

AN INTUITIVE MUSCLE-COMPUTER INTERFACE USING ULTRASOUND SENSING AND MARKOVIAN STATE TRANSITIONS

Ananya S. Dhawan¹, Jana Košecká¹, Huzefa Rangwala¹, Siddhartha Sikdar²

¹Department of Computer Science, George Mason University, Fairfax, USA

²Department of Bioengineering, George Mason University, Fairfax, USA

ABSTRACT

In recent work regarding gesture recognition and muscle computer interfaces, ultrasound-based sensing strategies have been demonstrated as a viable alternative to the pervasive surface electromyography (sEMG) modality. However, in order to facilitate switching between available gestures, both sEMG and ultrasound-based strategies have traditionally relied on unintuitive control mechanisms. The most common among these are: requiring the users to return to rest as an intermediary state between motions; mode switching through co-contraction or other ad-hoc user input; and switching based on muscle activations that are functionally unrelated to the desired motion. The unintuitive nature of such control has historically led to increased user frustration, and is often cited a major reason for device abandonment in the prosthetic control setting. In this work, we propose using an approach inspired by Hidden Markov Models (HMMs) with a novel continuous gesture recognition mechanism, for ultrasound-based sensing. We empirically calculate the average classification accuracy of our novel method during non-transitional periods to be 99%. We then demonstrate that including predictions made during transition periods reduces this value to 69%. Finally, by encoding the temporal dependency of the system within a Hidden Markov Model framework, we show that we can reduce the error caused by the instability of predictions during transitions, measured as the normalized Levenshtein distance from the true ordering, by approximately 98.8%.

Index Terms— Ultrasound, Muscle Computer Interface, Gesture Recognition, Hidden Markov Models

1. INTRODUCTION

With the recent and ongoing advancements in domains such as robotics, medicine, prosthetics, and general muscle computer interfaces [1-3], there is a clear need for robust and intuitive control strategies that are driven by non-invasive sensing of volitional motor intent of a human user. Surface Electromyography (sEMG) has found widespread use in these domains, and as a result a lot of research efforts have been devoted to building intuitive control around sEMG as a sensing modality [4-7]. However, low signal-to-noise ratio, inadequate specificity with regards to deep-seated muscle, sensor crosstalk and the absence of robust proportional control are all well-known shortcomings of this technology that limit functionality [8-12]. On the other hand, ultrasound-based sensing strategies have been demonstrated to enable a more natural motion paradigm, as well as a more intuitive control interface [13-17].

However, both the ultrasound and sEMG sensing strategies lack intuitive and directly responsive, continuous control. Although attempts have been made to overcome this problem for sEMG sensors [18-19], intuitive switching has yet to garner attention in ultrasound sensing paradigms. For a practical gesture recognition interface, a user must be able to transition seamlessly between motions without having to increase the amount of effort or cognitive load required to operate a device every time use of different hand gesture is desired. To this end, developing classification strategies that are robust to motion transitions and can make accurate classifications continuously, are key to the success of any muscle computer interface.

2. RELATED WORK

Sikdar et al. [13], Shi et al. [14], and Castellini et al. [15], have all demonstrated that ultrasound imaging can be used to classify individual finger flexion in able-bodied individuals with high accuracy. Akhlaghi et al. have shown that ultrasound based sensing strategies can be used to robustly classify significantly more complex hand motions, such as pronated grasps, in a computationally inexpensive manner with average offline classification accuracy of 91% and real-time average classification accuracy of 92% for 15 different motions of varying complexity [16]. Furthermore, the technology for ultrasound imaging has been undergoing significant miniaturization. Hettiarachchi et al. have utilized the images captured from eight custom built transducers, arranged into two wearable bands, to classify muscle activity in both able-bodied and amputee participants [17]. These advances make ultrasound a suitable alternative to sEMG as a wearable sensor for gesture recognition.

Of the work done in the gesture recognition domain the two most pertinent to the efforts of this paper are the works of Akhlaghi et al. on ultrasound based recognition [16], and the work done by Chan et al. on using HMMs for continuous prediction using sEMG [18]. The work described herein, builds directly upon the contributions of these two papers.

Akhlaghi et al. utilize a nearest neighbor classifier and demonstrate that the simplest of classifiers could yield very high discriminability among classes. However, they utilize a feature representation that depends on the motions beginning at rest which forces users to return to rest in order to switch between motions. To remedy this, we utilize a new feature representation so that the dependency on rest is removed. We demonstrate that the new feature representation offers high discriminability using the same classification model while allowing for continuous motion prediction and intuitive switching.

