Effect of Lateral Resolution on Classifying Individual Finger Flexions using Ultrasound

A. J. Fernandes[1] Y. Ono[1] and E. Ukwatta[1,2]

[1]Department of Systems and Computer Engineering, Carleton University, Ottawa, Canada
[2]School of Engineering, University of Guelph, Guelph, Canada

Abstract— **B-mode ultrasound imaging has recently shown promise in achieving higher classification accuracies than surface electromyography for predicting discrete hand gestures and individual finger movements. This preliminary study investigates the performance in classifying finger flexions when reducing the lateral sampling interval resolution of a conventional clinical ultrasonic imaging probe with data collected from one subject. An experiment using spatial and temporal features, extracted from ultrasound radio-frequency (RF) signals are used with linear discriminant analysis to classify individual thumb, index, middle, ring and pinky finger flexion movements. The spatial lateral sampling interval is increased from 315 µm to 10 mm (reduction in lateral resolution) by averaging four groups of 32 consecutively acquired A-mode ultrasound RF signals from a 40 mm probe. The results for the four averaged RF ultrasound signals with a 10 mm lateral sampling interval had an F_1 score ranging between 77-91% with a classification accuracy of 84% for all five finger flexions. This classification accuracy was similar when using the acquired 315 µm lateral resolution and decreases to a classification accuracy of 32% for no lateral resolution, when the full 40 mm width is averaged into a single RF signal. The results show motivation for using a wearable multichannel ultrasound device for predicting individual finger flexions for prosthetic devices.**

Keywords— **Ultrasound RF signal acquisition, lateral resolution, finger motion classification, prosthetic hand control.**

I. Introduction

Upper limb amputations have a significant impact on both the physical and mental aspects of daily lives for trauma-related amputees. Over one million Americans have suffered from trauma related upper limb amputations [1]. A study conducted by the National Physical and Sensory Disability Database in Ireland surveyed 148 people with major limb amputation found that 89% experienced some form of daily difficulty and 41% described that these difficulties interfered severely or extremely with their lives [2].

Modern prosthetic industrial devices such as “i-digits” by Touch Bionics uses surface electromyography (sEMG) as the sensing mechanism to detect muscle activity to perform up to 32 different hand gestures and grip motions [3]. Previous classification accuracies for using eight channel sEMG signals have been reported to have accuracies averaging around 89% when classifying 15 different finger movement positions [4]. Recently ultrasound as an alternative control sensor has been found to produce superior results in classifying individual finger flexions and hand gestures. Sikdar et al. were able to obtain 98% accuracy implementing B-mode ultrasound imaging when classifying complete individual finger flexions of four fingers, claiming that ultrasound techniques are superior by being able to spatially resolve the deep layered tissue muscle activity compared to the limited specificity from cross-talk occurring for multichannel sEMG [5]. McIntosh et al. used a multilayered perceptron neural network with 15 nodes in the hidden layer and optical flow image processing to extract spatial and time varying features from B-mode ultrasound imaging to differentiate between flexing at different joints and obtained an average accuracy of 97% when classifying discrete hand gestures [6]. Yang et al. used four A-mode ultrasonic transducers to achieve 95% accuracy in classifying 11 different hand gestures in real-time [7], however does not provide results for individual finger flexions proving to be useful for American Sign Language applications but undetermined for finger flexion prosthetic devices.

These high accuracies produced in [5, 6] use a bulky conventional ultrasonic clinical probe in B-mode imaging and may not be practical to implement in a prosthetic device for daily life activities outside clinics or laboratories. This paper presents a preliminary study that investigates spatial and temporal feature extraction techniques to classify finger flexions using ultrasound. A clinical ultrasound probe is used to study an effect of reducing the lateral resolution on the classification accuracy in motivation to explore a feasibility of wearable ultrasonic sensors for prosthetic hand control [8].

II. Methods And Materials

A. Data Acquisition Procedure

Ultrasound data acquisition for this study was approved by Carleton University’s Research Ethics Board. A clinical ultrasound imaging system (Model PICUS from ESAOTE Europe, Maastricht, Netherlands) acquired 127 ultrasound RF signals using a 40-mm linear array probe with a lateral sampling resolution of 315 µm per RF signal. Each RF signal was sampled at 33.3 MHz (30 nsec), corresponding to an axial (depth) sampling resolution of 23.1 µm, assuming an ultrasound speed of 1540 m/s in soft tissues.

