TONGUE STRUCTURAL MODEL: INTEGRATING MRI DATA AND ANATOMICAL STRUCTURE INTO A FINITE ELEMENT MODEL OF THE TONGUE

Chao-Min Wu

Biomed. Eng., Ohio State Univ.
Columbus, OH 43210, USA,
and ATR HIP Res. Labs.,Kyoto, Japan

Reiner Wilhelms-Tricarico

Speech Commun. Group,
Res. Lab. of Electronics, MIT
Cambridge, MA 02139, USA

ABSTRACT

A method is presented that has the potential to supplement MRI data with anatomical structural information from other resources that are not available in the MRI data. In the MRI data landmarks are specified which are then matched with landmarks in drawn specimens. Images or other geometric representations of specimens (such as fiber direction fields) are warped by using a thin-plate spline mapping in either two or three dimensions.

1 INTRODUCTION

The goal of this research is to obtain a 3-dimensional (3-D) tongue structural model as a preparation for dynamic simulations of the tongue during speech articulation. The tongue structural model, which is a new finite element representation of the human tongue, adapted to each individual's anatomy, will be developed by integrating tongue shape information from MRI scans and anatomical drawings to form an anatomical model, and incorporating into it the finite element structural framework.

We want to emulate closely the morphology and to some extent the biomechanics of a particular speaker, in a computational simulation of the tongue and other vocal tract structures. In particular, for the tongue, a finite element model is needed and the geometry of this model needs to be specified. The finite element model of the tongue will contain the intrinsic structural information that closely matches the speaker's anatomy (cf. [1]).

Some of the anatomical information can be extracted from magnetic resonance images. However, many important details, such as the directions of the muscle fibers, can only be obtained to a very limited extent from the MRI data. This information has to be supplemented from anatomical knowledge, which exists in the form of detailed and accurate anatomical drawings of specially studied specimens of tongue (see Miyawaki, [2]). These drawings have previously been used as reference to construct finite element models of the tongue (cf. [3] and [4]). This paper presents the basic methods that we use to combine MRI data with anatomical drawings.

2 MRI AND ANATOMICAL DATA

The MRI data consist of two stacks of transversal sections of the oropharyngeal region. Table 1 shows the specifications of the MRI data. These data were from a Shimadzu MRI machine, SMT-100GUX, which has a static magnetic field density of 1.0 T.

The data were obtained from two different speakers. During the MRI assessment, the two speakers articulated a constant vowel, Japanese /a/ in one case, and English /i:/ in the other case. From the data, areas of a size of 110 x 110 pixels surrounding the tongue were extracted. For the purpose of a larger and smoother display, they were interpolated and magnified, so that for further processing a stack of MRI images was used that consisted of 256×256 pixel images with a pixel size of (0.4196mm x 0.4196mm).

Table 1: *Specifications of the MRI data.*

Axial Imaging Parameters		
	Subj. KH (Japanese)	Subj. DO (English)
TR (ms)	800	800
TE (ms)	18	18
Images	26	26
Thickness	0.5cm	0.5cm
Interscan skip	0.0 cm	0.0 cm
View field	25 cm	25 cm
Matrix Size	256 x 256	256 x 256
Scan Time	4:05 min	4:05 min

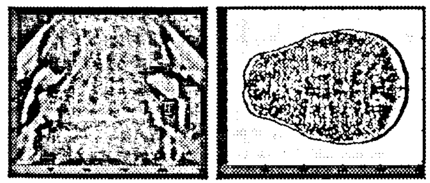

Figure 1: *(a) Enlarged oral portion of transverse transection at a level of the third vertebra from Japanese vowel /a/. (b) Full tongue sketch of transverse section at a compatible level of figures 1(a).*

Figure 1(a) shows one example of extracted regions around the tongue. It is from the Japanese subject.

Miyawaki's drawings ([2]) contain three different planes in which tongues were sliced. This makes it possible in principle to approximately reconstruct the three-dimensional structure of the tongue, including the fiber directions. Miyawaki's drawings were scanned and digitized in the computer and the transversal sections were used together with their mirror images. *Figure 1(b)* shows an example of the processed result.

Neglecting some deformations due to theslicing techniques used, Miyawaki's drawings can be aligned and stacked to obtain roughly the shape of the originally used tongue specimen. For the MRI data, the images are aligned properly. The stack of MR images can be used as a data base which allows the extraction of some landmark points in three dimensions.

Geometrically the mapping we are looking for is an interpolation mapping. We are given two sets of points: One set consists of landmarks in the stack of drawings of the tongue. The other set consists of the corresponding landmarks in the stack of MRI slices. The interpolation mapping has to map the first set of landmark points one-to-one onto the other set of landmark points. This can be illustrated easier in two dimensions.

Figure 2 shows in panel A a typical drawn tongue section and in B an outline of a tongue section as it may be obtained from MRI. In *Fig.* 2-C specified landmark pairs are shown. D shows the result of the mapping: The muscle fiber directions that are visible in A are warped and fill out the MRI contour. Panels E and F show the warping of a rectangular grid by the mapping. In the example of Fig. 2, the thin-plate spline mapping (*cf.* Bookstein,([5]) has been used to warp the figures. The thin-plate spline mapping is used for two- and three-dimensional interpolation because it has the advantage over other mappings that – among all possible interpolation mappings – it minimizes the bending energy of the interpolating function.

In the following, a short summary of the thin-plate spline mapping is given, starting at a definition of the thin-plate spline interpolation of a scalar field in two dimensions.

Let $\mathbf{P} = (x, y)$ be any point in the plane. Let $\{\mathbf{P}_i = (x_i, y_i)\}, i = 1, \ldots, N$ be a set of specific points in the plane and h_i a scalar associated with point $\mathbf{P}_i$.

The thin-plate spline interpolation of the field h_i is the function $f(x, y)$ which fulfills the following requirements:

1. $f(x_i, y_i) = h_i$ for all i.
2. The function f minimizes the following functional (bending energy):

$$I_f = \iint_{\mathbf{R}^2} \left(\left|\frac{\partial^2 f}{\partial x^2}\right|^2 + 2\left|\frac{\partial^2 f}{\partial x \partial y}\right|^2 + \left|\frac{\partial^2 f}{\partial y^2}\right|^2\right) dx\,dy.$$

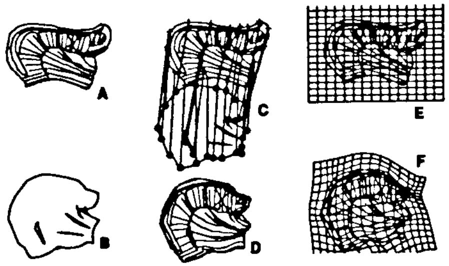

Figure 2: *Demonstration of the two-dimensional thin-plate spline mapping. (A) A typical hand-drawn sketch of a tongue section as in Miyawaki's tongue drawings. (B) Outline of tongue as could be obtained from MRI (hand-drawn). (C) Specification of corresponding landmarks (here 27 landmark pairs). (D) Panel A mapped onto panel B with thin-plate spline. (E) - (F) Demonstration of the mapping by warping a grid.*

It can be seen that in two dimensions the function f is based on the fundamental solution, U, of the biharmonic equation, $\Delta^2 U = \delta(x, y)$, where δ is the Kronecker delta function.

In two dimensions this fundamental solution function is $U(r) = r^2 log\, r$, where $r^2 = x^2 + y^2$, and in three dimensions the solution function is $U(r) = |r|$. For example, in two dimensions, the function $f(\mathbf{P})$ is

$$f(\mathbf{P}) = \sum_{i=1}^{N} w_i U(|\mathbf{P} - \mathbf{P}_i|) + a_0 + a_x x + a_y y$$

In three dimensions an additional factor $a_z z$ appears in the above formula, and $U(r) = |r|$ needs to be taken. The coefficients w_i and a_x, a_y, a_z are determined from a given the set of landmark pairs. This is shown in [5].

The extension of the method to interpolation of mappings between two three-dimensional sets of homologous points is straight-forward. One interpolating function such as the above $f(x, y)$ needs to be computed for each spatial direction.

In general, finding reliable landmarks turned out to be not so easy for the MRI data that we currently have. The location of the tip of the tongue, the path of a groove on the tongue and the rough geometry of the genioglossus and hyoglossus muscles could be assessed with confidence. An imaging anatomy atlas [6] was used as guidance to interprete the MRI data. The styloglossus muscle is also visible in the MRI data. However, the quality was insufficient to extract reliable landmark points.