Chan et al. utilize an HMM approach that encodes a higher probability of staying at a motion rather than transitioning, and an equal probability of transitioning to all possible classes. We adapt this approach due to its ability to garner high classification accuracies in their work. However, we modify their probabilistic model in order to better integrate it with the nearest neighbor classifier used by Akhlaghi et al and apply it to the ultrasound domain.

3. EXPERIMENTAL SETUP

The following sections give an overview of our approach. We will first describe our classification methodology and the evaluation metrics used to determine its merit; we then explain how we incorporate our classification into a HMM; and finally, we describe our experimental setup.

3.1. Classification Methodology

In order to perform continuous classification, we utilize a 1-nearest-neighbor classifier: Using Pearson's correlation, we compute the similarity between each incoming frame in an unknown sequence against all of the items in a predefined set of labeled images; each unknown frame is then assigned the label of the item to which it has the highest similarity. Pearson's correlation is utilized solely because the bounded distances it returns lend themselves well to an incorporation into our HMM approach discussed in the following section. Any bounded distance metric should suffice for this approach.

We build the set of labeled images, the training set, by asking the user to perform a pre-selected set of motions in a sequence: the user starts at rest, then performs each motion alternated with rest, and finally ends at rest. Fig. 1 provides a visual representation of the motions. From this sequence, our algorithm extracts the frames that correspond to the end of each motion where the forearm muscles are at maximal contraction. This is done by identifying regions in a dynamic ultrasound sequence where the Pearson's correlation to a rest state starts to plateau. For each motion in the sequence, the extracted frames corresponding to the end of the motion are averaged into a single representative frame and added to the training set. This feature representation removes the dependency on rest since they

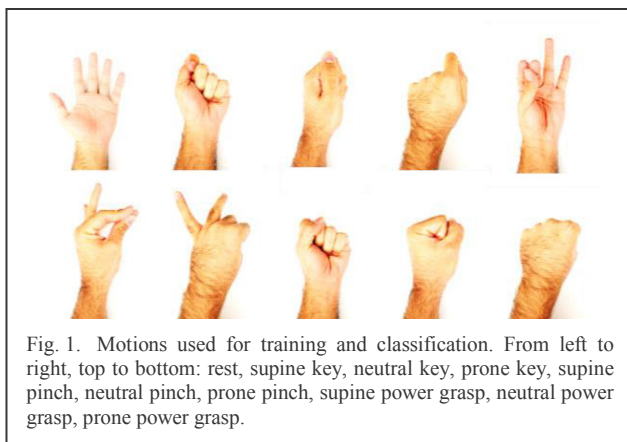


Fig. 1. Motions used for training and classification. From left to right, top to bottom: rest, supine key, neutral key, prone key, supine pinch, neutral pinch, prone pinch, supine power grasp, neutral power grasp, prone power grasp.

contain no temporal information and provide a mechanism for framewise-prediction and continuous classification. The frames are averaged to mitigate the effects of any small changes in probe positioning that may occur and the small transformations (translations, rotations) that may occur in the ultrasound images as a result. We repeat the image acquisition and training process until the desired number of representative frames per motion is reached.

3.2. Evaluation Metrics

In this paper, we report the framewise classification accuracy of our approach to give an intuitive measure of success. It is important to note however, that due to the continuous nature of muscle deformations, it is impossible to define the exact discrete frame where one motion switches to another. Furthermore, since the muscle deformations during transition periods are an amalgamation of the two motions they connect, and thus not exclusively like either motion, it is very difficult to determine the motion class to which a given transition frame belongs. Therefore, in order to provide an appropriate evaluation of our classifier to the best of our ability, we report the classification accuracies both while including the transition frames labeled as the motion state that follows them, and while removing them from consideration entirely. In fact, the exact temporal bounds of different motions are somewhat inconsequential; It is more important to determine whether a given gesture recognition system can correctly predict the relative ordering of motions with stability so that it may propagate movement to the end device –be it a physical or a virtual system– with the correct motions in a consistent and continuous manner.

Therefore, in order to supplement framewise accuracy, a metric that assumes the exact positions of the temporal bounds, we compute a second metric of evaluation to quantify the stability and correctness of motion predictions; We compute the squared Levenshtein distance [20] between a predicted ordering of motions, and the true ordering. The Levenshtein distance is a measure of the number of operations (replacing, deleting, or adding items) required to change one ordering to another. Therefore, if two orderings, predicted and true, are unlike one another, the number of operations required to change one ordering into the other will be high, thus making the Levenshtein distance between them large; conversely if two orderings are exactly the same, the number of operations required to change one to the other will be zero thus making the Levenshtein distance between them be zero as well. We report the Levenshtein distance for our HMM implementations after normalization such that the squared Levenshtein distance for the motion ordering predicted without the HMM is considered 100% erroneous, the ground truth ordering is considered 0% erroneous, and everything else is scaled proportionally in between. This allows us to intuitively quantify the extent to which the error is reduced by varying the parameters of our HMM implementation. Henceforth, we shall refer to this measure as the prediction error.

