observe the same anatomical section when minor probe shift occurs. In other words, the signal robustness of A-mode ultrasound is comparatively worse than B-mode ultrasound. A potential solution for enhancing classification robustness is to subtly segment preprocessed signals prior to feature extraction, as verified in the experiment.

Compared with EMG signals, there are also some inherent limitations of A-mode ultrasound signals. The EMG signals have relation to subjective movement intentions as the electrical manifestation of neural control commands. However, the echo signals of A-mode ultrasound detect the muscle movements as a morphological reflection of motion intention, which cannot be interpreted in the perspective of neural rehabilitation. From the aspect of signal sources, the sEMG signals are more intuitive. Since the sEMG signals are related to movement intentions and generated with muscle contraction, there are nearly no sEMG signals during rest. According to this phenomenon, it is relatively advantageous for sEMG-based HMI to recognize rest state. But for HMI-A, the rest state is regarded as a motion. The user was asked to keep a consistent gesture during rest, for the sake of a high-level recognition accuracy on the rest motion.

Moreover, the current A-mode ultrasound system is a little bulkier, compared with some wearable sEMG system like the Tringo Wireless (Delsys Inc, USA). To the best of our knowledge, advanced electronic technologies have compressed an ultrasound system into such a small size as a smartphone, for instance the Vscan (General Electric, USA) and the Lumify (Philips Ltd. USA). Therefore, it is believed that ultrasound system can be further miniaturized with the usage of highly-integrated chips. Besides, active transducer could be designed to integrate with signal generator/receiver to further compact the size of an A-mode ultrasound system.

In addition, the power consumption of the current A-mode ultrasound system is nearly 6 W. With a 3000 mAh (12 V, 127 g) battery, it can work for about 6 hours in the continuous signal acquisition mode. To be lightweight and wearable, the power consumption need be further reduced.

Furthermore, to obtain high-quality ultrasound signals, the liquid coupling gel filled between the skin and transducer is needful, which will cause some inconvenience, somewhat but not significantly. From state-of-the-art studies, solid coupling gel has been available [50]. Designing a customized transducer that integrates with the solid gel may improve the usability of HMI-A.

Also, another limitation of HMI-A is that it is sensitive for wrist pronation and supination because of the changes of the relative position between transducers and muscles under the skin. In this condition, the classification accuracy will degrade significantly.

E. Future Work

The following tasks have been targeted in our future work: a) Transradial amputee subjects will be employed for the offline and online experiments, promoting HMI-A into real-life implementation for prosthetic hand control; b) Wearable ultrasound system like Vscan (General Electric, USA) will be developed and integrated into the forearm prostheses; c) Deep learning based on small ultrasound data set, inspired by Shi *et al.* [51], [52], will be applied for feature learning to pursue better finger motion control; d) The robustness of the HMI-A under transducer shift, donning and doffing, and wrist rotation will be studied to realize a robust HMI for prosthesis control, sign language recognition, and interaction games, etc; e) A-mode ultrasound and sEMG will be combined to predict hand gestures and force simultaneously.

APPENDIX

APPENDIX TABLE I
PARAMETERS OF CUSTOMIZED ULTRASOUND TRANSDUCER

Backing layer	
Material	Mixture of epoxy and tungsten powder
Acoustic impedance	8 MRayl
Thickness	3-4 mm
Piezoelectric material	
Material	1-3 composite material
Diameter	6 mm
Thickness	0.34 mm
Matching layer	
Material	Mixture of epoxy and tungsten powder
Acoustic impedance	4.5 MRayl
Thickness	0.12 mm
Performance parameter	
Center frequency	4.9 MHz
-6dB Bandwidth	56.4%

