Observation and Simulation of Large-scale Deformation of Tongue

Jianwu DANG[1] Satoru FUJITA[1], Emi MURANO[2] and Maureen STONE[2]
[1]*Japan Advanced Institute of Science and Technology, Ishikawa, Japan*
[2]*Department of Biomedical Sciences*
University of Maryland Dental School, Baltimore, MD 21201

[jdang,fujita]@jaist.ac.jp, mstone@umaryland.edu

Abstract. *To develop a physiological articulatory model for speech production, mastication and swallowing, we proposed an analysis-by-synthesis (AbS) based estimation method for investigating contributions of the tongue muscles in both exterior movement and interior deformation using observations and model simulations. The validity of the method was confirmed by comparing the estimated muscle activation to known muscle activation via model simulation using a full three dimensional physiological articulatory model. Observations were conducted to measure several simple tasks of large-scale tongue deformations using tagged magnetic resonance imaging (tMRI), and the principal strain of the deformations were analyzed using HARP MRI analysis. The muscle activation was examined based on the principal strain pattern and anatomic knowledge. According to the observation, a muscle activation pattern was designed to generate a tongue movement with extreme protrusion-retraction using the model. Muscle activation in the movement was estimated by comparing the principal strain pattern and known muscle forces. The model simulation and tMRI observation showed consistent strain patterns with each other. The study showed that the muscle activation within tongue movement is able to be estimated from local deformations using the AbS method.*

1. Information

The tongue plays an important role in vital activities of human life such as mastication, swallowing and speech production. In those activities, skilled movements are realized by well-practiced muscular contraction, in which the human tongue has a complicated musculature that undergoes local internal deformations and then leads to changes in the shape of the tongue surface. Hence, understanding the movement of the tongue both on its surface and its interior is crucial in understanding speech control and the pathology of the tongue.

The surface shape of the tongue has been imaged using ultrasound, X-ray microbeam and electromagnetic articulographic systems (Hashi, *et al.* 1998; Okadome, *et al.* 2001; Stone, *et al.* 2003). The relation between tongue movement and muscle activation has been evaluated using experimental approaches such as electromyography (EMG) (Baer, *et al.* 1988). However, EMG observations succeeded only for a few large muscles such as the extrinsic tongue muscles. In contrast, tagged cine magnetic resonance imaging (tMRI) can observe both the exterior shape and interior strain of tongue movements. A number of researchers investigated the functions of the lingual muscles using tMRI (Epstein, *et al.* 2003; Stone, *et al.* 2001; Takano, *et al.* 2004). However, the accuracy of such investigations is questionable, since a number of muscles are co-activated even in forming a simple movement. Further, it is often difficult to determine accurately the mechanical load of a muscle because the load depends on the situation of many other muscles. On the other hand, a physiological model of human speech organs can provide an alternative means to understand characteristics of speech production.

In this study, we developed a full three-dimensional (3D) physiological simulator based on a partial 3D physiological articulatory model (Dang & Honda, 2004). A number of tongue movements are generated using the model by given muscle patterns, and the interior deformation is evaluated using the principal strains. An analysis-by-synthesis (AbS) based method is proposed

and applied to the model simulation. This method is evaluated by comparing model simulation and observation that is obtained from the tMRI data.

2. Physiological Model and Muscular Structure

The full 3D tongue model is derived from the partial 3D model (Dang & Honda, 2004), which was based on volumetric MRI data obtained from a male Japanese speaker. The model computationally employs a truss structure consisting of viscoelastic cylinders with both mass and volume, and thus achieves a "semi-continuum" model of the tongue tissue. The initial shape of the model adopts the tongue shape of the Japanese vowel [e]. The full 3D tongue model is shown in Figure 1 (a). The outlines of the tongue body were extracted from the volumetric MR images. Mesh segmentation of the tongue tissue roughly replicates the fiber orientation of the genioglossus muscle. The outline of the tongue body in the sagittal direction was divided into ten radial sections that fan out from the genioglossus' attachment on the jaw to the tongue surface. In the perpendicular direction, the tongue tissue was divided into six sections concentrically. A 3D mesh model was constructed by connecting the corresponding nodes in the adjacent sagittal planes. As a result, the width in the left-right direction of this model is narrower in the anterior portion and wider in the posterior portion, where the width of the tongue root is about 5cm. Based on the same MRI data set, the vocal tract wall and the jaw were constructed in 3D with a width of 2.8 cm in the left-right dimension.