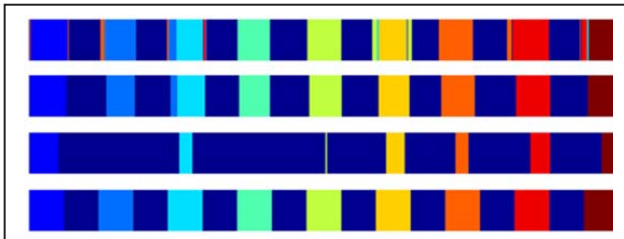


Fig 2. This figure demonstrates the noisiness of predictions given varying transition costs (top to bottom: 0, 2.5, 10) in comparison to the ground truth (bottom). The colors represent different motions and the horizontal axis represents time. For a HMM model with transition cost = 0, motion predictions are noisy around transitions; as transition cost increases the noise in the transitions is smoothed out; if the transition cost is too high, the model very rarely transitions to a new motion.

3.3. Hidden Markov Model Framework

A complete tutorial of Hidden Markov Models can be found at [21]; here, we only discuss Markov Models in terms of their adaptation for this work. Traditionally, Hidden Markov Models compute the probabilistic likelihood of being in a particular hidden-state at the current timestep when given: the current emissions of some variables that are probabilistically related to the hidden states, the probabilistic likelihood of transitioning from one state to another, and information about the hidden-state at the previous timestep. In our model, we consider a non-probabilistic, cost-based adaptation.

The emissions at a given timestep are directly computed from the correlation values between the unknown frame at the current timestep and the training set. We average the correlation values by class label and subtract them from 1 so that high similarity values (correlation values close to 1) become low cost (close to 0) and low similarity values (correlation values close to 0) become high cost (close to 1). The likelihood of transitioning in our model is also encoded non-probabilistically as a constant cost for transitioning to a new motion class and a cost of 0 for predicting the same motion class. The predicted motion class at a given timestep, t , can then be defined as: the class with the minimum corresponding cost at time t where cost C for a class i at time t is defined by eq. 1.

$$C_t^i = C_{t-1}^i + \min_j (E_{t-1}^j + T_{ji} + E_t^i) \quad (1)$$

Where C_{t-1}^i is the cost corresponding to class i , at time $t-1$; E_{t-1}^j is the emission corresponding to class j at time $t-1$; T_{ji} is the cost of transition from class j to class i and is equal to 0 when $i = j$ and some predefined constant otherwise; finally, E_t^i is the emission corresponding to class i at time t . The cost for all classes at time $t=0$ is 0. It is important to note that after the cost C_t^i is computed for all classes, that the costs are normalized according to eq. 2.

$$C_t^i = (C_t^i - \min_i C_t^i) / (\max_i C_t^i - \min_i C_t^i) \quad (2)$$

This ensures that the lowest computed cost of predicting a class is always 0 and the highest computed

cost is always 1, thus ensuring that a strong bias is not developed in favor of a given class solely because it has been predicted continuously for a long time.

3.4. Experiment

Three able-bodied participants were recruited for this study. All experiments were approved by the institutional review board of our university. A custom designed cuff held the ultrasound probe firmly against the participants' forearms to limit probe movement and ensure consistency in data collection. The probe was placed laterally on the anterior portion of the forearm at approximately 60% of the forearm length (closer to the elbow). Each participant was asked to perform nine motions alternated with rest (for a total of 10 classes). Five sequences were collected from each participant and 5-fold validation was performed for each participant such that four sequences were used for training (4 examples per class x 10 classes = 40 training examples) and the remaining sequence was used for testing. The sequences varied in length between 800 to 1100 frames. Classification accuracy is reported for each fold, for each participant. The prediction error for HMMs is computed at all possible combinations of transition cost between 0 and 10 at steps of size 0.1. A visual representation of the effects of the different cost settings on the noisiness of predictions is provided (Fig. 2). The cost yielding the lowest average prediction error across all participants and all trials is reported along with the prediction errors themselves (Fig. 3).

4. RESULTS AND DISCUSSION

The best transition cost was found to be 2.2. The average framewise classification accuracy if the transition periods are not considered, was found to be 98.8%. If the transition periods are considered, the average framewise classification accuracy was found to be 69.4%. At a transition cost of 2.2, the average prediction error was lowered to 1.2%.

