Proceedings of the 15th International Congress of Phonetic Sciences. Barcelona, Spain.
August 3–9, 2003. Universitat Autònoma de Barcelona.

Tissue strains and tongue shapes: Combining tMRI and Ultrasound.

Maureen Stone*, Vijay Parthasarathy+, Khalil Iskarous++, Moriel NessAiver**, Jerry L. Prince+

*University of Maryland Dental School and **Dept of Radiology, +Johns Hopkins University,
++Haskins Laboratories
E-mail: mstone@umaryland.edu

ABSTRACT

This project examines tongue surface shapes and internal strains using supine ultrasound and tagged Cine-MRI (tMRI) images. Mid-sagittal tongue slices were recorded for identical utterances and as similar a recording procedure as possible for both instruments. Tongue surface contours, extracted from ultrasound and conventional MRI, were registered. Local measures of Eulerian strain fields were extracted from the tMRI. Visual groupings of local internal strains were done.

1. INTRODUCTION

Ultrasound and tagged MRI provide complementary information about the tongue. The former provides an often crisp image of the edge of the tongue, while the latter is far more informative about tongue-internal kinematics. Previous work used ultrasound to examine tongue surface shape changes [1, 2] and tMRI to examine global and local tongue strains [3, 4]. However, we do not yet understand how changes in the internal strain field affect the shape of the tongue at the surface.

Due to the high empirical correlation between the edge of the tongue in both modalities (see below), a theoretical framework can be developed to allow us to make predictions about the shape of the tongue-edge from tongue-internal strains, and vice versa. This would be quite useful, since ultrasound is cheaper and easier to use than tagged MRI, yet information from the latter is required if we want to understand tongue internal kinematics—a prerequisite for theories of tongue dynamics and control.

There are two theoretical problems: the direct problem is the prediction of tongue-edge shape from tongue-internal strains, and the inverse problem is the prediction of the strains from the external shape. The inverse problem is the one we want to be able to solve, since ultrasound imaging is easier to perform than tagged MRI, but like other inverse problems, this one is highly under-constrained and much harder to solve. However if we are able to make headway in solving the forward problem, we can use its parameters to constrain an optimization or NN solution to the inverse problem.

In the forward problem, we are given a system of principal strains, mathematically represented by a vector field of eigenvectors scaled by their eigenvalues. From that we predict an overall configuration of the edge of the tongue by averaging over the local strains and establishing how they "build up" to global forces creating the surface, shape. This problem is analogous to physical problems which infer global behavior of a system from spatially-distributed local information. Many problems of thermal physics are of this sort. The system contains many molecules, each behaving somewhat differently, and we try to predict global overall behavior of the system, like its temperature, by averaging over the local kinematics. There are several methods of accomplishing this, one of the most successful being renormalization [5], where the system is divided into blocks and the local behavior over the block is averaged. The "atoms" of the system are now the blocks. The blocking is repeated as many times as necessary to obtain the global behavior of the system. If we continue blocking, we would ultimately arrive at a rigid body representation of the deformation.

Both ultrasound and MRI, especially, tagged Cine-MRI (tMRI), are well suited to speech research because they are non-invasive and provide extensive structural information. Ultrasound captures motion of surface contours for sagittal and coronal tongue sections including the pharyngeal regions. tMRI captures motion of tissue points within and on the surface of the tongue. Some technological development is needed, however, to extract strains, to interpret them, and to use them in a complementary manner. In addition, they have limitations that must be accommodated. Both instruments have poorer spatial and temporal resolution than point tracking instruments. In addition, ultrasound does not track tissue points and MRI has a supine gravitational orientation.

The present paper is part of a long-term study, which has three goals. The first is to register ultrasound and tMRI data in order to validate them. The second is to transform supine tMRI data into an equivalent upright data set. The third is to understand the relationship between the dynamic forces of muscles and the kinematics of the tongue surface. This study addresses the third issue and attempts to understand the muscle-surface mapping by looking at strain fields, which are directly related to force

ISBN 1-876346-48-5 © 2003 UAB

fields. In the present paper midsagittal data are considered. Supine ultrasound and MRI tongue contours are overlaid. In addition, principal strains are calculated from tMRI data of internal deformations for each frame of the motion using Harmonic Phase Imaging (HARP-MRI) [6]

HARP-MRI is a tool to estimate tissue motion from tMRI data using sinusoidal tags [7]. Its main use is to extract key indices of motion like Eulerian and Lagrangian strain, velocity fields and displacement fields. The basic assumption of HARP is that tissue motion is directly related to the phase of the sinusoidal tags. However, there are no tags in the airway beyond the tongue surface, therefore, HARP has reduced accuracy and robustness near the surface of the tongue. This is exactly where ultrasound gives its most useful information due to the impedance mismatch between the tissue and air. However, ultrasound, like other surface measurement modalities, is unable to provide information about the internal strains that cause the surface changes. The challenge is to register these complementary data sets to understand the relationsip between internal and external tongue kinematics.