REFERENCES

- [1] G. N. Saridis and T. P. Gootee, "Emg pattern analysis and classification for a prosthetic arm," *IEEE Transactions on Biomedical Engineering*, no. 6, pp. 403–412, 1982.
- [2] 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.
- [3] N. Jiang, J. L. Vest-Nielsen, S. Muceli, and D. Farina, "Emg-based simultaneous and proportional estimation of wrist/hand kinematics in uni-lateral trans-radial amputees," *Journal of neuroengineering and rehabilitation*, vol. 9, no. 1, p. 42, 2012.
- [4] Y. Geng, P. Zhou, and G. Li, "Toward attenuating the impact of arm positions on electromyography pattern-recognition based motion classification in transradial amputees," *Journal of neuroengineering and rehabilitation*, vol. 9, no. 1, p. 74, 2012.
- [5] D. Yang, L. Jiang, Q. Huang, R. Liu, and H. Liu, "Experimental study of an emg-controlled 5-dof anthropomorphic prosthetic hand for motion restoration," *Journal of Intelligent & Robotic Systems*, vol. 76, no. 3-4, p. 427, 2014.
- [6] Y. Fang, H. Liu, G. Li, and X. Zhu, "A multichannel surface emg system for hand motion recognition," *International Journal of Humanoid Robotics*, vol. 12, no. 02, p. 1550011, 2015.
- [7] Y. Fang, D. Zhou, K. Li, and H. Liu, "Interface prostheses with classifier-feedback-based user training," *IEEE Transactions on Biomedical Engineering*, vol. 64, no. 11, pp. 2575–2583, 2017.
- [8] M. J. Zwartz and D. F. Stegeman, "Multichannel surface emg: basic aspects and clinical utility," *Muscle & nerve*, vol. 28, no. 1, pp. 1–17, 2003.
- [9] J.-Y. Guo, Y.-P. Zheng, Q.-H. Huang, X. Chen, J.-F. He, and H. L.-W. Chan, "Performances of one-dimensional sonomyography and surface electromyography in tracking guided patterns of wrist extension," *Ultrasound in medicine & biology*, vol. 35, no. 6, pp. 894–902, 2009.
- [10] C. J. De Luca, "Surface electromyography: Detection and recording," *DelSys Incorporated*, vol. 10, p. 2011, 2002.
- [11] N. Akhlaghi, C. A. Baker, M. Lahlou, H. Zafar, K. G. Murthy, H. S. Rangwala, J. Kosecka, W. M. Joiner, J. J. Pancrazio, and S. Sikdar, "Real-time classification of hand motions using ultrasound imaging of forearm muscles," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 8, pp. 1687–1698, 2016.