The ultrasound data were acquired with a healthy male subject. The probe was setup to access the anterior side of the upper forearm held by a fixed apparatus shown in Figure 1. The probe was fixed 5-cm away from the wrist, and the 40-mm wide probing surface covered most of the forearm width. Ultrasound RF signals were acquired during the individual motions of the thumb, index, middle, ring and pinky finger flexions. The RF signal pattern changed due to the tissue motions caused by muscle contractions corresponding to the combination of flexor digitorum profundus, flexor digitorum superficialis and/or flexor pollicis longus for the chosen finger flexion. Each acquisition had the hand positioned in a natural resting state, the chosen finger completed a full 180° flexion, then returned to the natural resting state, over a 6 second duration. Such a motion was repeated to obtain three trials per finger at 30 frames per second, totaling at 180 frames per recording.

Fig. 1 Procedure setup with the ultrasound probe fixed laterally to the anterior side of the upper left forearm.

B. Signal Preprocessing and Feature Extraction

The acquired ultrasound RF signal data were structured as illustrated in Figure 2 a) such that a single frame along the temporal (frame) direction, *t*-axis, is composed of 127 RF signals along the lateral (scan) direction, *l*-axis, with each RF signal composing of 1516 points along the depth (axial) direction, *d*-axis. The acquired signals were then restructured into an N number of new averaged RF signals confined to the 40 mm width. The lateral sampling interval between the reconstructed RF signals was increased to $(127 / N) \times 315$ µm. The new RF signals were created by averaging every $127 / N$ group of acquired RF signals. Each ultrasound recording with $l \times d \times t$ axes, was reduced from $127 \times 1516 \times 180$ to $N \times 1516 \times 180$. For example, when $N = 4$, the lateral sampling interval distance between each RF signal is changed from 315 µm to 10 mm in the 40 mm lateral scan width.

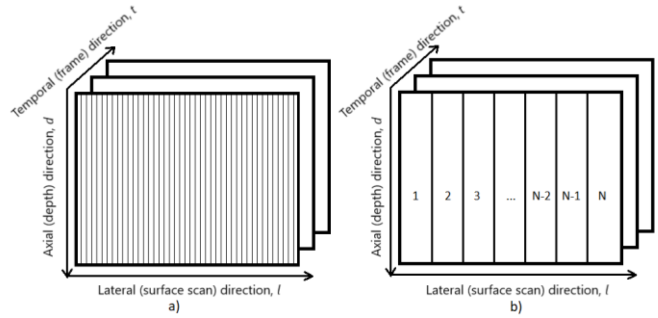

Fig. 2 Structure of the acquired ultrasound RF signal data a) with 127 lateral RF signals and b) N averaged RF signals.

Figure 3 shows signal processing and feature extraction techniques employed in this study, which were adapted from the method proposed by Yang et al. [7]. The acquired RF signals in a) were averaged to obtain the new RF signal which also is normalized and filtered using linear regression and Gaussian filtering in b). The envelope of the signal is then obtained using the absolute value of the Hilbert transform in c), and the amplitude is reduced using log compression in d). Linear regression is finally applied to every 15 data points to obtain 100 y-intercept b and slope m features in e). A total of $100 \times N$ spatial features were extracted from each frame such that the b and m features would be used as a substitute to pixels acquired in standard B-mode imaging.

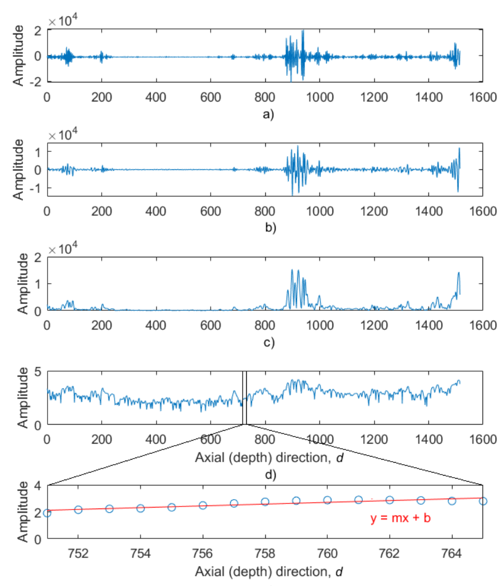

Fig. 3 Preprocessing and feature extraction on a) raw RF signal acquired, b) normalized and Gaussian filtered signal, c) envelope of signal, d) log compression, and e) linear fit features extracted using linear regression.

C. Feature Processing and Pattern Recognition

The extracted features were then processed to obtain temporal information using these features. This addition allows for classifying the frame motions for the thumb, index, middle, ring, and pinky finger flexions as opposed to the original intensions for recognizing hand gestures by classifying discrete hand positions [7].

Table 1 Feature set obtained for each frame.

