A-mode Ultrasound Driven Sensor Fusion for Hand Gesture Recognition

1st Peter Boyd  
School of Computing,  
University of Portsmouth  
Portsmouth, UK  
Peter.Boyd@port.ac.uk

2nd Honghai Liu  
School of Computing,  
University of Portsmouth  
Portsmouth, UK  
Honghai.Liu@port.ac.uk

Abstract—Traditionally, Surface electromyography (sEMG) has been the predominant method of sensing muscle activity in order to control myoelectric prosthesis. While many prosthesis control schemes used simple direct control, an ever increasing focus has moved to pattern recognition based approaches which promise a greater degree of natural control such as to further improve an amputees quality of life. Although pattern recognition based approaches have shown great promise, they have innate limitations due to changes that may occur during long term use which prevent clinical acceptance. Due to these limitations, researchers have increasingly investigated alternative modalities to provide more robust control schemes. A particular modality that has seen increasing interest is ultrasound based sensing due to its capability to better understand deep tissue activity.

Within this research, A-mode ultrasound based sensing is proposed not as a replacement for sEMG based sensing but instead to augment and drive sEMG based sensing during activities that may otherwise prove challenging to traditional sEMG based control schemes.

Index Terms—A-mode Ultrasound sEMG Hand Motion Recognition Rehabilitation

I. INTRODUCTION

Bio-Signal controlled prosthesis have long been seen as important tools in providing higher quality of life to amputees [1]. Classical approaches use simple direct control schemes that are functional although not fully intuitive, subsequently causing longer periods of rehabilitation and less freedom of use than desirable. In order to overcome this challenge, research has increasingly delved into various sensing modalities that can not only provide an intuitive and natural sense of control to amputees [2] but also one which is robust enough such that daily use can be achieved with minimal training or calibration.

Frequently in both clinical and academic environments, sEMG based sensing has been the chosen modality for these control schemes. The typical approach to detecting sEMG activity is to detect the period of rapid depolarization and repolarization seen within excited muscle fibres, creating an Action Potential (AP) [3]. Typically the relationship between muscle contraction and action potential is exploited in sEMG sensing to detect the collective impulse of the excited muscle fibres within a motor unit can be considered as the Motor unit action potential (MUAP). It can therefore be described that the raw sEMG signal is the output of detecting the recruitment of MUAPs across a muscle and their subsequent firing rates [4]–[6].

Although these MUAPs provide a reasonably high fidelity description of muscle activity provided the proper processing, they also suffer limitations due to the way in which transient changes may occur during long term use when using sEMG based sensing. These transient changes may include electrode shift [7], crosstalk, fatigue [8], changes to skin conductivity, time [9]–[11], and concept drift.

Ultrasound based sensing has long been used as a non-invasive diagnostic device in observing musculoskeletal disorders such as arthritis [12], [13]. Typically this method uses an array of piezoelectric transducers to project a focused wave of ultrasound into the body between a range of 2-20MHz. As these beams interact with varying tissue densities, echos are formed which can be measured in order to infer aspects of the muscles state during contraction [14], [15].

Through this capability to get a reliable stream as to tissue and muscle states within the human body, the application of ultrasound has been found viable in hand motion recognition in roles such as accurately predicting finger location [16], [17].

Within this research, A-mode ultrasound based sensing is proposed not as a replacement for sEMG based sensing but instead to augment and drive sEMG based sensing during activities that may otherwise prove challenging to traditional sEMG based control schemes. A common challenge found in sEMG based sensing is that of large arm movements and in situations where the location of an arm is not in the comfortable position frequently seen within laboratory experiments. As the arm moves into dynamic poses there exists a degree of shift and co-activation of muscle groupings surrounding the targeted muscles which subsequently impact recognition accuracy. The issue is further amplified when considering larger muscle groups in the biceps. Therefore it is proposed that the improved reading of deep muscle activity form ultrasound based sensing can help overcome issues that may otherwise occur during larger arm gestures, as to provide a robust multimodal solution for long term prosthesis control.

II. RELATED WORKS

The topic of provision of ultrasound based hand motion recognition has been explored by many researchers. Hodges
demonstrated that it is possible to observe low level of muscle contractions where the physical muscle architecture changes the most [18] making ultrasound promising when observing smaller muscle activation’s. A further aspect is the nature of ultrasound based sensing makes it robust to crosstalk that is often seen in fatigued muscles [19]. Perhaps one of the more promising aspects of ultrasound based sensing is that of the recorded performance in accurately detecting individual finger positions [20]–[22]. Expanding on this work, other researchers have found that ultrasound provides great promise in detecting multiple finger based hand motions [23].

While it has been shown that ultrasound may complement sEMG during wrist extension activities, the space required to mount both an ultrasound probe and sEMG electrodes in the same location was found to be difficult [24]. Within laboratory settings, B-mode ultrasound has been shown to be highly reliable in providing accurate hand motion recognition [25], [26]. As research has progressed, newer ultrasound based devices have been proposed which are increasingly [22], [27]. Further aspects of integrating multiple sensors into these devices has been explored by numerous researchers, largely towards combining sEMG and ultrasound capabilities into a single device [28], [29], especially with A-mode based devices.

