ARTICLE IN PRESS

Journal of Biomechanics xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Journal of Biomechanics

journal homepage: www.elsevier.com/locate/jbiomech
www.JBiomech.com

Model-based inverse estimation for active contraction stresses of tongue muscles using 3D surface shape in speech production

Narihiko Koike [a], Satoshi Ii [a,*], Tsukasa Yoshinaga [a], Kazunori Nozaki [b], Shigeo Wada [a]

[a] Graduate School of Engineering Science, Osaka University, 1-3 Machikaneyama, Toyonaka, Osaka 560-8531, Japan
[b] Division of Dental Informatics, Osaka University Dental Hospital, 1-8 Yamadaoka, Suita, Osaka 565-0871, Japan

ARTICLE INFO

Article history:
Accepted 4 September 2017
Available online xxxx

Keywords:
Tongue muscle
Active contraction stress
Inverse analysis
Variational data assimilation
Speech production
Nonlinear deformation

ABSTRACT

This paper presents a novel inverse estimation approach for the active contraction stresses of tongue muscles during speech. The proposed method is based on variational data assimilation using a mechanical tongue model and 3D tongue surface shapes for speech production. The mechanical tongue model considers nonlinear hyperelasticity, finite deformation, actual geometry from computed tomography (CT) images, and anisotropic active contraction by muscle fibers, the orientations of which are ideally determined using anatomical drawings. The tongue deformation is obtained by solving a stationary force-equilibrium equation using a finite element method. An inverse problem is established to find the combination of muscle contraction stresses that minimizes the Euclidean distance of the tongue surfaces between the mechanical analysis and CT results of speech production, where a signed-distance function represents the tongue surface. Our approach is validated through an ideal numerical example and extended to the real-world case of two Japanese vowels, /ʉ/ and /ɯ/. The results capture the target shape completely and provide an excellent estimation of the active contraction stresses in the ideal case, and exhibit similar tendencies as in previous observations and simulations for the actual vowel cases. The present approach can reveal the relative relationship among the muscle contraction stresses in similar utterances with different tongue shapes, and enables the investigation of the coordination of tongue muscles during speech using only the deformed tongue shape obtained from medical images. This will enhance our understanding of speech motor control.

1. Introduction

The tongue is an important articulatory organ that adjusts the shape of the oral cavity during speech production by deforming its own shape according to the active contractions of tongue muscle fibers. The tongue muscle fibers are distributed across the tongue and are locally confined in each muscle, and their orientations are smooth within each muscle group. The deformation behavior and associated mechanism of the tongue in speech has been explored by some experimental studies using electropalatography (EPG) (McAuliffe et al., 2001; McLeod et al., 2006), ultrasound imaging (Stone and Lundberg, 1996; Lundberg and Stone, 1999), and tagged magnetic resonance imaging (MRI) (Stone et al., 2001; Xing et al., 2016). In addition, the electromyography (EMG) measurements have revealed multiple muscle activations in speech (Miyawaki et al., 1975; Baer et al., 1988; Niimi et al., 1994). However, these activations were only measured for extrinsic tongue muscles, because it is difficult to resolve the intrinsic tongue muscles including the complexity of the fiber configurations.

To fully understand the detailed mechanisms of the tongue deformation and associated active contractions of the tongue muscles in speech production, some computational analyses have been conducted using a continuum model with anisotropic fiber components. Fang et al. (2009) estimated the tongue muscle contractions during Japanese vowel pronunciation by solving an optimization problem using a physiological articulatory model and MRI observation data on the midsagittal plane. Buchaillard et al. (2009) examined the tongue muscle contractions used to produce French cardinal vowels by solving the direct deformation problem under a given muscle contraction parameter and performing acoustic analyses using the deformed shapes for the upper mandible and tongue. However, the articulation occurs as the result of 3D tongue deformation, and thus the 2D shape matching approach may be insufficient for evaluating the full combination of muscle contractions. Moreover, it is not easy to find a unique combination of muscle contractions using direct deformation analysis because of their high degree of freedom. Stavness et al. (2012b) proposed

* Corresponding author.
E-mail address: sii@me.es.osaka-u.ac.jp (S. Ii).

https://doi.org/10.1016/j.jbiomech.2017.09.008
0021-9290/© 2017 Published by Elsevier Ltd.

an inverse prediction technique for muscle activations based on a forward-dynamics tracking simulation using discrete nodes on the tip of a mechanical tongue model. However, it may be difficult to associate the corresponding feature points uniquely from observation images. Additionally, their approach requires a regularization term to ensure the system is well-posed, and this can affect the estimation accuracy.

