On the Gesture Recognition of a Faint Phantom Motion for the Control of a Transradial Prosthesis amidst varying Contraction Forces

Ejay Nsugbe*
Nsugbe Research Labs, U.K., ennsugbe@yahoo.com

Abstract: The variation of the contraction force associated with the phantom motion used for the actuation of a bionic upper-limb prosthesis represents a scenario encountered regularly by amputees, while prior research appears to not have been able to succinctly address this problem. In this study, an extended prosthesis control system is proposed which is able to recognise gesture intent motions alongside the prediction of an associated contraction force as part of an advanced pattern recognition system. As part of this research topic, this paper introduces the proposed control architecture and is based on the solving of the gesture recognition problem amidst varying contraction forces for a transradial amputee with a seemingly faint phantom motion. The work involves the application of a novel decomposition algorithm and the use of a set of computationally effective features, alongside the contrast of the recognition capabilities of the proposed approach using various classification models. The results show an enhanced recognition of gesture motion intent with the use of the decomposition method, despite the faint phantom motion signal from the amputee.

Keywords: Prosthesis; Pattern recognition; Signal processing; Bionics; EMG

1. Introduction

The hand and upper-limb of an individual serve as means towards carrying out daily activities in addition to navigating through the environment, thus the loss of a portion of the upper-limb has steadily become a widespread problem in society due to reasons that span vascular diseases, accidents and also trauma from conflict-related incidents\(^1\). The loss of an upper-limb has implications concerning the independence and degree of autonomy of the individual, in addition to deeper rooted problems such as distress related to the phantom sensations, and an imbalance of the human motor control system and pathway\(^2\). Therapeutic measures exist to help alleviate the distress associated with uncomfortable phantom discomforts whilst the brain...
initiates the neuropsychiatric adaptation effect to reorganise the cortex and ultimately account for the loss of one of the body’s functional parts [2]. The bionic prosthesis which represents a functional replacement to the loss of an upper-limb has also shown signs of being able to compensate for the loss both from a functional and a neurological perspective, where the augmentation of an amputee with a bionic body part also serves to form a ‘cybernetic human motor control loop’, which helps to cater unpleasant phantom sensations, as described by Nsugbe et al. [2]. Thus, it can be said that the bionic prosthesis serves as a holistic augmentation where it caters to both physical and neurological voids in a human being.

There exist various kinds of bionic upper-limb prostheses depending on the level of upper-limb loss of the amputee, where aside from transcarpal amputees, the classes of amputees with a substantial amount of limb loss include transradial (below elbow amputation), transhumeral (above elbow amputation) and shoulder disarticulation, each with a unique challenge and kind of prosthesis [3,4]. Amongst the various kinds of bionic prostheses, common challenges include ergonomics, cost, and primarily the ability to sense and decode phantom motions [5].

The majority of the prosthesis literature appears to be centred around the research, development and design of transradial prostheses, which has served as an appealing platform for investigating prosthesis control approaches with various kinds of physiological sensing modules that later get carried over and further iterated with regards to the extreme amputee cohorts (transhumeral and shoulder disarticulation), who pose a greater challenge in terms of the sensing of the phantom motions and the decoding of the acquired stochastic physiological signal [6–14]. Of the available control architectures for bionic prosthesis, the pattern recognition architecture appears to be the favoured control scheme due to its overall intuitiveness, where the appropriate functionality of the architecture hinges heavily upon an appropriate solution to the gesture motion recognition problem [4,5]. As a result, this has understandably formed the core emphasis of the transradial prosthesis literature, where researchers have employed different sensing approaches and associated computational models towards effectively solving the recognition problem with a mixture of healthy and amputee subjects [15–21]. Also, to a lesser extent, work has also been done in the estimation of the contraction force associated with a gesture motion as part of means towards robustifying the control architecture to force variations, and also with a potential towards an advanced prosthesis control interface capable of closely mimicking the human upper-limb, where gestures can be performed with a varying level of contraction force and intensity [14,22,23].