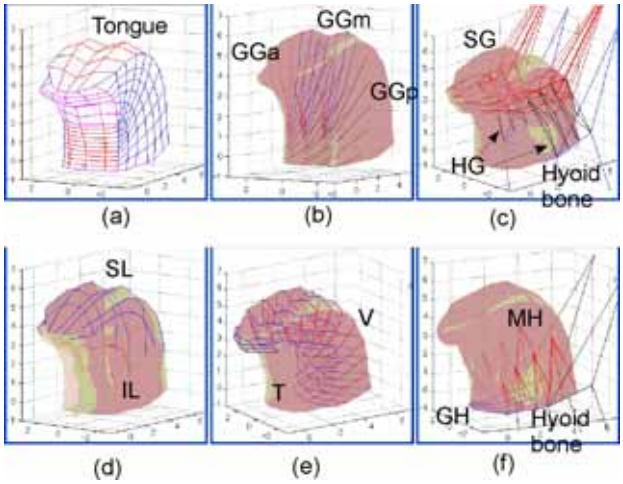

Figure 1. Configuration of the tongue (a) and the muscular structure of the model (b)-(e).

The muscular structure of the tongue model is shown in Figure 1 from (b) to (f), where the larger muscles were traced from the MR images obtained from the target speaker (Dang & Honda, 2001). The orientation of the tongue muscles was also examined with reference to the literature (Takemoto, 2001). Figure 1(b) shows the genioglossus (GG), which mainly runs midsagittally in the central part of the tongue and the same structure with thinner fibers are replicated in the bilateral layers next to the midsagittal layer. Since the fibers of the GG distributes in a triangular shape and different parts of the GG exert different effects on tongue deformation, it can be functionally separated into three segments: the anterior portion (GGa) indicated by the red lines, the middle portion (GGm) shown by the blue lines, and the posterior portion (GGp) denoted by the dark lines. To construct a realistic structure, the GG arrange in three layers while the muscle fibers are thinner in the bilateral layers than that in the midsagittal layer. Figure 1(c) shows the arrangements of the styloglossus (SG) and hyoglossus (HG) in the outer layers. The left and right bundles of the SG are attached to two bilateral fixed points respectively, which are outside the display region, while in the anterior part the fibers run from the outer layer to the inner layers towards the apex region. The HG is attached to the hyoid bone. Figure 1(d) illustrates the superior longitudinalis (SL) and the inferior longitudinalis (IL). Figure 1(e) shows the structure of the transversus (T) and verticalis (V), where the T runs in the left-right direction and V runs in the radial direction. Figure 1 (f) shows two tongue floor muscles of the geniohyoid (GH) and mylohyoid (MH). The top points of the mylohyoid bundles are attached to the medial surface of the mandibular body.

In normal speech, the tongue tends to move symmetrically while the tongue moves asymmetrically in mastication activities and maybe in disordered speech also. To realize those functions, all the muscles are designed symmetrically on the left and right sides, while they can be

controlled independently for the left side and right side. In this study, desired tongue movements are generated by activating the muscles and the generated movements are compared with observations of tMRI.

3. Force Field in Model Simulation

Since a number of muscles are co-activated even in forming a simple movement, it is often difficult to determine accurately the mechanical load of a muscle because the load depends on the situation of other muscles. In this study, we propose an analysis-by-synthesis (AbS) method to estimate muscle contributions by comparing model simulation and observation. For a preliminary test, we activate several muscles individually, and investigate the relation between the principal strain patterns and the muscle activation patterns.

3.1 Interpolation function for discrete structures

As shown in Fig. 1, the mesh structure of the model consists of discrete lattices. To make model simulation and tMRI observation compatible, the model lattices should be interpolated into a continuum form by means of shape functions. This section briefly describes the interpolation method used in this study.

