

MACHINE LEARNING FOR ESTIMATION OF ACTIVATION PATTERNS IN COMPUTATIONAL MODELS OF THE TONGUE

Arnold David Gomez (1), Amod Jog (1), Maureen L. Stone (2), Jerry L. Prince (1)

(1) Electrical and Computer Engineering Dept.
Johns Hopkins University
Baltimore, MD, USA

(2) Department of Neural and Pain Sciences
University of Maryland School of Dentistry
Baltimore, MD, USA

INTRODUCTION

Tissue motion determines many vital processes such as blood pressurization in the heart, locomotion with the limbs, and speech generation with the aid of the tongue. For this reason, the relationship between muscular activation, force generation, and motion is central to biomechanics and biomedicine. The tongue offers a unique example of the complexity of this relationship, because its muscles can activate independently to produce the majority of deformations without direct reliance on joints or other rigid support structures [1]. Despite its complexity, our society has a need for investigating tongue biomechanics given the increasing rates of oral cancer, and the profound detrimental effect that it can inflict [2-3]. An association between activation patterns and different activities can guide the development of surgical tumor removal techniques aimed at reducing functional impact. Thus, a robust method for measuring, or estimating, muscular activation across the tongue's volume would be beneficial.

While non-invasive motion estimation has been successful in different organs, it is currently impossible to perform direct experimental measurements of activation patterns [4-5]. Instead, these can be inferred by optimizing the (activation) input of a computational model so that the simulated output matches measured deformation. This approach, known as inverse modeling, has been demonstrated in the heart and the tongue using finite-element (FE) models [4-5]. Implementation of numerical optimization is challenging because it must include an effective objective function, be numerically stable, and be robust to local minima. For this reason, different strategies make use of regularization. For instance, a common approach consists of grouping contractile elements together and finding each group's activation magnitude [4-5]. However, such grouping, a form of regularization may not be optimal based on observations of the system. This research investigates the use of machine learning without such

regularization. Despite being digital in nature, computational models effectively approximate a continuum; thus, the number of possible configurations resulting from all activation patterns is extremely large. One of the main challenges is, then, designing an appropriate training dataset, which is the focus of this research. The goal of this study was to perform inverse estimation via machine learning with a limited training data set and to test the validity of the predictions.

METHODS

The experiments were carried out with an FE model because it can provide training data and test cases (with known activation patterns) outside the training dataset. The strategy (Fig. 1) uses a random forest algorithm to obtain activation predictions and compares them to a known benchmark [6].

The FE model used in this investigation included a representation of the tongue [7], which was constrained using rigid models of the hyoid bone and the jaw. The material was assumed to be isotropic

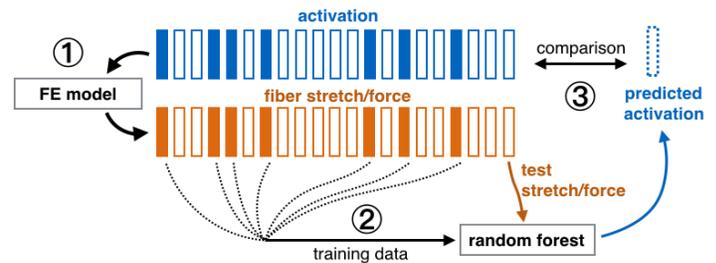


Figure 1: Test strategy. The FE model (1) was used to generate paired training data (2) for random forest regression, which generated predictions (3) to be compared to the known activations.

(neo-Hookean) with coefficients extracted from the literature [8]. To reduce the solution space, the geometry was grouped into the approximate location of 6 tongue muscles (Fig. 2). The training data set consisted of randomly sampling half of the permutations of muscular activation at 4 discrete levels of activation. The largest was 50% (of a maximum contractile stress of 0.35 Mpa [8]) in decrements of 12.5%. The total number of training simulations was 2,048 obtained over a 6-hr period using the FEBio Software suite [9].

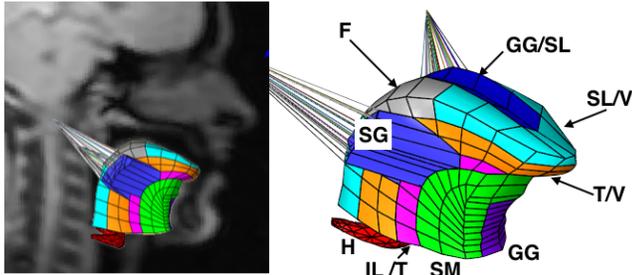


Figure 2: FE model of the tongue. Different colors represent regions shared by different muscles, which included: genioglossus (GG), superior longitudinal (SL), inferior longitudinal (IL), transverse (T), verticalis (V), and styloglossus (SG). The hyoid bone (H), mandible (not shown), and support material (SM) were also included. A flap region (F) was a test modification.