- [12] R. Vinjamuri, Z.-H. Mao, R. Scلابassi, and M. Sun, "Limitations of surface emg signals of extrinsic muscles in predicting postures of human hand," in *Engineering in Medicine and Biology Society, 2006. EMBS'06. 28th Annual International Conference of the IEEE*. IEEE, 2006, pp. 5491–5494.
- [13] B. Waryck, "Comparison of two myoelectric multi-articulating prosthetic hands." Myoelectric Symposium, 2011.
- [14] D. Yang, W. Yang, Q. Huang, and H. Liu, "Classification of multiple finger motions during dynamic upper limb movements," *IEEE journal of biomedical and health informatics*, vol. 21, no. 1, pp. 134–141, 2017.
- [15] D. R. Merrill, J. Lockhart, P. R. Troyk, R. F. Weir, and D. L. Hankin, "Development of an implantable myoelectric sensor for advanced prosthesis control," *Artificial organs*, vol. 35, no. 3, pp. 249–252, 2011.
- [16] E. A. Biddiss and T. T. Chau, "Upper limb prosthesis use and abandonment: a survey of the last 25 years," *Prosthetics and orthotics international*, vol. 31, no. 3, pp. 236–257, 2007.
- [17] J. A. Jensen, "Ultrasound imaging and its modeling," in *Imaging of Complex Media with Acoustic and Seismic Waves*. Springer, 2002, pp. 135–166.
- [18] D. Sierra González and C. Castellini, "A realistic implementation of ultrasound imaging as a human-machine interface for upper-limb amputees," *Frontiers in neurobotics*, vol. 7, p. 17, 2013.
- [19] P. Hodges, L. Pengel, R. Herbert, and S. Gandevia, "Measurement of muscle contraction with ultrasound imaging," *Muscle & nerve*, vol. 27, no. 6, pp. 682–692, 2003.
- [20] J.-Y. Guo, Y.-P. Zheng, Q.-H. Huang, and X. Chen, "Dynamic monitoring of forearm muscles using one-dimensional sonomyography system," *Journal of rehabilitation research and development*, 2008.
- [21] K. B. Kiesel, T. L. Uhl, F. B. Underwood, D. W. Rodd, and A. J. Nitz, "Measurement of lumbar multifidus muscle contraction with rehabilitative ultrasound imaging," *Manual therapy*, vol. 12, no. 2, pp. 161–166, 2007.
- [22] J. Shi, S.-x. Hu, Z. Liu, J.-Y. Guo, Y.-j. Zhou, and Y.-p. Zheng, "Recognition of finger flexion from ultrasound image with optical flow: A preliminary study," in *Biomedical Engineering and Computer Science (ICBECS), 2010 International Conference on*. IEEE, 2010, pp. 1–4.
- [23] J. Shi, Y.-P. Zheng, X. Chen, and Q.-H. Huang, "Assessment of muscle fatigue using sonomyography: muscle thickness change detected from ultrasound images," *Medical engineering & physics*, vol. 29, no. 4, pp. 472–479, 2007.
- [24] J. Shi, Y.-P. Zheng, Q.-H. Huang, and X. Chen, "Continuous monitoring of sonomyography, electromyography and torque generated by normal upper arm muscles during isometric contraction: sonomyography assessment for arm muscles," *IEEE transactions on biomedical engineering*, vol. 55, no. 3, pp. 1191–1198, 2008.
- [25] J. Shi, J.-Y. Guo, S.-X. Hu, and Y.-P. Zheng, "Recognition of finger flexion motion from ultrasound image: a feasibility study," *Ultrasound in medicine & biology*, vol. 38, no. 10, pp. 1695–1704, 2012.
- [26] C. Castellini and G. Passig, "Ultrasound image features of the wrist are linearly related to finger positions," in *Intelligent Robots and Systems (IROS), 2011 IEEE/RSJ International Conference on*. IEEE, 2011, pp. 2108–2114.
- [27] C. Castellini, G. Passig, and E. Zarka, "Using ultrasound images of the forearm to predict finger positions," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 20, no. 6, pp. 788–797, 2012.
- [28] C. A. Baker, N. Akhlaghi, H. Rangwala, J. Kosecka, and S. Sikdar, "Real-time, ultrasound-based control of a virtual hand by a trans-radial amputee," in *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the*. IEEE, 2016, pp. 3219–3222.
- [29] Y. Huang, X. Yang, Y. Li, D. Zhou, K. He, and H. Liu, "Ultrasound-based sensing models for finger motion classification," *IEEE Journal of Biomedical and Health Informatics*, vol. PP, no. 99, pp. 1–1, 2017.
- [30] 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.
- [31] N. Hettiarachchi, Z. Ju, and H. Liu, "A new wearable ultrasound muscle activity sensing system for dexterous prosthetic control," in *Systems, Man, and Cybernetics (SMC), 2015 IEEE International Conference on*. IEEE, 2015, pp. 1415–1420.
- [32] Y. Li, K. He, X. Sun, and H. Liu, "Human-machine interface based on multi-channel single-element ultrasound transducers: A preliminary study," in *e-Health Networking, Applications and Services (Healthcom), 2016 IEEE 18th International Conference on*. IEEE, 2016, pp. 1–6.
- [33] T. A. Kuiken, G. Li, B. A. Lock, R. D. Lipschutz, L. A. Miller, K. A. Stubblefield, and K. B. Englehart, "Targeted muscle reinnervation for real-time myoelectric control of multifunction artificial arms," *Jama*, vol. 301, no. 6, pp. 619–628, 2009.
- [34] B. Peerdeman, D. Boere, H. Witteveen, H. Hermens, S. Stramigioli, J. Rietman, P. Veltink, S. Misra *et al.*, "Myoelectric forearm prostheses: State of the art from a user-centered perspective," 2011.
- [35] T. L. Szabo, *Diagnostic ultrasound imaging: inside out*. Academic Press, 2004.
- [36] 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.
- [37] 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.
- [38] X. Zhu, J. Liu, D. Zhang, X. Sheng, and N. Jiang, "Cascaded adaptation framework for fast calibration of myoelectric control," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 3, pp. 254–264, 2017.
- [39] G. Li, A. E. Schultz, and T. A. Kuiken, "Quantifying pattern recognition-based myoelectric control of multifunctional transradial prostheses," *IEEE transactions on neural systems and rehabilitation engineering: a publication of the IEEE Engineering in Medicine and Biology Society*, vol. 18, no. 2, p. 185, 2010.
- [40] 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.
- [41] A. J. Young, L. J. Hargrove, and T. A. Kuiken, "The effects of electrode size and orientation on the sensitivity of myoelectric pattern recognition systems to electrode shift," *IEEE transactions on bio-medical engineering*, vol. 58, no. 9, p. 2537, 2011.
- [42] —, "Improving myoelectric pattern recognition robustness to electrode shift by changing interelectrode distance and electrode configuration," *IEEE transactions on bio-medical engineering*, vol. 59, no. 3, p. 645, 2012.
- [43] L. Hargrove, K. Englehart, and B. Hudgins, "A training strategy to reduce classification degradation due to electrode displacements in pattern recognition based myoelectric control," *Biomedical Signal Processing and Control*, vol. 3, no. 2, pp. 175–180, 2008.
- [44] J. He and X. Zhu, "Combining improved gray-level co-occurrence matrix with high density grid for myoelectric control robustness to electrode shift," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. PP, no. 99, pp. 1–1, 2016.
- [45] C. Disselhorst-Klug, T. Schmitz-Rode, and G. Rau, "Surface electromyography and muscle force: limits in semg-force relationship and new approaches for applications," *Clinical biomechanics*, vol. 24, no. 3, pp. 225–235, 2009.
- [46] H. S. Ryaite, A. Arora, and R. Agarwal, "Study of issues in the development of surface emg controlled human hand," *Journal of Materials Science: Materials in Medicine*, vol. 20, no. 1, p. 107, 2009.
- [47] Y. Fang, N. Hettiarachchi, D. Zhou, and H. Liu, "Multi-modal sensing techniques for interfacing hand prostheses: a review," *IEEE Sensors Journal*, vol. 15, no. 11, pp. 6065–6076, 2015.
- [48] 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.
- [49] T. R. Farrell *et al.*, "A comparison of the effects of electrode implantation and targeting on pattern classification accuracy for prosthesis control," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 9, pp. 2198–2211, 2008.
- [50] R. Chivers, "Coupling material," 2006.
- [51] J. Shi, S. Zhou, X. Liu, Q. Zhang, M. Lu, and T. Wang, "Stacked deep polynomial network based representation learning for tumor classification with small ultrasound image dataset," *Neurocomputing*, vol. 194, pp. 87–94, 2016.
- [52] J. Shi, X. Zheng, Y. Li, Q. Zhang, and S. Ying, "Multimodal neuroimaging feature learning with multimodal stacked deep polynomial networks for diagnosis of alzheimer's disease," *IEEE journal of biomedical and health informatics*, 2017.