To calculate the strain field for a quadrilateral, the four apexes are numbered as $(x_1, y_1; x_2, y_2; x_3, y_3; x_4, y_4)$ in counter-clockwise order. A common shape function shown in (1) is employed in this study.

$$f = a_1 + a_2 x + a_3 y + a_4 xy \tag{1}$$

At the apex *i*, the shape function becomes

$$a_1 + a_2 x_i + a_3 y_i + a_4 x_i y_i - f_i = 0 \tag{2}$$

Let $F = [f_1, f_2, f_3, f_4]^T, A = [a_1, a_2, a_3, a_4]^T, X = [1, x, y, xy]^T,$ coefficient A can be determined by a simultaneous equation, as shown in (3).

$$[X_1, X_2, X_3, X_4]^T A - F = 0 \tag{3}$$

As the result, A is expressed as a function of F. The interpolation $\Phi(x, y)$ for any point (x,y) within the quadrilateral can be obtained by an inner product of A and X, $\Phi(x, y) = A^T X$. Rearranging $\Phi(x, y)$ as a function of F, we get the following expression,

$$\Phi(x, y) = [N_1, N_2, N_3, N_4]F = \sum_{i=1}^{4} N_i f_i, \tag{4}$$

where $[N_1, N_2, N_3, N_4]$ is the displacement interpolation matrix. The strain-displacement matrix is a combination of the derivative of the displacement interpolation function N_i, as shown in (5).

$$D = \begin{bmatrix} \frac{\partial N_1}{\partial x} & 0 & \frac{\partial N_2}{\partial x} & 0 & \frac{\partial N_3}{\partial x} & 0 & \frac{\partial N_4}{\partial x} & 0 \\ 0 & \frac{\partial N_1}{\partial y} & 0 & \frac{\partial N_2}{\partial y} & 0 & \frac{\partial N_3}{\partial y} & 0 & \frac{\partial N_4}{\partial y} \\ \frac{\partial N_1}{\partial x} & \frac{\partial N_1}{\partial y} & \frac{\partial N_2}{\partial x} & \frac{\partial N_2}{\partial y} & \frac{\partial N_3}{\partial x} & \frac{\partial N_3}{\partial y} & \frac{\partial N_4}{\partial x} & \frac{\partial N_4}{\partial y} \end{bmatrix} \tag{5}$$

The relation between the strain at any point inside the quadrilateral and displacements of the four apexes can be calculated by (6).

$$\Sigma = DU \tag{6}$$

where $\Sigma = [\varepsilon_x, \varepsilon_y, \varepsilon_z]^T$ and $U = [u_1, v_1, u_2, v_2, u_3, v_3, u_4, v_4]^T$. ε_x and ε_y are the strains in x and y directions, and ε_{xy} is the shared strain. u_i and v_i are the x- and y-displacement at apex *i*. Stress can be expressed by strains using the following formula.

$$\begin{bmatrix}\sigma_x\\ \sigma_y\\ \sigma_{xy}\end{bmatrix}=\begin{bmatrix}\frac{E}{1-\nu^2} & \frac{E\nu}{1-\nu^2} & 0\\ \frac{E\nu}{1-\nu^2} & \frac{E}{1-\nu^2} & 0\\ 0 & 0 & \frac{E}{2(1+\nu)}\end{bmatrix}\begin{bmatrix}\varepsilon_x\\ \varepsilon_y\\ \varepsilon_{xy}\end{bmatrix}=S\begin{bmatrix}\varepsilon_x\\ \varepsilon_y\\ \varepsilon_{xy}\end{bmatrix} \quad (7)$$

Where σ_x and σ_y are the stresses in x and y directions, and σ_{xy} the shared stress. S is the strain-stress matrix. E and ν are the Young's modulus and Poisson's ratio respectively. In this study, the value is 3×10^5 Pa for E and 0.49 for ν.

The eigenvectors and eigenvalues of the strain-stress matrix S reflect the tissue deformations, which may correspond to the muscle activation. If a local compression occurs in a certain direction the corresponding eigenvalue would be negative, while if a tissue expands in a certain direction the corresponding eigenvalue would be positive. Therefore, we can determine directions of compression and expansion for a deformation by calculating the eigenvalue and eigenvector. The eigenvectors give the direction of compression/expansion, while the absolute value of the eigenvalue shows the amplitude of compression/expansion. If we suppose that the shared strain ε_{xy} equals zero, the compression and expansion are always located in orthogonal directions.

3.2 Estimation of Principal Strain in a Simple Situation

In the preceding section, we described the relation between the deformation (compression/expansion) and the principal strain that is estimated from the principle of the finite element method. Generally speaking, any deformation of soft tissue is caused by an internal/external force. However, there is no guarantee that all of local deformations are driven by an active muscle force since some passive deformation may be caused by the hydrostatic property of the tongue tissue.