Amongst the various research done in this area, key work has been done by Al-Timemy et al. who investigated the recognition of phantom gesture motions amidst a varying force contraction in a group of amputees [24]. As part of Al-Timemy’s work, prosthesis control architectures were trained and validated with a strategic combination of various contraction forces using various feature sets, where they obtained a variety of results and showcased quantitatively the effects of varied contraction forces on the transradial prosthesis control interface [24].

The work done in this study builds on the results obtained by Al-Timemy et al., where the open-source data was used to perform a pilot study on a multiphase advanced prosthesis control interface capable of first recognising an input gesture motion (irrespective of the contraction force), and subsequently followed by an estimation of the contraction intensity with which to perform the desired gesture motion by. In the first instance, this paper focuses on the gesture recognition amidst a varying contraction force intensity. An image and flow diagram of the proposed advanced control architecture can be seen in Figure 1.
Figure 1: Image of the proposed multistage prosthesis control architecture.

This pilot work was done with the use of a transradial subject who had undergone traumatic amputation and had been amputated for the longest time relative to the other amputee subjects. The subject utilised a cosmetic prosthesis, from which it can be inferred that the phantom motions would be expected to produce weak contractions due to the onset of neuroplasticity and the reorganisation of the motor cortex, when a certain limb pathway ceases to be used [25]. As part of means towards dealing with the problem of a potentially faint phantom motion and a likely low signal-to-noise ratio, we employed a novel and computationally effective signal decomposition method previously utilised in various aspects of medical signal processing and capable of deconvolving a signal as part of a search process to find the optimal region in the signal which minimises uncertainty and maximises the prediction capability [26–31]. Specifically speaking, the contributions of this paper are as follows:

- The use of a novel decomposition alongside a set of selected features to investigate the gesture recognition of a faint phantom motion
- A contrast between various classification models spanning decision tree (DT), linear support vector machine (L-SVM) and quadratic support vector machine (Q-SVM)

2. Methods

2.1. Data Collection

The data from the amputee subject was collected from an army base in Iraq by Al-Timemy et al., where the amputee was missing a segment of his left upper-limb, an image of which can be seen in Figure 2 [24]. Details regarding the amputee can be seen in Table 1 [24].

Figure 2: Image of a transradial amputee missing a segment of his left upper-limb. [24]

Table 1: Details on the amputee subject.

<table>
<thead>
<tr>
<th>Age</th>
<th>Gender</th>
<th>Cause of Amputation</th>
<th>Stump Length</th>
<th>Stump Circumference</th>
<th>Time Since Amputation</th>
<th>Prosthesis Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>30</td>
<td>Male</td>
<td>Traumatic</td>
<td>20 cm</td>
<td>23.5 cm</td>
<td>28 years</td>
<td>Cosmetic</td>
</tr>
</tbody>
</table>

Ethical approval was approved for the study by the local authority, while a written consent was provided by the participant prior to the commencement of the study [24]. The skin of the subject was cleansed with a combination of alcohol and an abrasive skin preparation gel prior to the placement of electrodes from an electromyography (EMG) sensing system where 12 electrodes were placed on the amputee’s stump, which were ordered in two rows around the arm of the amputee with the elbow joint serving as the reference point, and where the European recommendations were followed [24]. A custom EMG acquisition system was used in the acquisition of the neuromuscular signals with a sampling rate of 2000 Hz [24].

The gesture motions conducted as part of the data collection included fine digit flexions as well as gripping motions, all of which can be said to be important hand movements in day-to-day activi-
ties [24]. The list of these gestures is as follows: Spherical Grip, Index Flexion, Hook Grip, Thumb Flexion, Tripod Grip, and Fine Pinch; while a visual interactive system allowed for an overview of how much contraction force they were exerting [24]. The aim of the force variation was to mimic a real life operating scenario where the contraction force used to produce gesture motion varies [24]. For each gesture motion, three distinct contraction levels were produced, namely, Low, Medium and High, where each force and associated gesture motion was held for 8–12 seconds across a varied number of iterations. The acquisition protocol utilised for each contraction force level is as follows:

**Low Force:** the low contraction force represents a contraction intensity which is below a determined baseline level of contraction; this was held for a timeframe spanning 8–12 seconds for multiple iterations.