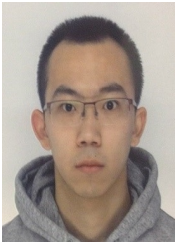
Xingchen Yang received the Bachelor's degree in mechanical engineering from Tianjin University, Tianjin, Shanghai, China, in 2015. He is currently pursuing the Ph.D. degree in the School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai, China.

His research interests include biological signal processing and machine learning.



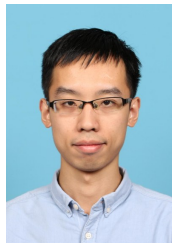
Xueli Sun received the Bachelors degree in mechanical engineering from Xi'an Jiao Tong University, Xian, China, in 2015. He is currently working toward the Masters degree in the School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai, China.

His research interests lies in ultrasound sensing technology and biological signal processing.



Dalin Zhou received the B.S. degree in automation from the University of Science and Technology of China, in 2012. He is currently pursuing the Ph.D. degree with the Intelligent System and Biomedical Robotics Group, University of Portsmouth, Portsmouth, U.K.

His research interest lies in ultrasound signal processing on hand gesture and prediction and the development of the ultrasound transducers.



Yuefeng Li received the Master's degree from the School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai, China, in 2017.

His research interests are signal processing and machine learning.



Honghai Liu received the Ph.D. degree in robotics from King's College London, London, U.K., in 2003.

He is the Chair Professor of Intelligent Systems and Robotics, University of Portsmouth, Portsmouth, U.K. His research interests include biomechatronics, pattern recognition, intelligent video analytics, intelligent robotics, and their practical applications with an emphasis on approaches that could make contribution to the intelligent connection of perception to action using contextual information.

Prof. Liu is a fellow of the Institution of Engineering and Technology. He is an Associate Editor of the IEEE TRANSACTIONS ON HUMAN MACHINE SYSTEMS, IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS, and IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS.