

EEG-TO-F0: ESTABLISHING ARTIFICIAL NEURO-MUSCULAR PATHWAY FOR KINEMATICS-BASED FUNDAMENTAL FREQUENCY CONTROL

Himanshu Goyal^{*1,3}, Pramit Saha^{†2}, Bryan Gick^{‡3}, and Sidney Fels^{§2}

¹Department of Computer Science, University of British Columbia

²Department of Electrical and Computer Engineering, University of British Columbia

³Department of Linguistics, University of British Columbia

1 Introduction

Speech-related Brain Computer Interfaces (BCI) are primarily targeted at finding alternative vocal communication pathways for people with speaking disabilities. However, most of the works in this field are centered around recognition of words from imagined speech, given a particular, small vocabulary [1–3]. Lesser efforts have been invested towards investigating voice synthesis from the information decoded from EEG signals [4, 5]. In this work, as an extension of the previous work [6], we aim at controlling the fundamental frequency of voice based on information derived from brain signals for an elbow movement task, through a biomechanical simulation toolkit ArtiSynth [6, 7]. To the best of our knowledge, this is the first attempt targeting such a brain-to-vocal frequency mapping incorporating a biomechanical control pathway.

2 Overview of proposed technique

The fundamental frequency (F0) of human voice is generally controlled by changing the vocal fold parameters (including tension, length and mass), which in turn is manipulated by the muscle exciters, activated by neural synergies. In order to begin investigating the neuromuscular to F0 control pathway, we simulate a simple biomechanical arm prototype (instead of an artificial vocal tract) that tends to control F0 of an artificial sound synthesiser based on the elbow movements. The intended arm movements are decoded from the EEG signal inputs (collected simultaneously with the kinematic hand data of the participant) through a combined machine learning and biomechanical modeling strategy.

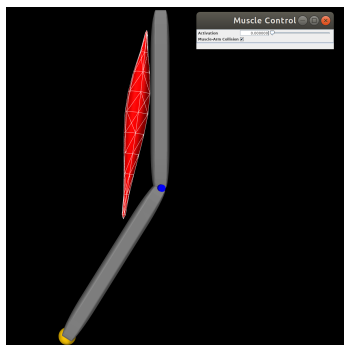


Figure 1: Simulated single-muscle arm model in ArtiSynth

3 Data collection and preparation

We record EEG data as shown in Fig. 2, corresponding to hand movements using a 10 electrode dry EEG headset with bluetooth connectivity, manufactured by Avertus. It has a sampling frequency of 1000 Hz per electrode at 20 bits/sample. We take data from the electrodes FP1, FP2, F7, F8, T3, T4, T5, T6, O1 and O2 with FCz as reference and FPz as the ground electrode. We particularly choose this headset because of the ease of use and portability of the equipment, which makes our strategy practically implementable. We collect 500 EEG data samples from single participant: 250 each for the hand going up and down, both with the eyes closed and the head fixed. We do not perform any pre-processing on the data and split the data into train:test as 70:30. We also simultaneously collect the angular displacement values or the elbow angles correspondingly with the EEG.

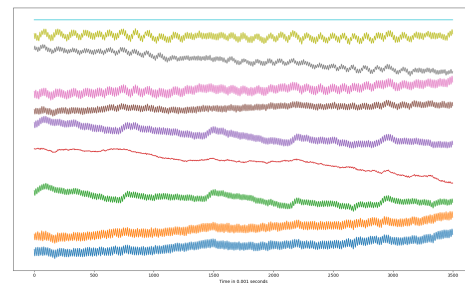


Figure 2: Sample of 10 channel EEG data collected (0.001 s)

4 Proposed methodology

In order to achieve this EEG to F0 mapping, we first attempt to map the EEG signals corresponding to the elbow movements to the discretized muscle activation of the simulated one-muscle arm model (in the biomechanical toolkit ArtiSynth as shown in Fig. 1) that would create same elbow angular displacement. The temporal resolution of the muscle activations in ArtiSynth is .01 second. Hence, we set a time-window of 0.01 second for EEG signal as well, thereby, each EEG sample (to be mapped to muscle activation) becomes a 10 x 10 matrix . Each matrix is fed to a random forest based classifier with min sample leaf = 1 and min sample split = 2, number of estimators = 10 and random state = 42. This is a 10-class classification task, with the output classes being {0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 and 1} which represent the discretized muscle activations between 0 and 1. Using inverse biomechanical modeling in ArtiSynth,

*h.goyal@alumni.ubc.ca

†pramit@ece.ubc.ca

‡gick@mail.ubc.ca

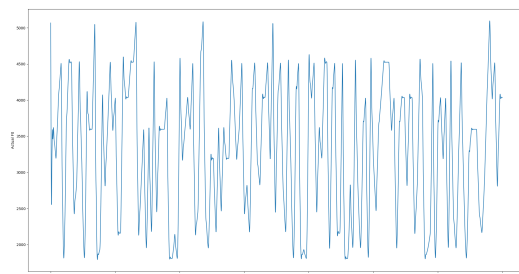
§ssfels@ece.ubc.ca

we calculate the required muscle activation corresponding to targeted angular movement and optimise the loss function between the intended muscle activation (from EEG signal) and simulated muscle activation (from ArtiSynth) to achieve the best possible mapping from EEG to the muscle activation.