**Moderate Force:** the moderate force in this case represents the baseline and nominal force that an amputee produces when they make a gesture motion, and can be expected to be higher than the prior Low force contraction.

**High Force:** this force level was deemed to be the maximum contraction force that the amputee was comfortable making, but was less than the maximum voluntary contraction force which could not be sustained by the amputee due to a lack of use, and typically resulted in pain and discomfort in the amputee.

It was noted by Al-Timemy et al. that the process of producing the Low and High contraction intensity was deemed to be challenging as these represented muscular contraction levels that had become scarcely used by the amputees [24]. Thus, these contractions caused discomfort, an increased cognitive loading and occasional tremor in the amputee during the data acquisition and can also be thought to contribute to uncertainty within the acquired physiological signals [24].

In this work, a candidate trial spanning 10 seconds was utilised for each gesture alongside the accompanying force contraction level. This was further split using a disjointed windowing scheme spanning 250 ms to produce samples used for the decomposition exercise and subsequent feature extraction.

### 2.2. Signal Decomposition

As mentioned, the concept of signal decompositions involves the methodical and algorithmic separation of a candidate signal in an attempt to find a region of interest which minimises the overall uncertainty in the signal and allows for an accurate prediction. Applications of signal decomposition methods span areas such as the analysis of physiological signal in medicine, econometrics, and seismic explorations, to name a few [32].

The Linear Series Decomposition Learner (LSDL) represents a metaheuristics decomposition method devised by Nsugbe et al. as part of source separation exercises involving the estimation of particle size distributions in powder mixtures using high frequency acoustic emission signals [26–31]. The decomposition has been seen to surpass the wavelet decomposition and is also characterised by being computationally efficient due to working in the time domain and the utilisation of a linear threshold as the basis function for the decomposition. Aside from its original inception case study, the LSDL has seen broad applications in various areas of clinical medicine and physiological signal processing spanning rehabilitation, pregnancy medicine, and adolescent schizophrenia [30,31,33].

The threshold parameters for the LSDL can be seen in Table 2, and a comprehensive list of the heuristics used in tuning the series of linear thresholds for the decomposition can be seen in Nsugbe et al [30]. The various parameters used for the implementation of a threshold given absolute representation of a signal $|S_n|$, a tree-like decomposition flow of the process, can be seen in Figure 3.

**Table 2:** Threshold parameters for the LSDL.
Where $T_{\text{Upper}, n}$ and $T_{\text{Lower}, n}$ are the thresholds for the upper and lower amplitude regions of the signal.

From a mathematical perspective, the LSDL decomposition series can be expressed as follows:

$$|S_n| = e^{-\sum \left( \frac{1}{2} (c_{\text{mean}})(S_{n-1})^2 + \sum (c_{\text{mean}})(S_{n-1})(S_n) \right) + \sum (c_{\text{mean}})(S_{n-1})(S_{n+1})}$$

2.3. Feature Extraction

The features used in this study comprise an ensemble of select statistical features capable of modelling these kinds of signals and used in various capacities in previous studies where, as mentioned, they comprise of mainly statistical features which are computationally efficient [35]. The list of features is as follows: mean absolute value (MAV), 4th order autoregression (AR), simple square integral (SSI), enhanced mean absolute value (EMAV), log detector (LD), Wilson amplitude (WAMP), variance (VAR), root mean square (RMS), kurtosis (Kurt), and maximum cepstrum coefficient (Ceps)[38].

2.4. Classification Methods

- Decision Tree (DT): the DT represents a form of non-parametric classification model which sorts data into various classes using sorting rules garnered from the input feature samples, and is regarded as a white box modelling method due to its high interpretability appeal [36].

