

Change Measures for Ultrasound: Relating Pixel Difference on Raw Data to Nearest Neighbour on Splines

Pertti Palo¹, Sonja Schaeffler¹, and James Scobbie¹

¹Clinical Audiology, Speech and Language (CASL) Research Centre, Queen Margaret University, Scotland, UK

pertti.palo@taurlin.org, sshaeffler@qmu.ac.uk, jscobbie@qmu.ac.uk

Background: Using appropriate dimension reduction methods can help in making the analysis of large datasets tractable. Our theoretical interest is in delayed naming (Rastle et al., 2005) and speech initiation (Schaeffler et al., 2014). We have used a dimension reduction method called Pixel Difference (PD) to aid in annotating articulatory reaction time (movement onset detection) in Ultrasound Tongue Imaging (UTI) videos (Palo et al., 2015). PD calculates the Euclidean distance between consecutive raw (uninterpolated) ultrasound images and yields a curve that gives change rate as a function of time over the duration of a video.

The reaction times based PD appear to be on average shorter than those measured by manually analysing ultrasound videos (Fig. 1 and Palo, 2019). Given that PD takes into account changes at any of the pixels in the raw (uninterpolated) images, it seems likely that while human video annotators look for movement of the major structures visible in an ultrasound frame PD is sensitive to smaller local changes. In this study we seek to determine if this is indeed the case by comparing manual PD curve annotation to manual annotation of Nearest Neighbour Distance (NND) (Zharkova and Hewlett, 2009) curves.

NND is a method of calculating the distance between two splines by taking a sum over the squared distances from each node on one spline to their nearest neighbours on the next spline. We use it to create a curve like the PD curves.

Materials: The data we analyse comes from Experiment 3 as reported by Palo (2019). We recorded one 40-year-old native Finnish speaking participant (the first author) in a delayed naming experiment which combined simultaneous acquisitions of audio and UTI and was controlled with AAA.

Each trial began with the target word being displayed on a computer screen. The participant was instructed read the word internally while remaining at rest until he heard the go-signal (50 ms long 1 kHz beep), which was played out after a random delay of 1.2-1.8 s from the beginning of the UTI recording. He was instructed to produce the target word as soon as possible after he observed the beep.

Discussion: The PD annotation is already done and we will next automatically spline the data in AAA, and calculate and annotate the NND curves. Results will be available by the time of the conference and the code

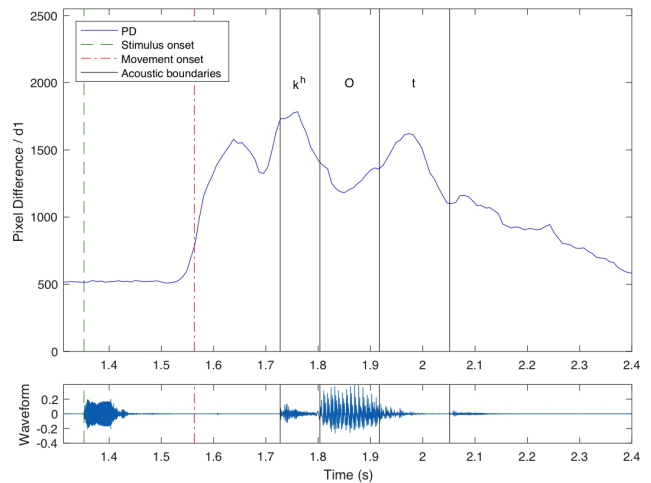


Figure 1: Example of a pixel difference contour and the corresponding acoustic waveform: A participant says the word 'caught'. Stimulus onset marks the go signal (a 50 ms long 1 kHz beep visible at the beginning of the waveform), movement onset was *manually labelled based on the UTI video*.

used in the NND and PD calculations and annotations will be made publicly available.

References

- Palo, P. (2019). *Measuring Pre-Speech Articulation*. PhD thesis, Queen Margaret University, Edinburgh.
- Palo, P., Schaeffler, S., and Scobbie, J. M. (2015). Effect of phonetic onset on acoustic and articulatory speech reaction times studied with tongue ultrasound. In *Proceedings of ICPHS 2015*, Glasgow, UK.
- Rastle, K., Harrington, J. M., Croot, K. P., and Coltheart, M. (2005). Characterizing the motor execution stage of speech production: Consonantal effects on delayed naming latency and onset duration. *Journal of Experimental Psychology: Human Perception and Performance*, 31(5):1083 – 1095.
- Schaeffler, S., Scobbie, J., and Schaeffler, F. (2014). Measuring reaction times: Vocalisation vs. articulation. In *Proceedings of 10th ISSP*, pages 383 – 386.
- Zharkova, N. and Hewlett, N. (2009). Measuring lingual coarticulation from midsagittal tongue contours: description and example calculations using English /t/ and /a/. *Journal of Phonetics*, 37:248 – 256.

Keywords: Methodological research, pixel difference, nearest neighbour distance.