Goals of the present study:

- Compare MRI and ultrasound surface data.
- Calculate principal strains for /i/ to /ɑ/ deformations.
- Identify local regions of deformation.

2. METHODS

2.1 Subject and speech materials

This paper presents data from a single male subject aged 24, with Tamil accented English (the second author). The subject repeated the vowel-vowel (VV) utterances /i/-/ɑ/, /ɑ/-/u/, and /i/-/u/. These VV utteances represent the boundaries of the vowel space in English and should provide a fairly large range of motion for the tongue. Subject precision is critical in the tMRI data collection because a single image sequence is the summation of several repetitions. Considerable exploration of pretesting methods has been done, and a specific test has been determined as optimal for the current protocol [8]. Therefore, the subject was pretested according to this method. The pretest consisted of repeating the proposed speech materials (and others) to a metronome composed of tMRI noise bursts at 1 sec intervals. The subject was trained to time his repetitions to the burst cycle. The pretest is virtually identical acoustically to the tMRI experience. The data were subsequently analyzed acoustically to determine temporal precision of repetition time. The present subject was very precise, with standard deviations of 20 ms or less.

2.2 Data Recording and Analysis.

Fort the ultrasound recording, the ultrasound transducer was attached to a cervical collar in a midsagittal orientation. The collar was positioned around the subject's neck; he was seated in a dental chair and lowered into supine position. The subject spoke to the beat of a digital metronome set at a 1 second repetition time. He repeated each VV utterance eight times. A digital recording and a back-up videotape were made of the ultrasound and acoustic data. Repetitions 2-6 were measured. Upright data and para-sagittal data were also collected but will not be reported here.

For the tMRI recording, an identical cervical collar was fitted to the subject's neck, but without the ultrasound transducer. This was done so that the jaw restrictions would be the same in both recording conditions. The subject repeated each VV eight times per slice; the even numbered repetitions were used to construct the slice, the odd numbered repetitions were 'set-up' measures. The result was a mid sagittal image sequence of one (summed) repetition of each VV. The one second repetition time contained two components: 800 ms recording time and 200 ms set up time for the next recording. Thus the tMRI data, collected at 15 Hz (half the ultrasound frame rate), yielded 12 time-phases per utterance. Non-tagged images sequences were also collected. The tMRI data were collected using a new procedure, called MICSR (Magnitude Image C-SPAMM Reconstruction Method) [9]; principal strains were extracted using HARP and then further analyzed using principal strains.

3. RESULTS

3.1 Ultrasound-to-MRI surface registration

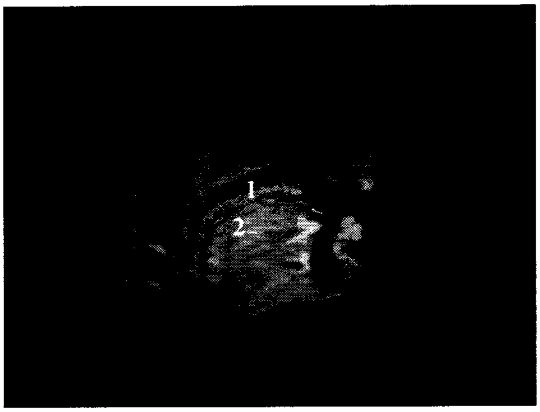

Figure 1. Overlay of thresholded ultrasound "noise" on MRI image of /ɑ/.

Figure 1 shows that the lower edge of (1) the ultrasound noise (i.e. the tongue surface), and (2) the upper edge of the MRI tongue align quite well. This suggests that comparisons of ultrasound surfaces with tMRI strains will be a useful methodology.

3.2 Tag Results

Figure 2 shows MICSR data for midsagittal frames 4 and 5. In frame 4, the anterior tags are slightly compressed vertically consistent with GGA compression. The upper tongue surface is lowering, but the tongue back is not yet moving backward. In frame 5, maximum /ɑ/, the vertical compression continues and the lower tongue expands passively, backing the tongue root for /ɑ/.

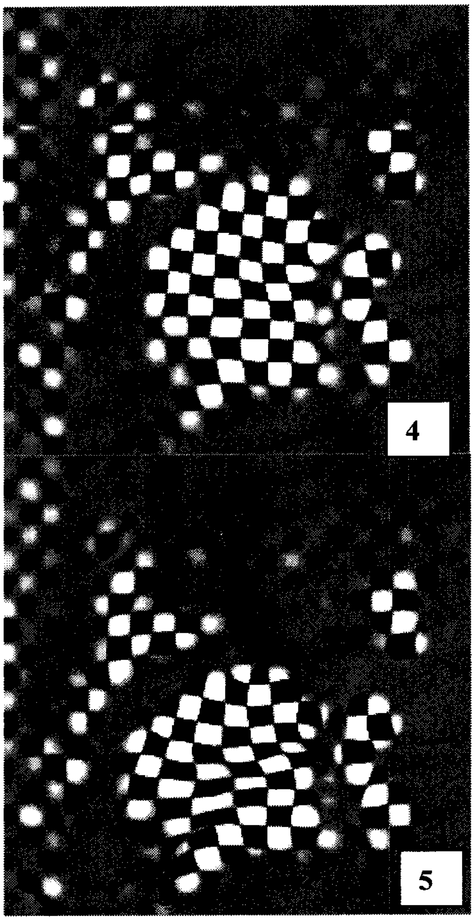

Figure 2. Midsagittal deformed tags at frames 4 and 5 (max /ɑ/). Non-square shapes reflect expansion and compression.

3.3 Principal Strains.

Figure 3 shows principal strain 2 (compression) for frame 4. The strains combine active muscle contraction and passive tissue compression. Over 20% compression is seen in the center and tip (circled), in the line of action of genioglossus anterior.

In Figure 4, maximum /ɑ/, vertical compression is seen throughout the lower tongue with very strong force at the center and obliquely in back. The oblique angle of the force (not the direction) is consistent with the backward and downward expansion of /ɑ/ and with HG contraction.

In figure 5, horizontal expansion during maximum /ɑ/ is seen. Expansion is smaller and not exactly identical to the compression pattern. This is because some expansion occurs laterally.

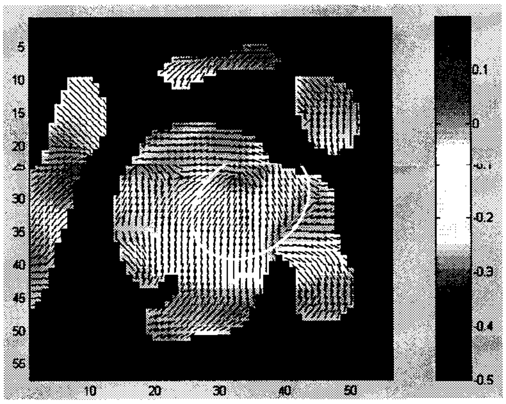

Figure 3. Compression at midsagittal frame 4.

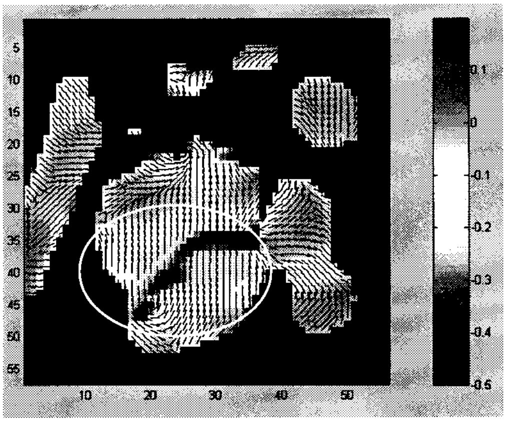

Figure 4. Compression at midsagittal frame 5.

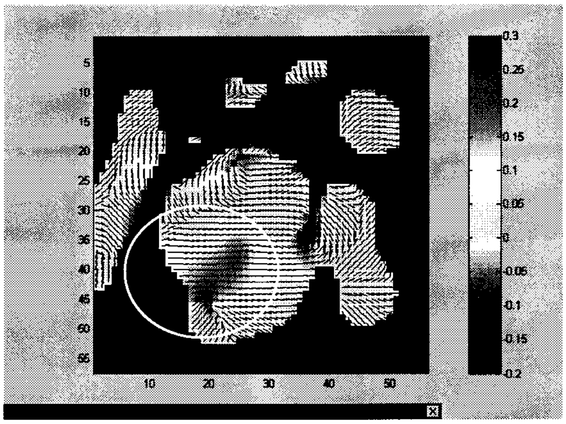

Figure 5. Expansion at midsagittal frame 5.