- Support Vector Machine (SVM): this classification model is based around the higher dimensional projection of the data while using a sub-portion of the dataset known as support vectors, and in an iterative manner the instillation of class boundaries are formed and subsequently followed by a downward projection to a lower dimensional space while retaining the formed class boundaries in a feat regarded as the ‘kernel trick’[30].

Two variations of the SVM were used in this work, namely, the linear SVM denoted as L-SVM and the quadratic SVM denoted as Q-SVM, both of which represent iterative classifiers with a low model complexity.

Both classifiers were validated using the hold-out approach where 80 % of the data was used to train the classifiers, with the remained 20 % serving as the test-set to obtain the final classification accuracy of the trained classifier.

3. Results

The results of the LSDL decomposition can be seen in Table 3, where it can be noted that the optimal threshold region is seen to be in the first iteration of the lower amplitude region of the signal. All subsequent signals were decomposed using the parameters corresponding to the optimal threshold.
region as per the framework of the LSDL.

**Table 3: LSDL decomposition results.**

<table>
<thead>
<tr>
<th>Iteration</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Upper</td>
<td>2.2432</td>
<td>n/a*</td>
<td>n/a*</td>
<td>n/a*</td>
</tr>
<tr>
<td>Lower</td>
<td>2.8000</td>
<td>2.7803</td>
<td>2.7536</td>
<td>2.7167</td>
</tr>
</tbody>
</table>

* Where n/a indicates decompositions that could not be carried out due to a constrained number of samples

### 3.1. Pattern Recognition Design and Test with Combined Forces

The results for the training and testing of the classifier of the pattern recognition system using all three force levels can be seen in Table 4, where it can also be intuitively noted that this exercise represents a realistic reflection of how the prosthesis would be used in a real-life scenario, where the control and actuation can be expected to be initiated with a variety of forcing levels. From Table 4, it can be seen that a generally low accuracy is obtained when the raw signal is used, with the highest accuracy seen to be the Q-SVM with 56.8 %. A sharp increment is noted for when the signal is pre-processed with the LSDL, where there is an improvement in the range of around 15–25 % depending on the classifier, of which the best performance was seen to be the DT, and closely followed by the Q-SVM. From this it can also be inferred that a non-linear sorting and decision boundary is needed for a good classification performance for this problem.

As mentioned, the training of the pattern recognition system with a range of forces as described also serves as a form of robustness to the system, as this can also help cater to the common problem of gesture misclassification due to muscular fatigue, and therein adds an extra layer of technical capability to the pattern recognition system.

**Table 4: Raw signal.**

<table>
<thead>
<tr>
<th></th>
<th>DT (%)</th>
<th>L-SVM (%)</th>
<th>Q-SVM (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classification Accuracy (Raw Signal)</td>
<td>46.1</td>
<td>45.8</td>
<td>56.8</td>
</tr>
<tr>
<td>Classification Accuracy (LSDL Decomposed Signal)</td>
<td>72.2</td>
<td>62.6</td>
<td>71.8</td>
</tr>
</tbody>
</table>

A principal component analysis (PCA) plot, which qualitatively compares the cluster separation extent of the raw data and the LSDL decomposed signal, can be seen in Figure 4.

Figure 4: PCA plot of the LSDL pre-processed decomposition (top) and the raw signal (bottom), with 99 % variability explained.

Figure 4 shows the clusters for three gestures comprising a variation of the force levels, where it can be seen that with the LSDL decomposition there is a greater degree of cluster separability, while in the case of the raw signal there exists a substantial amount of cluster overlap, which makes the classification between data clusters challenging.

### 3.2. Pattern Recognition Design and Test with Individual Forces

#### 3.2.1. Low Force

The result for the training and testing with the Low force can be seen in Table 5, from which an
improvement can firstly be noted (especially in the case of the raw signal) when compared with Table 4, where all the forces were used in the training and test process. This is thought to be due to a simplification of the problem by the use of a homogeneous/single force level in the training and testing phase, which simplifies the recognition problem in a sense. The LSDL has also undergone a slight improvement in the classification accuracy due to the simplification of the problem, where the Q-SVM is seen to produce the highest classification accuracy at 77.5\%.