The purpose of this study is to develop a novel inverse estimation approach for evaluating the active contraction stresses of tongue muscles during speech. The proposed method is based on variational data assimilation using a mechanical tongue model and the 3D tongue surface shapes observed in speech production. A key point of this approach is the use of a signed-distance function (SDF) of the 3D target tongue shape, as this ensures that the data assimilation does not require any feature points or regularization technique to stabilize the system. The proposed approach is validated through ideal numerical examples and extended to the real-world speech productions problem of two Japanese vowels, /ʉ/ and /ɯ/. We discuss the estimation accuracy of the present approach in the ideal example, and discuss the validity of the estimated shape and muscle contraction stress for the actual vowel cases through a comparison with the previous studies.

2. Methods

2.1. Three-dimensional tongue model

The vocal tract imaging protocol was approved by the ethics committee at the Faculty of Dentistry, Osaka University (H26-E39). Volumetric images of the vocal tract of a 42-year-old male Japanese subject in the rest position were measured with a 320-row Area Detector CT (Toshiba Medical Systems Corporation, Ibaraki, Japan), resulting in 320 slices of 512×512 pixels and $0.488 \times 0.488 \times 0.5$ mm voxels. A 3D tongue model was created by extracting the tongue geometry from the CT images using the image-processing software AMIRA v.5.5.0 (FEI VSG, ZIB, Germany), and a tetrahedral volumetric mesh of 293,841 cells was created by ANSYS ICEM CFD v.13.0 (ANSYS Inc, PA, US).

The muscle type was classified into nine categories, with five types of extrinsic tongue muscles (GGa: anterior part of the genioglossus, GGm: middle part of genioglossus, GGp: posterior part of genioglossus, HG: hyoglossus, Sty: styloglossus) and four types of intrinsic tongue muscles (IL: inferior longitudinalis, SL: superior longitudinalis, Trans: transversalis, Vert: verticalis). We assumed that the tongue muscle fibers only play a role in generating contraction stress in the direction of each fiber orientation, and that the magnitude of the contraction stress in each muscle is the same. At the discrete level, the fiber orientation was assigned at the centroid of each tetrahedral cell and the contraction stress was imposed in the direction of the fiber orientation. The reconstructed tongue domain was implicitly divided into regions for extrinsic and intrinsic tongue muscles, and the fiber orientations were ideally determined based on anatomical data (Miyawaki, 1974; Takemoto, 2001). Note that the Trans and Vert muscles occupy the same space in our ideal description; therefore, we randomly assigned the muscle type (Trans or Vert) in each tetrahedral cell over the domain. Fig. 1 shows the reconstructed geometry of the tongue in the rest state and each muscle fiber.

Note that determining the *exact* reference (or stress-free) state is not easy, and therefore we regarded the 3D tongue model at the rest position as the reference configuration.

2.2. Direct analysis for mechanical tongue model

Let us denote by $\Omega \subset \mathbb{R}^3$ the tongue domain in the reference configuration $\mathbf{X}$. The tongue deformation is obtained by solving the following boundary value problem for a nonlinear finite deformation:

$$\nabla_0 \cdot (\mathbf{S} \cdot \mathbf{F}^T) = 0, \quad \text{in } \Omega, \tag{1a}$$
$$\mathbf{u} = 0, \quad \text{on } \Gamma_D, \tag{1b}$$
$$\mathbf{N} \cdot (\mathbf{S} \cdot \mathbf{F}^T) = 0, \text{on } \Gamma_N, \tag{1c}$$

where $\nabla_0 = \partial/\partial\mathbf{X}$ is the gradient operator in the reference configuration, $\mathbf{S}$ is the second Piola–Kirchhoff stress tensor, $\mathbf{u}$ is the displacement vector, and $\mathbf{F}$ is the deformation gradient tensor, given as $\mathbf{F} = \mathbf{I} + \nabla_0\mathbf{u}^T$. We imposed a fixed displacement condition on the boundary Γ_D and a traction-free condition on the boundary Γ_N with the unit normal vector $\mathbf{N}$ in the reference configuration. The stress tensor $\mathbf{S}$ is composed of a passive part $\mathbf{S}_{\text{passive}}$ and an active part $\mathbf{S}_{\text{active}}$ for the tongue tissue and muscle fibers, namely,

$$\mathbf{S} = \mathbf{S}_{\text{passive}} + \mathbf{S}_{\text{active}}. \tag{2}$$

The tissue was assumed to be an isotropic hyperelastic body expressed by a compressible Mooney–Rivlin model that considers material nonlinearity up to the second order. Thus, the passive stress is given by

$$\begin{aligned} \mathbf{S}_{\text{passive}}(\mathbf{u}) &= \frac{\partial W}{\partial \mathbf{E}}, \\ W &= c_1(\widetilde{I}_C - 3) + c_2(\widetilde{I}_C - 3)^2 + \frac{K}{2}(J-1)^2, \end{aligned} \tag{3}$$

where W is the strain energy, which is a function of the first reduced invariant $\widetilde{I}_C$ of the right Cauchy–Green deformation tensor, $\mathbf{C} = \mathbf{F}^T \cdot \mathbf{F}$, and the volumetric Jacobian J; $\mathbf{E} = (\mathbf{C} - \mathbf{I})/2$ is the Green–Lagrange strain tensor. The shear moduli c_1, c_2 were set to $c_1 = 1037$ Pa, $c_2 = 486$ Pa (Buchaillard et al., 2009), and the bulk modulus K was empirically set to $K = 2000$ Pa. The active stress is defined as

$$\mathbf{S}_{\text{active}}(F) = F\mathbf{e}_0 \otimes \mathbf{e}_0, \tag{4}$$

where F and $\mathbf{e}_0$ are the muscle contraction stress and unit basis vector of the fiber orientation in the reference configuration, respectively. We introduced nine groups of muscles, and thus nine corresponding contraction stresses are defined in terms of F as driving components of the deformation. We set Γ_D on the bottom surface of the tongue model and Γ_N on the other surfaces (Fig. 2(a)).

The nodal displacement-based Galerkin finite element formulation was employed to discretize Eq. (1) using the tetrahedral linear element (P_1 element), resulting in the weak form:

$$\int_\Omega \delta\mathbf{E} : (\mathbf{S}_{\text{passive}}(\mathbf{u}) + \mathbf{S}_{\text{active}}(F))d\mathbf{X} = 0, \quad \text{in } \Omega, \tag{5a}$$
$$\mathbf{u} = 0, \quad \text{on } \Gamma_D, \tag{5b}$$

where $\delta\mathbf{E} = \delta(\mathbf{F}^T \cdot \mathbf{F})/2 = (\nabla_0\delta\mathbf{u} + \nabla_0\delta\mathbf{u}^T + \nabla_0\delta\mathbf{u} \cdot \nabla_0\mathbf{u}^T + \nabla_0\mathbf{u} \cdot \nabla_0\delta\mathbf{u}^T)/2$ consists of the virtual displacement vector $\delta\mathbf{u}$. By linearizing the nonlinear (5) for $\mathbf{u}$, we finally obtain the following linear equation:

$$\mathbf{L} \cdot \Delta\mathbf{u} = \mathbf{r} + \sum_{\alpha=1}^{9} \mathbf{s}^{(\alpha)} F^{(\alpha)}, \quad \text{in } \Omega, \tag{6}$$

where $\Delta\mathbf{u} = \mathbf{u}^{n+1} - \mathbf{u}^n$ is the displacement increment vector (n indicates the nonlinear iteration index), $\mathbf{L} = \mathbf{L}(\mathbf{u}^n)$ is the stiffness matrix, $\mathbf{r} = \mathbf{r}(\mathbf{u}^n)$ is the right-hand-side vector, and $\mathbf{s}^{(\alpha)}, F^{(\alpha)}$ are the coefficient vector and muscle contraction stress of each fiber $\alpha(= 1, 2, \ldots, 9)$. To relax the numerical instability caused by nonlinearity, the displacement is updated as

$$\mathbf{u}^{n+1} = (1 - \xi)\mathbf{u}^n + \xi(\mathbf{u}^n + \Delta\mathbf{u}) = \mathbf{u}^n + \xi\Delta\mathbf{u}, \quad \text{in } \Omega, \tag{7}$$

where $\xi \in [0, 1]$ is the relaxation parameter, which we set to $\xi = 0.1$.

Please cite this article in press as: Koike, N., et al. Model-based inverse estimation for active contraction stresses of tongue muscles using 3D surface shape in speech production. J. Biomech. (2017), https://doi.org/10.1016/j.jbiomech.2017.09.008