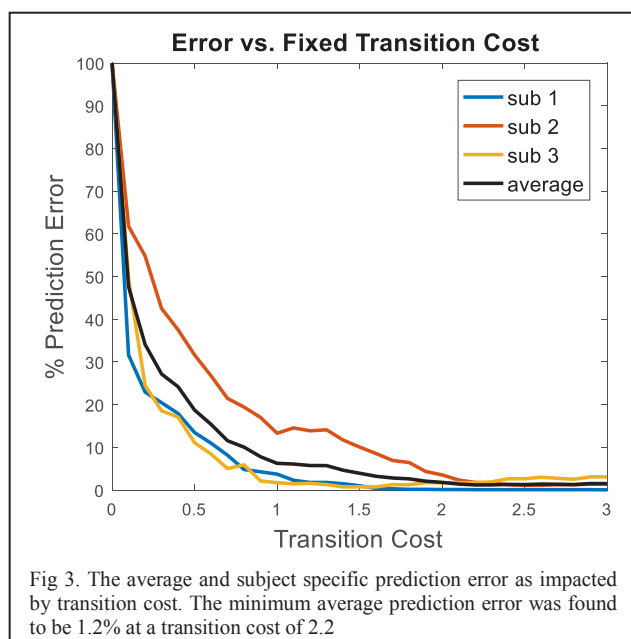
TABLE I. CLASSIFICATION ACCURACY WITH AND WITHOUT TRANSITIONS

	Subject 1		Subject 2		Subject 3	
	With	Without	With	Without	With	Without
Trial 1	100.0%	70.8%	100.0%	71.8%	100.0%	72.7%
Trial 2	100.0%	70.3%	100.0%	67.9%	96.8%	63.1%
Trial 3	98.6%	69.4%	95.4%	66.1%	100.0%	70.6%
Trial 4	100.0%	68.3%	91.8%	64.7%	100.0%	69.5%
Trial 5	100.0%	75.5%	100.0%	70.0%	100.0%	70.6%
Average	99.7%	70.9%	97.4%	68.1%	99.4%	69.3%

TABLE II. PREDICTION ERROR AT OPTIMAL TRANSITION COST

	Subject 1	Subject 2	Subject 3
	With	Without	With
Trial 1	0.2%	2.4%	6.6%
Trial 2	0.0%	1.9%	1.0%
Trial 3	0.3%	1.0%	1.1%
Trial 4	0.0%	0.5%	0.0%
Trial 5	0.0%	3.2%	0.0%
Average	0.1%	1.8%	1.7%

Our study shows that using a cost-based adaptation of HMMs, the noisy predictions of this classifier during transition periods can be easily overcome (Table 2, Fig 2



and Fig 3). Our predictions were made in an online fashion, i.e. we did not change past predictions once they had been made, or use information from future frames to make current predictions. However, we did not perform testing in real-time in this study nor did we address the problem of changes in probe positioning that may result from the user doffing the device and donning it on another day. This means that the important question of prediction latency caused by the HMM bias against transitioning is still unevaluated, as is the ability of the user to adjust to the response of the device in real-time. Furthermore, the effect of utilizing one day's data to make prediction on another day where the sensor has been removed and replaced is also absent in this work. We plan to continue evaluating these factors and improve upon this work by testing it in real-time settings with donning and doffing on a larger number of able-bodied participants as well as amputees.

5. CONCLUSIONS

In this work, we propose a novel algorithm that enables continuous prediction of the volitional motor intent of a user with high fidelity using ultrasound-based sensing of muscle activity. Our study demonstrates the feasibility of a continuous ultrasound based gesture recognition system that facilitates intuitive switching with low prediction errors.

6. ACKNOWLEDGEMENT

This work was supported by NSF grant #1329829

7. REFERENCES

- [1] M. Lech and B. Kostek, "Gesture-based computer control system applied to the interactive whiteboard," in 2010 2nd International Conference on Information Technology (ICIT), 2010, pp.75–78.
- [2] J. H. Tan, C. Chao, M. Zawaidh, A. C. Roberts, and T. B. Kinney, "Informatics in Radiology: Developing a Touchless User Interface for Intraoperative Image Control during Interventional Radiology Procedures," *RadioGraphics*, vol. 33, no. 2, pp.E61–E70, Mar. 2013.