The LSDL has also undergone a slight increment in the classification accuracy due to the simplification of the problem, where the Q-SVM is seen to produce the highest classification accuracy at 77.5\%.

Table 5: Results of the pattern recognition exercises with a Low force.

<table>
<thead>
<tr>
<th></th>
<th>DT (%)</th>
<th>L-SVM (%)</th>
<th>Q-SVM (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classification Accuracy (Raw Signal)</td>
<td>63.4</td>
<td>53.9</td>
<td>67.6</td>
</tr>
<tr>
<td>Classification Accuracy (LSDL Decomposed Signal)</td>
<td>74.4</td>
<td>76.4</td>
<td>77.5</td>
</tr>
</tbody>
</table>

3.2.2. Medium Force

The results for the Medium force exercise can be seen in Table 6, where it can be seen that the result of the raw signal appears to have gone down slightly. This can be assumed to be due to the fact that the probability distribution gets more dispersed as contraction force increases and therein adds further variability to the overall sample, thus potentially contributing to the observed reduction in the classification accuracy. This notion does not extend to the LSDL results as a signal decomposition is carried out as a pre-processing strategy, where it can be seen that the results appear to have undergone a slight increase. The reason for this is thought to be linked to the collection of the data, where it is noted that the Medium contraction force represents the nominal baseline contraction that the amputees produce on a frequent basis, whereas anything outside of this (i.e., Low and High contraction forces) caused discomfort and pain in the amputees, and therein a source of uncertainty. The Q-SVM produced the best classification accuracy with 81.7\%.

Table 6: Results of the pattern recognition exercises with a Medium force.

<table>
<thead>
<tr>
<th></th>
<th>DT (%)</th>
<th>L-SVM (%)</th>
<th>Q-SVM (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classification Accuracy (Raw Signal)</td>
<td>55.2</td>
<td>46.1</td>
<td>62.7</td>
</tr>
<tr>
<td>Classification Accuracy (LSDL Decomposed Signal)</td>
<td>79.0</td>
<td>77.9</td>
<td>81.7</td>
</tr>
</tbody>
</table>

3.2.3. High Force

The results for the final force category can be seen in Table 7 for the High force, where for the raw signal the trend continues, with the classification results dipping yet again in this case due to the probability distribution function further expanding with an increased level of contraction force and therein being more variable as a whole, and contributing to the perceived drop in the accuracy. In the case of the LSDL, although a considerable increment is observed in the classification accuracy when compared with the raw signal, a slight accuracy degradation has been seen to occur, which is due to the aforementioned effects of the discomfort and pain encountered by the amputee in the production of contraction forces separate from the nominal Medium level contraction force. Nevertheless, the
LSDL continues to be able to recognise gestures with a high accuracy, which in this case the best accuracy is seen to be in the order of 77% for the Q-SVM.

**Table 7: Results of the pattern recognition exercises with a High force.**

<table>
<thead>
<tr>
<th></th>
<th>DT (%)</th>
<th>L-SVM (%)</th>
<th>Q-SVM (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classification</td>
<td>49.9</td>
<td>47.7</td>
<td>60.3</td>
</tr>
<tr>
<td>Accuracy (Raw Signal)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>73.5</td>
<td>70.9</td>
<td>77.0</td>
</tr>
<tr>
<td>Classification</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accuracy (LSDL</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decomposed Signal)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.3. Computational Time Evaluation

Computation times were conducted based on a single patient’s dataset, where the metrics were computed with a laptop of Intel(R) Core™ i5-3210M CPU @ 2.50 GHz, with a 64-bit operating system and 6GB RAM.