To clarify the accuracy of the estimation, the relation of a tongue deformation and muscle activation is investigated using a model simulation, where muscle activation directly induces a local compression along its fiber orientation. For this purpose, several simple movements were generated by activating one or two muscles alone, and the relation of tissue deformation and muscle activation is investigated using the above method. Figure 2 shows the principal strains of the tongue deformation under different conditions. The blue lines indicate the compression and the red lines show the expansion. The upper panels illustrate the results on the midsagittal plane 80ms after 2N force is applied on GGa, GGm, and GGp respectively. One can see that the compression area and direction are consistent with the orientation of the activated muscles. In (a), the strong compression pattern reflects the GGa activation, while the fourth column from the apex shows smaller compression amplitude because GGa is thinner in that region. In (b), the compression pattern almost corresponds to the GGm orientation. When GGp is activated, the compression pattern appears in the muscle distribution region, while some unexpected compression is also seen in the area of the tongue tip in Fig 2(c). In the simulation, this unexpected compression was caused by a passive deformation not by muscle activation. For the passive compression, it was found that the amplitude of the compression was smaller than that of the expansion. This might be a criterion for us to judge whether the compression is passive or not.

The lower panels of Fig. 2 show the compression and expansion in the outer layer, in which the fibers of HG and SG are mainly located. When HG is activated, vertical compression vectors appear in the posterior portion of the tongue, which reflect the activation of HG. In Fig. 2(e), the compression vectors roughly illustrate the fiber orientation of the SG when it is activated, while some unexpected compression appears in the lingua-mandibular attachment. Fig. 2(f) shows the compression pattern of the upper layer when the two muscles, SL anterior and Verticalis middle, are activated simultaneously using 2N of force. One can see that in the anterior part compression is almost in the horizontal direction, consistent with the contraction of SL anterior. In the middle part of the tongue, the compression vectors are consistent with the contraction of the Verticalis middle. In the second and third layers from the attachment there is less compression because there are less active muscle fibers.

By comparing the estimation to the known muscle activation, one can see that the compression pattern shows the distribution and orientation of the muscle fibers if they are activated. This confirms that the proposed method is capable of estimating a muscle activation pattern from soft

tissue deformation. However, some unexpected compression is seen in the analysis of the principle strain. The unexpected compression is passive and is mainly caused by the incompressibility of the tongue's soft tissue. As clarified above, if a compression is induced by a local expansion, in general, the amplitude of the expansion is larger than the compression. With this fact we can roughly distinguish whether a compression is active or not.

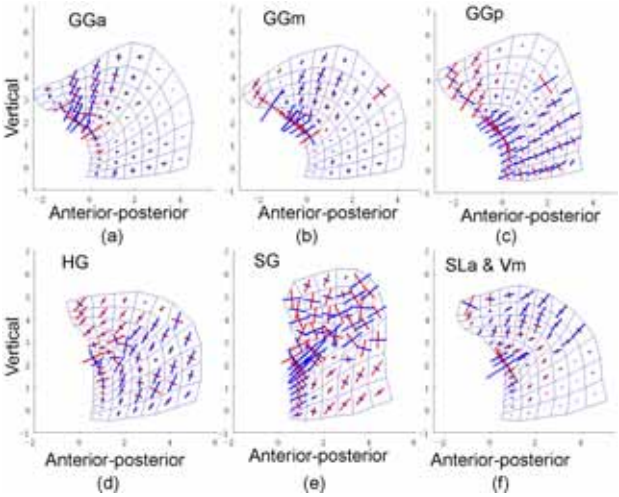

Figure 2. Principal strains from model simulations where the blue lines show the compression and the red lines the expansion. The pattern in the upper panels is generated when GGa, GGm, and GGp are activated by 2N force respectively. Panel (d) and (e) are the strain pattern when HG and SG are activated. (f) is the pattern when SL anterior and verticalis middle are activated simultaneously.

4. Observation of Large-scale Deformation of the Tongue

As mentioned in the earlier section, the tagged cine-MRI (tMRI) is widely used to investigate the interior deformation of the tongue. In this study, we attempt to estimate muscle contributions in a large-scale deformation using the tMRI data.