ISBN 1-876346-48-5 © 2003 UAB

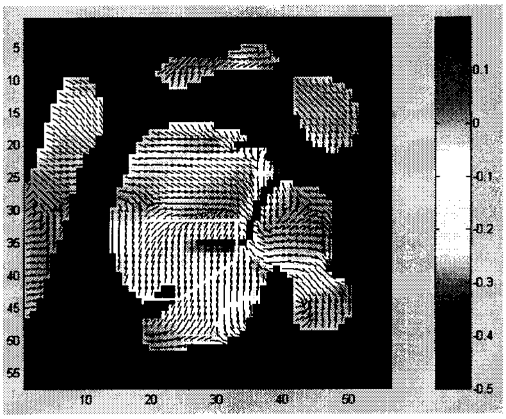

Figure 6 Compression at LEFT sagittal frame 4.

Left sagittal data were also collected. Although the maximal backward expansion occurred at frame 5, the maximum compression in the direction of HG did not occur at frame 4 (above), rather it was earlier, at frame 3 below in Figure 7.

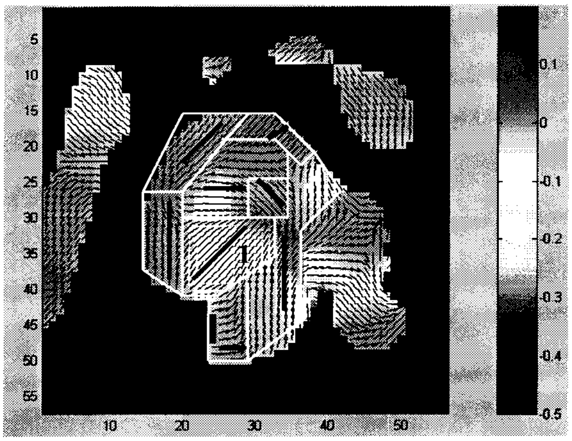

Figure 7 Compression patterns (black bars) in LEFT sagittal frame 3. Regions of common direction are outlined in white. (1) shows compression in the HG direction.

The first step in the process of relating local-strains to global shape is representing each direction of strain by the angle it makes with the horizontal direction. This gives a single parameter on which to do the blocking. To start we have drawn regions of common compression direction in Figure 7. Strain directions are indicated with black lines.

4. CONCLUSIONS

Comparisons of MRI and ultrasound surface data look promising and will be useful in mapping surface shapes to internal strains as well as improve the interpretation of noisy images.

MICSR and HARP successfully allow calculation of principal strains from tMRI, thus rendering it a useful tool for extracting tissue point motion.

Extracting tag deformations and strain information from tMRI sequences is complicated by the interaction of active and passive compression. Thus even though local regions of compression can be identified in the principal strains, aligning them appropriately and unambiguously with specific muscle groups will be more difficult. Extracted compression patterns can be entered into a predictive model, however, for mapping of surface to muscle activity. Ultrasound and tMRI data can then validate the model.

REFERENCES

[1] Stone, M., Lundberg, A., Davis, E., Gullapalli, R., and Ness Aiver, M. (1997): "Three-dimensional coarticulatory strategies of tongue movement," *Proceedings of the 5th European Conference on Speech Communications_and Technology*, Rhodes, Greece, **1**, 31-34.

[2] Slud, E., Smith, P., Stone, M., and Goldstein, M. (2002) Principal Components Representation of the Two-Dimensional Coronal Tongue Surface. *Phonetica* vol. 59, Nos. 2-3, pp. 108-133.

[3] Stone, M., Davis, E., Douglas, A., Ness Aiver, M., Gullapalli, R., Levine, W., and Lundberg, A. (2001) Modeling the motion of the internal tongue from tagged cine-MRI images. *J. Acoust. Soc.Am.* 109(6), 2974-2982.

[4] Stone, M., Davis, E., Douglas, A., Ness Aiver, M., Gullapalli, R., Levine, W., and Lundberg, A. (2001) Modeling Tongue Surface Contours from Cine-MRI images. *J. Speech, Lang, Hearing Research*, **44**, 1026-1040.

[5] Kadanoff, L., (2000) *Statistical Physics: Statics, dynamics and Renormalization.* Singapore: World Scientific.

[6] Parthasarathy, V, NessAiver, M. Prince, J.L., and Stone, M. Tracking Tongue Motion from Tagged Magnetic Resonance Images Using Harmonic Phase Imaging (HARP_MRI). In this Proceedings.

[7] Osman, N.F., and Prince, J.L. Visualizing Myocrdial Function Using HARP MRI. *Phys. Med. Biol.* 45 (6) : 1665-1682, 2000.

[8] Epstein, M., and Stone, M. Speech task and timing considerations in MRI research *J. Acoust. Soc. Amer.* Nashville Conference, June. 2003

[9] NessAiver, M., Stone, M., Partharasathy, V., Prince, J. Using Tagged Cine-MRI to Extract Muscle Activity Within the Tongue. *Radiol. Soc. of N. Am. (RSNA) meeting*, Chicago, IL, USA, November, 2002.