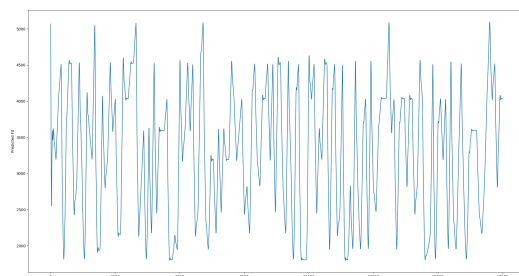
This muscle activation is loaded into the one-muscle biomechanical model of the hand in the ArtiSynth. The resultant angular movement is recorded and the elbow joint angle is then linearly mapped to F0 between 1500 Hz and 5150 Hz. We use the F0 value mapped from the actual kinematic hand data as the ground truth and compare the F0 estimated from brain signal to evaluate the performance of our method.

5 Results

In this section, we present our step-wise performance for the test data. In the first step, *i.e.*, mapping EEG to muscle activations, the overall mean accuracy (the percentage of times, the predicted value matches the actual value) achieved is 85%, with a root mean square error (rmse) of 0.04. After the next step, *i.e.* mapping muscle activations to angles, the cumulative mean accuracy is 68.8% with rmse of 3.4 degrees. The final rmse of the F0 estimation (ranging from 1.5 KHz to 5.15 KHz) is 102.7 Hz. For qualitative comparison of the final predicted F0 with actual F0, a randomly chosen sample is shown in Fig. 3.



(a) Actual F0 values



(b) Predicted F0 values

Figure 3: Comparison of actual and predicted (from EEG) fundamental frequency values

6 Conclusion

We report a novel successful mapping scheme from EEG to fundamental frequency, ranging from 1.5 KHz to 5.15

KHz, with a negligible rmse of 102.7 Hz. This is an initial prototype which proved to be satisfactory in testing the idea of EEG-to-muscle activation-to-kinematics-to-acoustics mapping. Instead of using the complex biomechanical space of the vocal folds, as an initial work in this direction, we tested the concept using simple elbow rotation and single-muscle arm model to simulate the movement, where upwards (or downwards) elbow motion represented increase (or decrease) in frequency. Further work will be directed towards incorporating the ideas using multiple muscle biomechanical arm model, extending it to a vocal fold model as well as replacing the random forest-based classifier with neural network models to investigate any increase in performance accuracy.

Acknowledgments

This work was funded by the Natural Sciences and Engineering Research Council (NSERC) of Canada and Canadian Institutes for Health Research (CIHR).

References

- [1] Primit Saha, Muhammad Abdul-Mageed, and Sidney Fels. Speak your mind! towards imagined speech recognition with hierarchical deep learning. *arXiv preprint arXiv:1904.05746*, 2019.
- [2] Primit Saha, Sidney Fels, and Muhammad Abdul-Mageed. Deep learning the eeg manifold for phonological categorization from active thoughts. In *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2762–2766. IEEE, 2019.
- [3] Primit Saha and Sidney Fels. Hierarchical deep feature learning for decoding imagined speech from eeg. *arXiv preprint arXiv:1904.04352*, 2019.
- [4] Gopala K Anumanchipalli, Josh Chartier, and Edward F Chang. Speech synthesis from neural decoding of spoken sentences. *Nature*, 568(7753):493, 2019.
- [5] Lindy Comstock, Ariel Tankus, Michelle Tran, Nader Pouratian, Itzhak Fried, and William Speier. Developing a real-time translator from neural signals to text: An articulatory phonetics approach. *Proceedings of the Society for Computation in Linguistics*, 2(1):322–325, 2019.
- [6] Primit Saha, Debasish Ray Mohapatra, SV Praneeth, and Sidney Fels. Sound-stream ii: Towards real-time gesture-controlled articulatory sound synthesis. *Canadian Acoustics*, 46(4):58–59, 2019.
- [7] John E Lloyd, Ian Stavness, and Sidney Fels. Artisynt: A fast interactive biomechanical modeling toolkit combining multi-body and finite element simulation. In *Soft Tissue Biomechanical Modeling for Computer Assisted Surgery*, pages 355–394. Springer, 2012.