- [3] M. Obaid, F. Kistler, M. Häring, R. Bühling, and E. André, "A Framework for User-Defined Body Gestures to Control a Humanoid Robot," *Int. J. Soc. Robot.*, vol. 6, no. 3, pp.383–396, Apr. 2014.
- [4] G. Li, A. E. Schultz, and T. A. Kuiken, "Quantifying pattern recognition-based myoelectric control of multifunctional transradial prostheses," *IEEE Trans. Neural Syst. Rehabil. Eng. Publ. IEEE Eng. Med. Biol. Soc.*, vol. 18, no. 2, pp.185–192, Apr. 2010.
- [5] F. V. G. Tenore*, A. Ramos, A. Fahmy, S. Acharya, R. Etienne-Cummings, and N. V. Thakor, "Decoding of Individuated Finger Movements Using Surface Electromyography," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 5, pp.1427–1434, May 2009.
- [6] H. Daley, K. Englehart, L. Hargrove, and U. Kuruganti, "High density electromyography data of normally limbed and transradial amputee subjects for multifunction prosthetic control," *J. Electromyogr. Kinesiol. Off. J. Int. Soc. Electrophysiol. Kinesiol.*, vol. 22, no. 3, pp.478–484, Jun. 2012.
- [7] Scheme, Erik and Kevin Englehart. "Electromyogram Pattern Recognition for Control of Powered Upper-Limb Prostheses: State of the Art and Challenges for Clinical Use." *Journal of Rehabilitation Research & Development*, vol. 48, no. 6, Sept. 2011, pp.643-659. EBSCOhost
- [8] E. A. Clancy, E. L. Morin, and R. Merletti, "Sampling, noise-reduction and amplitude estimation issues in surface electromyography," *J. Electromyogr. Kinesiol. Off. J. Int. Soc. Electrophysiol. Kinesiol.*, vol. 12, no. 1, pp.1–16, Feb. 2002.
- [9] H. S. Ryait, A. S. Arora, and R. Agarwal, "Study of issues in the development of surface EMG controlled human hand," *J. Mater. Sci. Mater. Med.*, vol. 20 Suppl 1, pp.S107–114, Dec. 2009.
- [10] E. Biddiss, D. Beaton, and T. Chau, "Consumer design priorities for upper limb prosthetics," *Disabil. Rehabil. Assist. Technol.*, vol. 2, no. 6, pp.346–357, Nov. 2007.
- [11] "Prosthesis rejection in acquired major upper-limb amputees: a population-based survey."
- [12] P. J. Kyberd and W. Hill, "Survey of upper limb prosthesis users in Sweden, the United Kingdom and Canada," *Prosthet. Orthot. Int.*, vol. 35, no. 2, pp.234–241, Jun. 2011.
- [13] 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 Trans. Neural Syst. Rehabil. Eng.*, vol. 22, no. 1, pp.69–76, Jan. 2014.
- [14] J. Shi, J.-Y. Guo, S.-X. Hu, and Y.-P. Zheng, "Recognition of finger flexion motion from ultrasound image: a feasibility study," *Ultrasound Med. Biol.*, vol. 38, no. 10, pp.1695–1704, Oct. 2012.
- [15] C. Castellini, G. Passig, and E. Zarka, "Using ultrasound images of the forearm to predict finger positions," *IEEE Trans. Neural Syst. Rehabil. Eng. Publ. IEEE Eng. Med. Biol. Soc.*, vol. 20, no. 6, pp.788–797, Nov. 2012.
- [16] N. Akhlaghi, C. Baker, M. Lahlou, H. Zafar, K. Murthy, H. Rangwala, J. Kosecka, W. Joiner, J. Pancrazio, and S. Sikdar, "Real-time Classification of Hand Motions using Ultrasound Imaging of Forearm Muscles," *IEEE Trans. Biomed. Eng.*, vol. PP, no. 99, pp.1–1, 2015.
- [17] N. Hettiarachchi, Z. Ju, and H. Liu, "A New Wearable Ultrasound Muscle Activity Sensing System for Dexterous Prosthetic Control," in 2015 IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2015, pp.1415–1420.
- [18] D. C. Chan and K. B. Englehart, "Continuous myoelectric control for powered prostheses using hidden Markov models," in *IEEE Transactions on Biomedical Engineering*, vol. 52, no. 1, pp.121–124, Jan. 2005.
- [19] K. Englehart and B. Hudgins, "A robust, real-time control scheme for multifunction myoelectric control," in *IEEE Transactions on Biomedical Engineering*, vol. 50, no. 7, pp.848–854, July 2003.
- [20] Wagner, Robert A. The String-to-String Correction Problem. *Journal of the ACM*. (01/1974) , 21 (1), p. 168 - 173.
- [21] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," in *Proceedings of the IEEE*, vol. 77, no. 2, pp.257–286, Feb 1989