Figure 3 shows a first result of the computed mapping (i.e., the anatomical model). To the left is the tip of the tongue (anterior). The thick line shows the trace of a groove on the tongue surface that can be seen in the MRI data. The line shown on the right side (posteriorly) marks the midsagittal plane on the back of the tongue, obtained from interpreting the MRI data. The circles represent the landmark points after the 3-D thin-spline mapping was applied.

3 CONCLUSIONS AND PLANNED WORK

A method was presented that has the potential to supplement MRI data with structural information from other resources that are not available in the MRI data. In the MRI data landmarks are specified which are then matched with landmarks in drawn specimens. Images or other geometric representations of specimens (such as fiber direction fields) are warped by using a thin-plate spline mapping in either two or three dimensions.

For future work, a finite element grid will be enscribed into the volume and mapped as a whole onto the MRI volume data to obtain a finite element model of an individual tongue in one (arbitrary) articulatory configuration. The finite element model can be used as a reference model for dynamic simulations of tongue movements.

(This work was supported in part by ATR HIP Res. Lab., an NSF grant to O. Fujimura, and a gift from ATR ITL to O. Fujimura. We would like to thank Dr. K. Honda from ATR-HIP

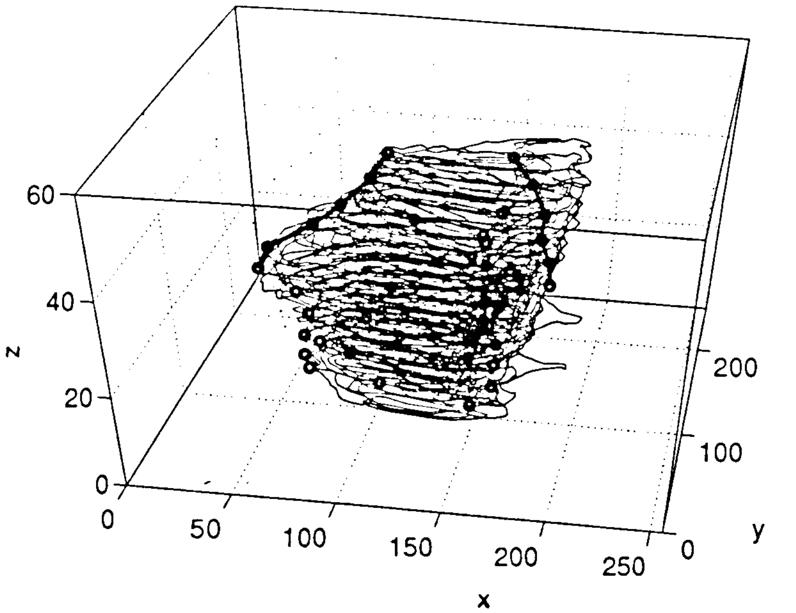

Figure 3: *The result of a 3-d mapping shown in a lateral view for data from one subject articulating the Japanese vowel /a/(in perspective projection.)*

and Dr. J. Negulesco from OSU for helpful discussion on interpreting MRI data.)

4 REFERENCES

[1] Wilhelms-Tricarico, R. (in press), Physiological modeling of speech production: methods for modeling of soft-tissue articulators. *J. Acoust. Soc. Am.*.

[2] Miyawaki, K. (1974), A study of the musculature of the human tongue. *Ann. Bul. Research Institute of Logopedics and Phoniatrics, Univ. Tokyo*, 8:23–50.

[3] Kiritani, S., Miyawaki, K., Fujimura, O., and Miller, J.E. (1976), A computational model of the tongue. *Ann. Bul. Research Institute of Logopedics and Phoniatrics, Univ. Tokyo*, 10:243–251.

[4] Hashimoto K., and Suga, S. (1986), Estimation of the muscular tensions of the human tongue by using a three-dimensional model of the tongue. *J. Acoust. Soc. Japan*, (E) 7,1:39–46.

[5] Bookstein, F.L. (1989), Principal warps: thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-11(6):567–585.

[6] Schnitzlein, H.N., and Murtagh, F.R. (1990), *Imaging Anatomy of the Head and Spine: A photographic Color Atlas of MRI, CT, Gross and Microscopic Anatomy in Axial, Coronal, and Sagittal Planes.* Urban & Schwarzenberg, Baltimore.