4.1. Task design and data recording

To investigate stress and strain of the tongue with a large-scale deformation, two tasks were designed for the tongue: protrusion-retraction movement and left-right bending. A male subject participated in this experiment. Precision of the repetition is critical in the tMRI data collection because a single image sequence is the summation of several repetitions. To obtain high quality tMRI data, the subject practiced the tasks a great number of times. He was instructed to move the anterior part of the tongue while keeping the posterior part as invariant as possible. A bilateral molar bite block was used to fix the jaw. The height of the block was about 1.6 cm. In the protrusion experiment, a wooden tongue depressor was set as an endpoint in front of the lips, which the subject touched in each repetition in the protrusion, while no depressor was used in the left-right experiment. The movement was made to a metronome composed of tMRI noise bursts at 2 sec intervals. The subject was trained to time his repetitions to the burst cycle.

For the tMRI recording, the subject repeated each task eight times per slice, with the bite block. Among the eight repetitions, the even numbered repetitions were used to construct the slice, and the odd numbered repetitions were 'set-up' measures. The tMRI data, collected at 18 Hz, yielded 25 time-phases per utterance. In the sagittal plane, nine side-by-side slices were recorded with a spatial resolution of $1.56\times 1.56 \times 5$ mm^3, totally 225 frames were recorded for each task. In the axial and coronal directions the spatial resolution was $2.19 \times 2.19 \times 7$ mm^3. The tMRI sequences were recorded in the sagittal and axial planes for the protrusion-retraction task, and in the axial and coronal planes for the left-right task. The tMRI data were collected using a procedure, called MICSR (Magnitude Image C-SPAMM Reconstruction Method).

The tMRI of the tongue protrusion-retraction-protrusion task is shown Figure 3. Figure 3(a) is the first frame of the midsagittal plane, where the tongue performs an extreme protrusion. Fig. 3(b) shows the parasagittal plane 1cm apart from the midsagittal planes, where the tongue retracted back. Panel (c) shows the second tongue protrusion in the same parasagittal plane as that of the

panel (b), where the different plane is used for investigating the muscles running in the different planes.

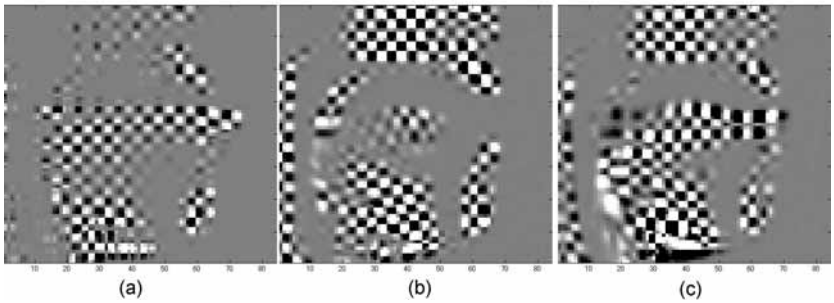

Figure 3. Tagged MR image of a protrusion-retraction-protrusion movement, (a) Frame 1 of the midsagittal plane where tongue is in an extreme protrusion, (b) Frame 12 of a parasagittal plane the tongue in a retraction position, and (c) Frame 25 of the same parasagittal plane where the tongue protruded again.

4.2 Analysis of Principal Strain

Motion of the tongue was measured from tagged MR images using Harmonic Phase (HARP)-MRI, which is an image processing method that allows for fast automated processing of tagged MR images. The HARP-MRI tongue analysis method was developed by Pathasarathy, et al. (2006) for estimating the muscle activation from the interior deformation of the tongue.

Figure 4 shows the principal strains of the three frames that correspond to the frames shown in Fig. 3. As seen in Fig. 4(a), there is very little compression at Frame 1 in the midsagittal plane, because the tags were laid down when the tongue was maximally protruded, only 16 ms earlier, and it is still quite protruded. In those 16 ms, the tip has expanded and the balance of the tongue is starting to compress. Since the tip was pressed against the tongue depressor, the tip expansion is probably due to the release of pressure.