Feature	Description
$b - b_{ref}$	Spatial intensity difference from resting frame (y-intercept)
$m - m_{ref}$	Spatial intensity difference from resting frame (slope)
Δb	Instantaneous change between frames (y-intercept)
Δm	Instantaneous change between frames (slope)

The extracted b and m features were normalized to represent spatial changes from the rest state by subtracting the corresponding b and m features of the first frame for each frame in the recording (b_{ref} and m_{ref}). The first frame was chosen because all recordings had the hand start in a rested position. A novel method for extracting temporal features was obtained by applying a first order difference equation to the b and m features such that for each frame, the corresponding b and m values would be subtracted by the previous frame's b and m values.

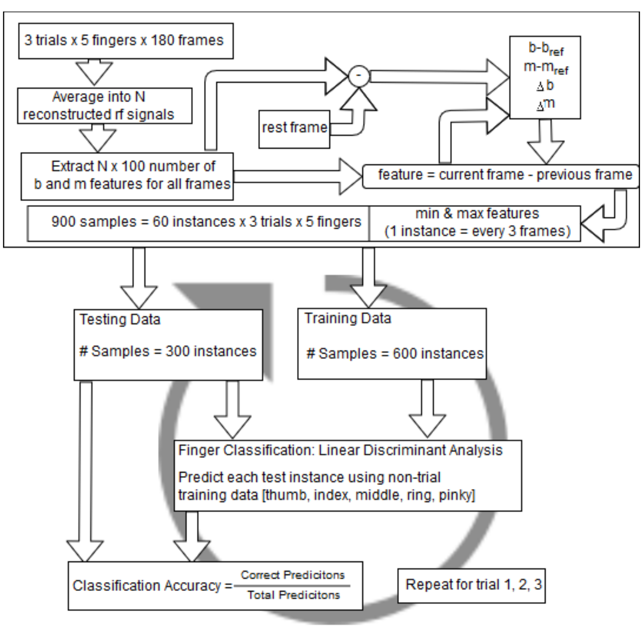

Fig. 4 Flowchart of procedure to estimate classification accuracies.

Figure 4 illustrates the flowchart of the procedure to classify the individual finger flexion activity and to estimate the classification accuracy. To reduce the number of computations, every three frames in each recording of 180 frames was treated as a single sample, resulting in a total of 900 samples (3 trials × 5 fingers × 180 frames / 3 sequential frames). From the three sequential frames in each sample, the minimum and maximum b-b_{ref}, m-m_{ref}, Δb, and Δm values for all 100 depth segments and N averaged RF signals was extracted as the features resulting in 800 × N total features per sample (2 values as min & max × 4 feature set values × 100 depth segments × N averaged RF signals). Linear discriminant analysis was used to classify the data using three folds based on the trial number, such that the test data would consist of one trial recording of all 5 fingers and the training data would consist of the other trials. For example, to classify the frame motions for index trial 1 (6 seconds of recording), the training data would use the other two 6 second recording trials for the index finger along with two trails from the other four fingers, ensuring class balance when determining the prediction results for index trial 1.

III. Results And Disucssions

The classification accuracy is defined to be the percentage of the total number of correct finger predictions out of the total number of samples in the data set (900 samples total). The number of averaged RF signals was varied from 1 to 127 and resulted in the following asymptotic trend in Figure 5 with a coefficient of determination to be 0.9332. As the number of reconstructed RF signals (N) is increased (across a 40 mm width of the ultrasound probing surface), the classification accuracy plateaus around N = 4, which is represented as having a 10 mm lateral sampling interval between the RF signals, showing that the full lateral resolution may not be needed when predicting individual finger motions.

