

A Statistical Shape Space Model of the Palate Surface Trained on 3D MRI Scans of the Vocal Tract

Alexander Hewer^{1,3} Ingmar Steiner^{1,3} Timo Bolkart³ Stefanie Wuhrer² Korin Richmond⁴

¹DFKI Language Technology Lab

³Cluster of Excellence MMCI,
Saarland University

²INRIA Grenoble Rhône-Alpes

⁴Centre for Speech Technology Research,
University of Edinburgh

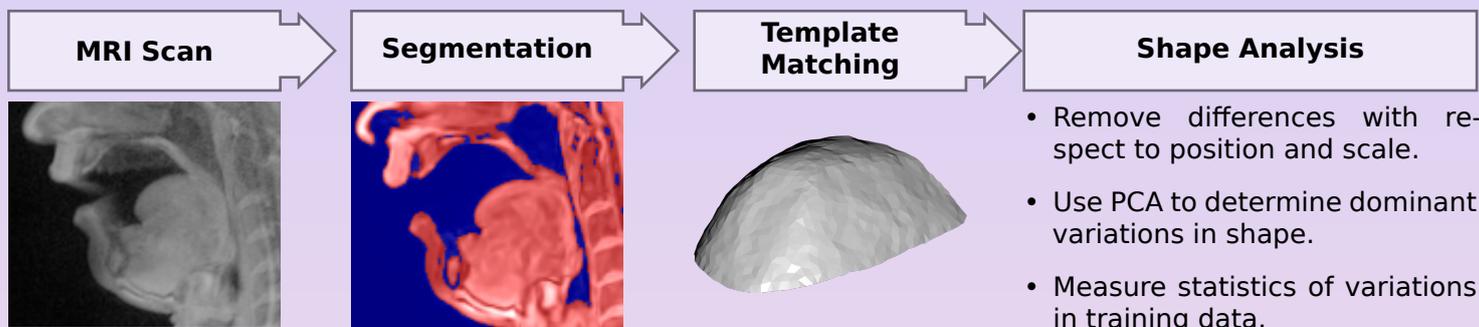
1 Motivation

- The palate plays an important role in articulation, e.g., in the production of obstruents such as /ʒ, ʃ, j/.
- Training a statistical palate shape space can help with, e.g.,
 - creating unseen shapes that may be used in computer graphics;
 - registering unknown data that might be incomplete, such as palate traces acquired by electromagnetic articulography (EMA).

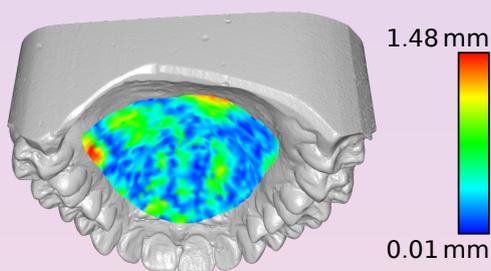
2 Datasets

- Volumetric magnetic resonance imaging (MRI) is regarded as the state-of-the-art modality for imaging the vocal tract.
- We use two MRI datasets for training: The one of Baker [2] (1 speaker) and the full dataset of the Ultrax project [1] (11 speakers).
- For evaluation, we use the volumetric MRI subset of the mngu0 corpus [3]. It contains data from one male speaker, including high-resolution 3D scans of a plaster cast of his teeth and palate.

3 Training the Model

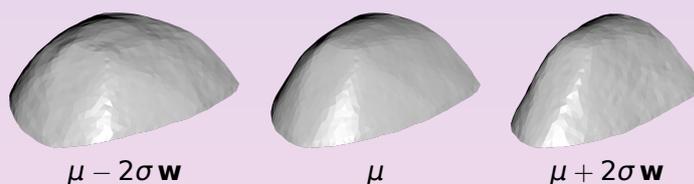


4 Registering Unknown Data



Registered palate shape in the /b/ scan of mngu0. Colors show the distances to the 3D scan of the plaster cast of the speaker's teeth and palate.

5 Generating New Shapes



Generated shapes by manipulating the coefficient for the principal component \mathbf{w} related to the palate width. μ is the mean shape of the space model and σ the observed standard deviation of the coefficient in the training set.

6 Conclusion

- Our model trained on 12 speakers is already able to register unknown data.
- It also allows to generate new palate shapes.
- In the future, we plan to increase our training set to improve the coverage.
- Additionally, we want to investigate how to use the palate model to efficiently obtain the full palate shape from EMA palate traces.

References

- [1] Ultrax: Real-time tongue tracking for speech therapy using ultrasound. <http://www.ultrax-speech.org/>, 2014.
- [2] A. Baker. A biomechanical tongue model for speech production based on MRI live speaker data. <http://www.adambaker.org/qmu.php>, 2011.
- [3] I. Steiner, K. Richmond, I. Marshall, and C. D. Gray. The magnetic resonance imaging subset of the mngu0 articulatory corpus. *Journal of the Acoustical Society of America*, 131(2): EL106--EL111, 2012. doi:10.1121/1.3675459.