In Fig. 4(b), Frame 12 shows the tongue is maximally retracted in a parasagittal plane 1cm apart from the midsagittal plane. The upper portion of the tongue shows horizontal compression (blue) consistent with SG contraction. The expansion (red) between the tongue and velum appears to be artifact, due to air. In the base of the tongue there is vertical compression. This may be a passive response to the retraction, or HG may be acting with SG to pull the tongue straight back. Other compressions are seen in the tongue base and may act to stabilize the tongue. The bottom layer is the jaw muscles. The jaw muscles are not expected to be active in tongue retraction, as they lower the jaw, elevate the hyoid, or push up the tongue. None of those things happen in tongue retraction, so the compression is probably passive; the volume preserving tongue moves downward as it retracts.

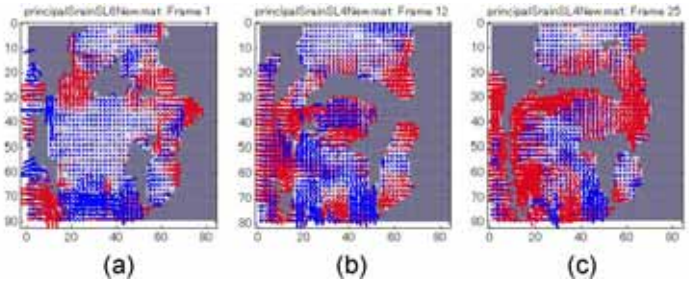

Figure 4. The principal strain pattern obtained using the HARP MRI analysis method. The frames are the same as those in Fig. 3.

Frame 25 shows the tongue in the process of protruding. Because the principal strains are all defined by the reference frame (i.e., maximum protrusion) the tongue appears to be horizontally compressed, when it is actually just less expanded than the reference. A comparison with Fig 4(b) confirms this; the horizontal compression of 4(c) is smaller than 4(b). In addition, vertical compression appears in 4(b) that could be verticalis extending the tongue.

5. A Large-Scale Deformation in Model Simulation

According to the observation, we designed a set of muscle force patterns to generate a protrusion-retraction-protrusion movement. The muscle force used in the simulation is shown in Table 1, where the force unit is Newton. The extensions of the muscle name indicate the independent controllable muscle units in this model. *l* and *r* denotes the left and right sides respectively. *u* represents the outer layer near the tongue surface and *s* for the bilateral layers. *a*, *m*, and *p* indicate the anterior, middle, and posterior regions, respectively. Using this force pattern, a large-scale protrusion motion is generated for the tongue. Figure 5 shows an oblique view for the extreme protrusion, top views of the extreme protrusion and the initial shape. One can see that the tongue gets thinner in vertical and left-right directions in the movement.

Table 1. Muscle force pattern (N)

Duration (s)	0.2	0.2	0.2		0.2	0.2	0.2
hg_l	0	0.1	0	hg_r	0	0.1	0
sg_l	0	0.1	0	sg_r	0	0.1	0
sl_l_u_a	4	0	4	sl_r_u_a	4	0	4
sl_l_u_m	4	0	4	sl_r_u_m	4	0	4
sl_l_s	1	0	1	sl_r_s	1	0	1
vert_l_a	4	0	4	vert_r_a	4	0	4
vert_l_m	2	0	2	vert_r_m	2	0	2
tran_l_a	4	0	4	tran_r_a	4	0	4
tran_l_m	2	0	2	tran_r_m	2	0	2

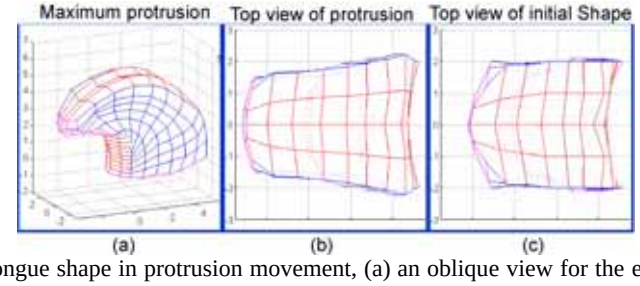

Figure 5: The tongue shape in protrusion movement, (a) an oblique view for the extreme protrusion, (b) a top view of the extreme protrusion, and (c) a top view of the initial shape