The computation time for the LSDL can be seen to be around 5 ms, which is a distinctively low figure given its structure and is what has come to be expected from the approach. The feature extraction time produced for a windowed segment of 250 ms can be seen to be under 80 ms, which is a relatively efficient figure given the number of features, with the computation time benefits from the feature group being largely statistical. The selection time was evaluated for the three different models, where it can be seen that the DT produced the lowest amount of time for the computation metric, while the different SVM models produced higher computation time due to the iterative nature of the model architecture. Given the relatively low computation time of the DT, which can be likened to that of the discriminant analysis – the favoured classification architecture in this area due to its optimal computation time – it can be recommended that the DT be further explored in subsequent studies on the basis of the results obtained in this research.

**Table 8: Computational time results.**

<table>
<thead>
<tr>
<th>LSDL Decomposition (for 1 channel) (ms)</th>
<th>Feature Extraction (for a 250ms segment) (ms)</th>
<th>Selection Time-DT (for the prediction of a single instance) (ms)</th>
<th>Selection Time-L-SVM (for the prediction of a single instance) (ms)</th>
<th>Selection Time-Q-SVM (for the prediction of a single instance) (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.46 ± 0.48</td>
<td>79.0 ± 40.0</td>
<td>19.0 ± 15.3</td>
<td>59.4 ± 8.5*</td>
<td>50.2 ± 7.0*</td>
</tr>
</tbody>
</table>

* Although the selection time for both the L-SVM and Q-SVM are of somewhat similar timeframes, it can be noted that the training time of the Q-SVM is considerably longer than that of the L-SVM (although not displayed here)

4. Conclusion

The bionic prosthesis represents the most functional alternative to the loss of an upper-limb, and due to its functionality, it also helps in alleviating the symptoms of an unpleasant phantom sensation, where the pattern recognition control scheme is seen to be the favoured control architecture due to its overall intuitiveness. The majority of the research in the literature that explores the design of the pattern recognition control interface is primarily centred around gesture recognition of phantom motions, and to a lesser extent, the prediction of an associated extent of the contraction force.

As part of an ongoing study, the design of an advanced pattern recognition prosthesis control system capable of recognising an input gesture from a phantom motion, followed by a successive esti-
mation of the contraction force was investigated, where in the first instance the gesture recognition portion of the control system was studied amidst a varying level of contraction forces from the phantom motion. The work utilised an open-source transradial amputee database, where data from an amputee whose phantom is believed to produce a faint contraction due to the time since amputation was utilised for the exercises carried out in this study. This involved the use of a computationally efficient decomposition algorithm whose results were contrasted against the raw signal with a set of computationally efficient statistical features alongside a contrast of the recognition performance with three different classifiers. Various case studies were conducted within this study where the first involved the recognition of gestures amidst varying forces and also a single force only.

The results from the varying force exercise, which represent a realistic depiction of what is likely to be encountered in the daily life of an amputee, showed an improved accuracy when the decomposition algorithm was used as a pre-processing mechanism. In the case of the single force, the Medium contraction force produced the best accuracy amidst the continued use of the decomposition algorithm, with a lower accuracy obtained for the other forces due to the lack of familiarity alongside the associated discomfort accompanied in the production of those forces. For the various classification models used in this study, the DT produced the best combination for computation time and classification accuracy, and thus is recommended to be utilised in the area of gesture recognition and prosthesis control.

Subsequent work in this area will now involve further exercise on the prediction of an accompanying contraction force using classification algorithms, and also the exploration of regressions as a means towards a continuous estimation of the contraction force.

Acknowledgement

The author would like to thank Brian Kerr from Kerr Editing for proofreading this manuscript.

References


[16] Barry D, Leonard J, Gitter A, Ball RD. Acoustic myography as a control signal for an externally powered prosthesis. Archives of physical medicine and rehabilitation. 1986;


[34] Nsugbe E, Al-Timemy AH, Samuel OW. Intelligence Combiner: A Combination of Deep Learning and Handcrafted Features for an Adolescent Psychosis Prediction using EEG Signals. In peer review.