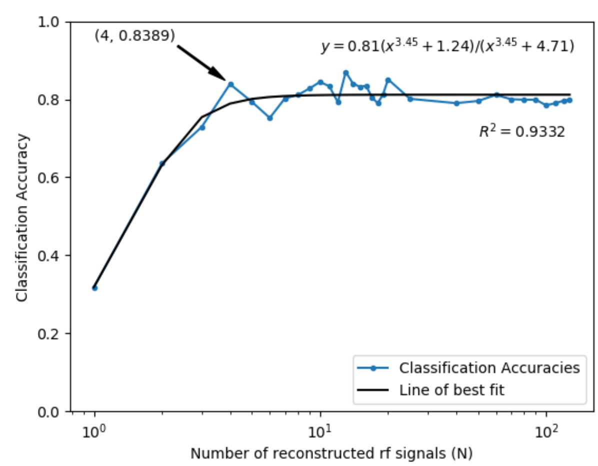

Fig. 5 Classification accuracies verses number of reconstructed RF signals

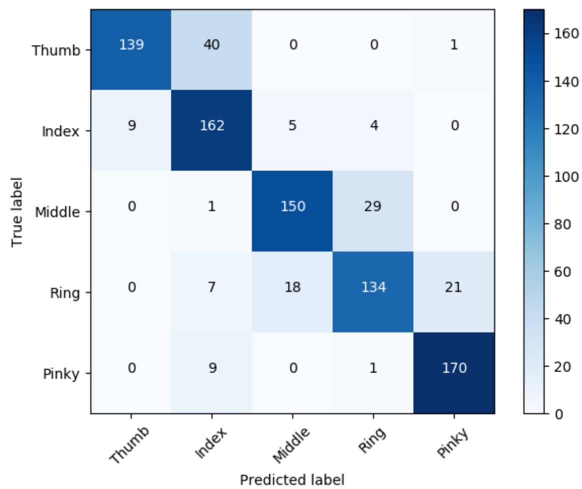

Fig. 6 Confusion matrix for N = 4 reconstructed RF signals.

Looking at the specific example of when N = 4 shown by the confusion matrix in Figure 6, the effective lateral sampling interval between each reconstructed RF signal becomes 10 mm and results in a classification accuracy of 83.89%. The precision and recall metrics are calculated in Table 2, which shows that the ring finger has the lowest calculated F_1 score at 0.7701 and the pinky finger to have the highest calculated F1 score at 0.9135.

Table 2 Precision and recall metrics for N = 4 reconstructed RF signals

	Precision	Recall	F_1 Score
Finger	$\frac{TP}{(TP + FP)}$	$\frac{TP}{(TP + FN)}$	$2 \times \frac{Precision \times Recall}{Precision + Recall}$
Thumb	93.9%	77.2%	0.847
Index	74.0%	90.0%	0.812
Middle	86.7%	83.3%	0.850
Ring	79.8%	74.4%	0.770
Pinky	88.5%	94.4%	0.914

The results produced in this preliminary study gives evidence that a high lateral spatial resolution may not be required for a possible wearable ultrasound system composed of a reduced number of transducers. However, when the effective lateral resolution is reduced to a single RF signal (no spatial variation in the lateral axis) the classification accuracy reduces significantly, giving evidence on the importance of being able to spatially distinguish deep layer muscle activity along the lateral axis. The features extracted in this study would be improved in the future by using more than just the envelope of the RF signals, and compared alongside traditional sEMG methods.

IV. Conclusion

Although these results may be limited in this preliminary study from acquiring data from only one subject, these results show motivation for using wearable multichannel ultrasound to predict individual finger flexions for prosthetic devices rather than using a more complex bulky ultrasonic imaging probe and system. Implementation of multiple wearable ultrasonic sensors [8] has a potential to perform individual finger flexions in a less costly and more practical manner than a convention clinical ultrasound probe and imaging system.

Acknowledgment

This work was supported by Natural Sciences and Engineering Research Council of Canada.

References

1. Ziegler-Graham K, MacKenzie E, Ephraim E et al. (2008) Estimating the prevalence of limb loss in the United States: 2005 to 2050. Arch Phys Med Rehabil 89:422-9.
2. Gallagher P, O'Donovan M, Doyle A et al. (2011) Environmental barriers, activity limitations and participation restrictions experienced by people with major limb amputation. Prosthet Orthot Int. 35:278-84.
3. Touch Bionics Ltd (2018) i-digits: User Manual. [Online]. Available: http://oc4assets.azurewebsites.net/library/39600
4. Geethanjali P, Raunak V (2018) Identification of a feature selection based pattern recognition scheme for finger movement recognition from multichannel EMG signals. Phys Eng Sci Med 22:549-559.
5. Sikdar S, Rangwala H, Eastlake E et al (2014) Novel method for predicting dexterous individual finger movements by imaging muscle activity using a wearable ultrasonic system. IEEE Trans Neural Syst Rehabil Eng. 22:69-76.
6. McIntosh J, Marzo A, Fraser M et al (2017) EchoFlex: Hand Gesture Recognition using Ultrasound Imaging. CHI Conference on Human Factors in Computing Systems, Denver, CO, USA, 2017, pp 1923-1934.
7. Yang X, Sun X, Zhou D et al. (2018). Towards wearable A-mode ultrasound sensing for real-time finger motion recognition. IEEE Trans Neural Syst Rehabil Eng. 26:1199-1208.
8. AlMohimeed I, Turkistani H, Ono Y (2013) Development of wearable and flexible ultrasonic sensor for skeletal muscle monitoring. Proc. IEEE Ultrasonics Symposium, Prague, Czech Republic, 2013, pp 1137-1140.

Author: Alexander Fernandes
Institute: Carleton University
Street: 1125 Colonel By Dr.
City: Ottawa
Country: Canada
Email: alexanderfernandes@cmail.carleton.ca