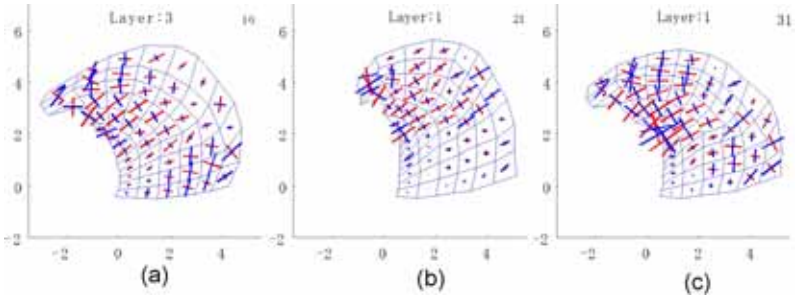

Figure 6. The principal strain in a protrusion-retraction-protrusion, (a) the first protrusion in the midsagittal plane, (b) retraction in the parasagittal plane, and (c) second protrusion in parasagittal plane.

Using the proposed method, the principal strain patterns are obtained for the protrusion-retraction-protrusion and shown in Figure 6. Panel (a) shows the principal strain of the midsagittal plane. As known from the give force, the muscle SL, V, and T were activated in forming this shape. The compression is seen in the tongue tip area and the tongue root area, which is similar to Fig. 4 (a). In the retraction, the HG and SG were activated with a small force of 0.1 N. The compression pattern of the Panel (b) reflects this muscle activation, which is similar to Fig. 4 (b). When the tongue protruded again, the compression is almost seen in the same areas in the simulation (c) and observation in Fig. 4 (c). As shown in (a), some compression vectors are seen in the bottom area of the tongue, while a similar compression pattern was also seen in the observation. According to the known force pattern, it is clarified that both the compression patterns are artifact by means of the AbS analysis.

6. Conclusions

This study proposed a method to estimate muscle activation from local deformations of the tongue using the analysis-by-synthesis (AbS) method. The validity of the method was confirmed by comparing the estimated muscle activation pattern with a known pattern. A simple task of the tongue protrusion-retraction-protrusion was observed in tMRI and was analyzed by a HARP MRI analysis. According to the observation, we designed a force pattern and used it to generate a protrusion-retraction-protrusion. The compression pattern obtained using the proposed method was almost consistent with the observation. Since the muscle activation pattern is known in model simulations, it provides reliable evidence for estimating muscle activation patterns from tMRI using the AbS method.

However, there are couple problems in the current study. One is concerned with the implicit reference used in the HARP MRI analysis, which makes comparison of observation and simulation more complicated. Another problem is the tongue shape in observation is different from that used in the simulation. These problems are remaining in the further study.

Acknowledgements:
This study is supported in part by Grant-in-Aid for Scientific Research of Japan (No. 17300182).

References

Baer T, Alfonso, J., & Honda, K. (1988) Electromyography of the tongue muscle during vowels in /?pvp/ environment. Ann. Bull. R. I. L. P., Univ. Tokyo *7:*7-18.

Dang J, & Honda, K. (2001) A physiological model of a dynamic vocal tract for speech production. Acoustical Science and Technology *22,:*415-425.

Dang J, & Honda, K. (2004) Construction and control of a physiological articulatory model. J. Acoust. Soc. Am. *115:*853-870.

Epstein M, & Stone, M. (2003) Speech task and timing considerations in MRI research. ASA Nashville Conference.

Hashi M, Westbury, J., & Honda, K. (1998) Vowel posture normalization. JASA *104:*2426-2437.

Okadome T, & Honda, M (2001) Generation of articulatory movements by using a kinematic triphone model. JASA*:*453-463.

Parthasarathy V (2006) CHARACTERIZATION OF HARMONIC PHASE MRI: THEORY, SIMULATIONS, AND APPLICAT. Eng., The Johns Hopkins University, Baltimore, Maryland.

Stone M (2003) Tissue strains and tongue shapes: Combining tMRI and Ultrasound. 15th ICPhS*:*273-276.

Stone M, Davis, E., Douglas, A., Ness Aiver, M., Gullapalli, R., Levine, W., and Lundberg, A. (2001) Modeling motion of the internal tongue from tagged cine-MRI images. J. Acoust. Soc. Am. *109:*2974-2982.

Takano S, & Honda, K. (2004) Observation of tongue deformation during vowel production by tagging-MRI. Proc. ASJ Fall Meeting, pp. 263-264.

Takemoto H (2001) Morphological Analyses of the Human Tongue Musculature for Three-dimensional Modeling. J. SLHR *44:*95-107.