Estimation of activation was framed as a regression problem, but element interactions make it highly nonlinear. The random forest algorithm [6] has been used successfully as a nonlinear regression learning framework in many applications [10]. Given a test simulation k with known fiber strains λ_{ik} and fiber forces f_{ik} (i.e. the strain and stress tensors along the fiber directions, respectively), the trained random forest R_i predicted the activation \hat{a}_{ik} for a given element i ($i = 1, 2, \dots, n$). Borrowing machine learning notation, the set of activations in all elements is $\hat{\mathbf{a}}_k$. To train R_i , a feature vector \mathbf{b}_j was calculated from each training simulation j ($j = 1, 2, \dots, m$), where $\mathbf{b}_j = [\boldsymbol{\lambda}_j \ \mathbf{f}_j]$, such that $\boldsymbol{\lambda}_j \{\lambda_{1j}, \lambda_{1j}, \dots, \lambda_{nj}\}$ was concatenation of fiber strains at all elements for a training simulation j , and $\mathbf{f}_j \{f_{1j}, f_{1j}, \dots, f_{nj}\}$ was a concatenation of fiber forces at all n elements. The feature vector \mathbf{b}_j was paired with activation of a_{ij} to create a training dataset $\langle \mathbf{b}_j, a_{ij} \rangle$ for all m training simulations. A random forest R_i was trained on $\langle \mathbf{b}_j, a_{ij} \rangle$. The number of trees in R_i was chosen empirically to be 30. Each tree was trained as described in previous work [6]. The minimum number of samples to be maintained in the leaf of a tree was empirically determined to be one. Given the fiber strains and fiber forces observations from a test case k , a feature vector $\mathbf{b}_k = [\boldsymbol{\lambda}_k \ \mathbf{f}_k]$ was generated, and \hat{a}_{ik} resulted from applying the trained random forest to the new feature vector, $R_i(\mathbf{b}_k)$.

Test data was constructed using the same muscular groupings, but with random 100 magnitude permutations in the (continuous) range from 0–50%. An additional case used one of these permutations, but with a modification in the spatial grouping: the flap region (labeled F in Fig. 2) was given zero activation. The difference between the approximated activation and the benchmark (for each k) was quantified as a percentage from $\|\mathbf{a}_k - \hat{\mathbf{a}}_k\| / \|\mathbf{a}_k\|$.

RESULTS

Overall, the random forest algorithm provided accurate predictions of activations (pattern and magnitude) across all the magnitude test cases. The mean error across these cases was 10.3% with a standard deviation of 5.8%. Figure 3 (top row) shows at typical solution from these test cases—few differences can be seen visually.

Introducing the flap modification to the mesh (Figure 3, bottom row), affected the prediction by lowering the activation magnitude across the entire muscle region (in this case the SL/V grouping) instead of revealing the expected low-magnitude region. The error in this case was 91.1%.

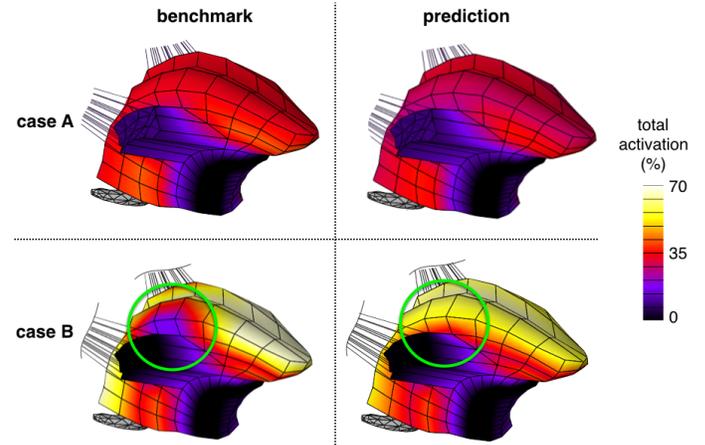


Figure 3: Differences between benchmark and predicted activation patterns. Case A (top row) corresponds to one of the tests simulations differing in magnitude only. Case B (bottom row) resulted from modifying the magnitude and spatial configuration. Both of these were outside the training dataset.

DISCUSSION

Despite using a small training dataset, the algorithm was successful in predicting activation patterns when they differed in magnitude. This can be attributed to the relative smoothness in variation between configurations of varying magnitude. The same was not true when introducing a change in spatial configuration, where the prediction fell back to a type of “best fit” within the spatial groups already in the training dataset. This is a direct consequence of the training approach, which implicitly implemented spatial regularization within the algorithm.

These findings are encouraging because a large number of magnitude and pattern variations can be captured by the regression as is. However, the results also highlight the need for training sets that encompass both magnitude, and spatial variations. Therefore, increasing the size of the training database and improving sampling of the solution space are required to achieve better results.

ACKNOWLEDGEMENTS

The authors would like to acknowledge grants R01DC014717 and 2R01NS055951 from the National Institutes of Health in the United States.

REFERENCES

- [1] Takemoto H, *J Speech Lang Hear Res.* 44(1):95-107, 2001.
- [2] Edge SB, et al., *Ann Surg Oncol.* 17(6):1471-4, 2010.
- [3] Patel SG, et al., *CA Cancer J Clin.* 55(4):242-58, 2005.
- [4] Sun K, et al. *J Biomech Eng.* 131(11):111001, 2009.
- [5] Stavness, et a. *J Biomech.* 15;45(16):2841-8, 2012.
- [6] Breiman L, *Machine Learning.* 24(2):123–140, 1996
- [7] Fujita S, et al., *Oral Sci. Int.* 4(2):97-109, 2007 (via artisynth.org)
- [8] Teran JE, et al., *IEEE Trans. Vis. and Comp. Graphics.* 11(3):317-328, 2005.
- [9] Maas, G et al., *J Biomech.* 34:859-871, 2001.
- [10] Breiman L, et al. *Classification and Regression Trees*, 1984.