Interest in ultrasound based hand motion recognition has increased a wearable ultrasound based devices for hand motion recognition continue to be developed and show great promise in the realm of proportional control and accurate detection in comparison to sEMG based sensing. However, these methods are frequently used as comparisons to sEMG based sensing and are seldom integrated with sEMG based sensing [23], [30], [31].

This paper is structured as follows: Firstly the system and experimental design for this study are described. Secondly a description is given of methods used to perform data processing and evaluation is given. Finally the results of this study are detailed, a discussion is made as to the implications of these results, and finally a conclusion is given with potential future avenues of research into this topic.

III. HYPOTHESIS

The Hypothesis of this study is therefore that the addition of a-mode ultrasound based sensing to sEMG based sensing can provide a meaningful increase to hand motion recognition during hand motion activation with larger arm movements.

IV. SYSTEM AND MATERIALS

A. Participants

Two able bodied participants aged 26 and 32 were used in this study. Both participants had no history of limb injury and had previous experience with sEMG but not ultrasound. Prior to the study all participants were informed of the contents of the experiments and the series of gestures to be performed. All work done in this study was approved by and done in accordance with the local ethical committee.

B. A-Mode Ultrasound Hardware

A-mode Ultrasound manifests itself as a linear set of time points and amplitudes that describe the echo intensity from the Ultrasound diode.

The ultrasound data used in this study was collected through a 4 channel A-Mode Ultrasound device that collected 1000 data points (or time dots) at a rate of 10MHZ with an inspection depth of around 39mm. The device was placed on the muscle grouping on the lower forearm, above the wrist of the candidate. To maintain a secure fitting for the US sensors, an elastic skin tape was used to secure the sensors in their correct location along the subjects arm. Hypoallergenic Ultrasound gel was applied prior to each set of data collection as to ensure improve the transmission quality of the ultrasound signals and to reduce interference from air gaps. Data transmission was performed via Ethernet from the ultrasound hardware to a windows 10 based machine.

C. sEMG Hardware

The sEMG based sensing utilized an sEMG data collection device capable of recording 16 channels with 12 bits ADC resolution and 1Khz sampling frequently as is typically used within literature [32]–[34]. Gathered data was within the 10Hz to 500Hz range through a hardware based band pass filter. As power-line noise can degrade the quality of the collected signals, 50Hz power-line noise was removed through a hardware based notch filter and a software based comb filter, the usage of these two filters also aid in removing motion artefacts from the larger arm movements. Data transmission was performed via USB and processed through the same Windows 10 based machine as the ultrasound data. Further information about the hardware device can be found in [35], [36]. The participant was fitted with 6 pairs of bi-polar wet
Fig. 2. Multimodal Ultrasound and sEMG experimental setup

electrodes distributed on the major muscle sites of the arm as to provide muscle activation information during larger arm movements. The chosen electrode locations are the lateral deltoid, biceps brachii shorthead, supinator, and pronator teres and ticeps brachii. The fitting of the wet electrodes and a-mode ultrasound hardware can be seen in figure 2. Each electrode location was cleaned with an alcohol wipe prior to electrode fitting.

V. METHODOLOGY

A. Data Capture

The data capture was divided into 5 trials with ten gestures performed per trial. In order to acquire a dynamic range of motion activity across the whole arm five motion primitives were selected and performed in their open pose and inverse totalling in ten hand and arm motions. The chosen primitives were as follows in order Hand Open (HO), Hand Closed (HC), Forearm Pronation (FP), Forearm Supination (FS), Rotation In (RI), Rotation Out (RO), Humerus Forward (HF), Humerus Backward (HB), Wrist Flexion (WF), Wrist Extension (WE). Each individual motion was performed for five seconds before shifting to its inverse motion and then a rest period of ten seconds between motion primitive. Each trial lasted for a total of 110 seconds including the initial rest period with a total of 50 seconds of motion activity. There was no change of the electrode or ultrasound sensor placement between trials as to introduce a degree of the natural shift that may form with daily use of this combined sensing modality while also removing potential larger shift from donning and doffing.

B. Data Processing

All data processing was completed using Matlab r2017b. The first step of pre-processing for both the ultrasound and sEMG data was to remove 0.5 second of motion data from the beginning and end of each 5 second gesture performance as to remove transition data. The resulting data set provided

4 seconds of stable motion data from each 4 second gesture performed with 5 trials in total per participant of all gestures providing a total of 20 seconds of gesture data across all trials.

The A-mode ultrasound data consists of a single frame every 100ms containing 1000 time points, these time points indicating the muscle activity at a given depth from the US Diode. The starting and ending 20 time points of each frame of ultrasound data were removed as the information carried here was not considered meaningful with the initial 20 time points holding largely skin boundary data and final 20 time points being unlikely to properly represent any usable muscle boundaries that it comes into contact with, resulting in each frame holding 960 time points. Presently, there exists little comparative US feature selection strategies or comparisons. Therefore, traditional feature extraction methods for EMG data were to be modified to better exploit the generalizable traits of the data. The approach to feature extraction was to operate directly on the time points within each frame, as opposed to across multiple frames, using a 120ms window and a 30ms sliding window. It was shown in [37] that the Waveform Length feature provides a robust performance in A-mode ultrasound based hand motion recognition and is resilient to diode shift.

