Real-time feedback of speech movements based on optopalatography

Simon Preuß, Christiane Neuschaefer-Rube, Peter Birkholz

July 8, 2013

Abstract

Speech is the product of articulation and information on the sound forming articulatory gestures of the tongue and the lips is of major interest not only in speech therapy and phonetic research, but in speech synthesis and coding as well. Various attempts have been made to describe these important articulators using established imaging techniques like ultrasound, electromagnetic articulography (EMA), or X-ray microbeam [3], but so far none of these has proven to be fast, precise, and at the same time cheap enough to yield satisfying results that can be applied in day-to-day applications. However, the relatively little-known glossometry [2, 4], also called optopalatography (OPG) [6], has shown some good results in measuring the tongue contour in the anterior oral cavity [1] and holds the most promise as a clinically relevant tool as it is relatively cheap to produce, easy to employ, and reasonably precise. Its biggest limitation, though, is that its coverage does not extend to the large part of the tongue in the posterior region of the oral cavity. To remedy this drawback, we introduce in this poster an animated 2D articulation model controlled by real-time optopalatographic measurements of the positions of the upper lip and the tongue in the anterior oral cavity that are extended by linear prediction of the posterior part of the tongue. The measurement system is an improvement on a previous prototype with increased spatial resolution and an enhanced close-range behavior. The prediction coefficients were determined and evaluated using a corpus of vocal tract traces of 25 sustained phonemes. The model represents the tongue motion and the lip opening physiologically plausible during articulation in real-time. Finally we discuss how the model could be improved by speaker adaptation or by incorporating additional data like electropalatographic (EPG) measurements in our ongoing project to develop a 3D real-time representation of the vocal tract [5].

References

[1] Peter Birkholz and Christiane Neuschaefer-Rube. A new artificial palate design for the optical measurement of tongue and lip movements. In

Matthias Wolff, editor, *Studientexte zur Sprachkommunikation: Elektronische Sprachsignalverarbeitung 2012*, pages 89–95. TUDPress, Dresden, Germany, 2012.

[2] Chiu-Kuang Chuang and William S.Y. Wang. Use of optical distance sensing to track tongue motion. *Journal of Speech and Hearing Research*, 21:482–496, 1978.

[3] Margaret M. Earnest and Ludo Max. En route to the three-dimensional registration and analysis of speech movements: instrumental techniques for the study of articulatory kinematics. *Contemporary Issues in Communication Science and Disorders*, 30:2–25, 2003.

[4] Samuel G. Fletcher, Martin J. McCutcheon, Stephen C. Smith, and Wilson H. Smith. Glossometric measurements in vowel production and modification. *Clinical Linguistics and Phonetics*, 3(4):359–375, 1989.

[5] Simon Preuß, Peter Birkholz, and Christiane Neuschaefer-Rube. Prospects of EPG and OPG sensor fusion in pursuit of a 3D real-time representation of the oral cavity. In Petra Wagner, editor, *Studientexte zur Sprachkommunikation: Elektronische Sprachsignalverarbeitung 2013*, pages 144–151. TUDPress, Dresden, 2013.

[6] Alan A. Wrench, A. D. McIntosh, and William J. Hardcastle. Optopalatograph (opg): A new apparatus for speech production analysis. In *4th International Conference on Spoken Language Processing (ICSLP 1996)*, pages 1589–1592, Philadelphia, PA, USA, 1996.