Waveform length can be considered to be the measure of complexity of the sEMG signal over a specific time segment [38]

\[ WL = \sum_{i=1}^{n-1} |x_{i+1} - x_i| \] (1)

As stated above, transient changes in the signal were removed as to assure that only stable signal is used for training and testing. The feature extraction for sEMG signal was performed using a 250ms window with a 50ms sliding window, this time was selected as it provided the maximum degree of time for activity detection while also being below the quantity of time that can create a perceivable delay to a person [39].

the Root mean square (RMS) feature was selected to represent the sEMG signal in this study. The RMS feature can be considered as the square root of the mean values for a set of squared raw sEMG values over a period of time. This method
RMS $= \sqrt{\frac{1}{N} \sum_{i=1}^{N} x_i^2}$  \hspace{1cm} (2)

C. Evaluation

Within this study, the three conditions were evaluated as unimodal sEMG, unimodal Ultrasound, then finally multimodal Ultrasound driven sEMG. Synchronisation of the sEMG signal and Ultrasound signal was performed through manual recognition of the first muscle activation of each gesture set in sequence during the multimodal condition.

All classification was performed using LDA with naive bayes. The decision to select LDA is due to the classifier’s robustness to changes in the transient changes in the signal when compared to other traditional classifiers, such as SVM, as demonstrated in literature [43]–[46]. Classifier training was performed using a single trial that was evaluated against other 4 trials for each participant.

VI. RESULTS

To verify the performance of the ultrasound led method, each set of 5 trials were evaluated on the basis of sEMG alone, a-mode ultrasound alone, and finally the a-mode ultrasound led approach.

As seen in fig.5, the sEMG unimodal approach demonstrated the poorest performance overall, achieving an accuracy of 71.77% across both participants and all trials. The A-mode ultrasound unimodal approach demonstrated some improvement over unimodal sEMG with an accuracy of 77.43%. The strongest modality for performance was the ultrasound led multimodal approach with an accuracy of 80.21%.

It can be inferred from this result that sEMG unimodal sensing is more susceptible to reduced classification accuracy during larger arm motions where crosstalk and interference from cabling may be more apparent. Furthermore, while the sEMG sensor locations and electrodes used should ensure decent coverage and signal collections, it is quite possible that the activating muscle groupings in larger gestures may become less separable than lower arm gestures alone. An interesting result is that of the performance from the unimodal ultrasound approach, which although only focused on the forearm area had managed to achieve reasonable accuracy during larger arm motions. The cause of this performance increase may likely come from smaller muscle changes that occur due to physiological changes of the arm as opposed to active MU activity, something which would explain the poorer results from unimodal sEMG. The results of Ultrasound led multimodal approach are quite promising in displaying that the ultrasound can provide a high degree of accuracy in the hand motions while further improving sEMG based sensing. The potential for ultrasound to better recognize subtle changes in muscle boundaries is promising for future usage of a-mode ultrasound.

VII. DISCUSSION

As demonstrated in the results section, the proposed multimodal a-mode ultrasound led control scheme showed promising performance over unimodal implementations of sEMG and ultrasound based sensing. The implications that can be inferred from this is that while sEMG based solutions are capable of providing high fidelity information into muscle group activity, when undesirable arm positions are obtained then robustness may not always be guaranteed. Much like other researchers who have seen improved performance in monitoring muscle grouping activity for smaller adjustments within the hand and arm [16], [17], [20]–[22]. Furthermore this promising result reaffirms findings of other researchers towards augmenting sEMG based sensing through ultrasound [23]. Although this study differs by using sEMG to bolster ultrasound in providing
extra context as to larger arm movements while the ultrasound signal focuses on the hands activity.

Subsequently, through the usage of a-mode ultrasound led sEMG based sensing there was a clear improvement in robust sensing in larger arm movements beyond that of the constituent modalities, placed in context of the investigation into ultrasound diode shift [37], [47] and present research into sEMG sensor shift [48], there exists an argument that the ultrasound based approach provides robustness to larger arm movements, whereas sEMG based sensing may improve the robustness to sensor shift during daily use.

VIII. CONCLUSION

In this study, a multi modal a-mode ultrasound led sensing platform was implemented and demonstrated to be feasible during larger arm motions particularly in comparison to unimodal sEMG and furthermore was shown to contribute to improving the quality of sEMG based sensing in this area.

It is suggested that future directions in Ultrasound hand motion recognition is to investigate whether the inclusion of more channels when considering sensor shift may further improve the classification accuracy alongside the impact of larger arm motions. Furthermore, it is suggested to further investigate the relative comparison of Ultrasound based hand Motion Recognition, in comparison and in complement to sEMG when considering Long Term Use. Finally, it is suggested to explore the capability of ultrasound led hand and arm motion recognition during more dynamic tasks as a multi modal approach with sEMG sensing.

REFERENCES


