# Comparison of distance measures in tongue contour traces of ultrasound images

*Tamás Gábor Csapó*[1,2], *Andrea Deme*[2,3], *Tekla Etelka Gráczi*[2,4], *Alexandra Markó*[2,3]

[1] Dept. of Telecomm. and Media Informatics, Budapest Univ. of Technology and Economics, Budapest, Hungary
[2] MTA-ELTE Lendület Lingual Articulation Research Group, Budapest, Hungary
[3] Department of Phonetics, Eötvös Loránd University, Budapest, Hungary
[4] Research Institute for Linguistics, Hungarian Academy of Sciences, Budapest, Hungary

csapot@tmit.bme.hu, {deme.andrea,marko.alexandra}@btk.elte.hu, graczi.tekla.etelka@nytud.mta.hu

**Background:** Numerous measures have been used to quantify the distance between tongue contours extracted from ultrasound images (traced manually or automatically). One type of these measures are comparing two contours directly, e.g. MAD - Mean Absolute Difference (Stone, 2005), RMSD - Root Mean Squared Distance (Stone, 2005), MSD - Mean Sum of Distances (Li et al., 2005) and NND - Nearest Neighbor Distance (Zharkova et al., 2009) [the latter being basically the same as the mean sum of distances]. All of these methods offer advantages and disadvantages, which will be compared here and extended with a novel distance measure to calculate the difference between two tongue contours.

**Methods:** Five Hungarian subjects with normal speaking abilities were recorded while reading aloud sentences and nonsense words (Csapó et al. 2017). The tongue movement was recorded in midsagittal orientation using a "Micro" ultrasound system (Articulate Instruments Ltd.) with a 2–4 MHz / 64 element 20mm radius convex ultrasound transducer at 80–100 fps. After the recordings, the ultrasound frames were extracted as raw scan line data and converted to JPG images. To illustrate the comparison of distance measures we traced the tongue manually on two images (http://apil.parsertongue.com/draw).

An advantage of MAD and RMSD is that they are simple and quick to calculate. However an equal number of data points are necessary on the two contours (denoted by U and V), because if one of the contours in the comparison is significantly longer, not all parts will be involved in the calculation of the difference (see Fig. 1). The advantage of RMSD over MAD is that more weight is given to large differences, and thus global differences are more apparent, since the distances are squared (Stone 2005). Both of the measures are calculated only on the common part of the radial grid (i.e., between indices *a* and *b*):

$$MAD(U,V) = \frac{1}{b-a-1}\sum_{j=a}^{b}|u_j - v_j|$$

$$RMSD(U,V) = \sqrt{\frac{1}{b-a-1}\sum_{j=a}^{b}(u_j - v_j)^2}$$

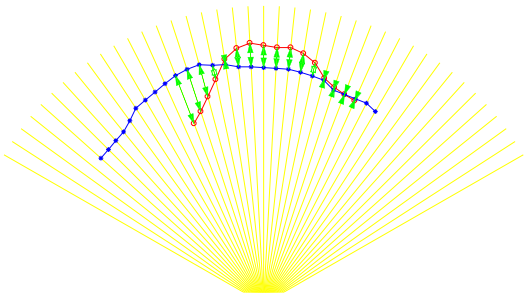

Figure 1: *Calculation of MAD and RMSD using a radial grid*

The advantage of MSD is that it can measure the distance between tongue contours that are not equal in length (i.e. do not have the same number of points, see Fig. 2). MSD is asymmetric in a sense that the U-to-V distance might be different from V-to-U, as shown in Fig. 2. For this reason, both ways are calculated and averaged:

$$MSD(U,V) = \frac{1}{m+n}\left(\sum_{i=1}^{m}\min_j|v_i - u_j| + \sum_{i=1}^{n}\min_j|u_i - v_j|\right)$$

However, MSD can be improved further by taking the squared distances to give more weight to large differences (similarly to RMSD). Therefore, we propose to use the novel RMSSD (Root Mean Sum of Squared Distances) measure:

$$RMSSD(U,V) = \sqrt{\frac{1}{m+n}\left(\sum_{i=1}^{m}\min_j(v_i - u_j)^2 + \sum_{i=1}^{n}\min_j(u_i - v_j)^2\right)}$$

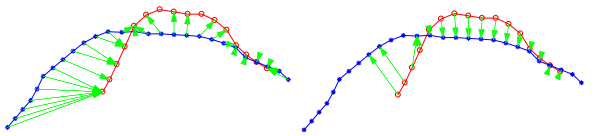

Figure 2: *Calculation of MSD and RMSSD.*
*Left: blue-to-red (U-to-V), right: red-to-blue (V-to-U)*

**Results and discussion:** For simple demonstration, we measured all the distances on the above pair of tongue contours, resulting: MAD = 24.79 pixel, RMSD = 31.51 pixel, MSD = 37.03 pixel and RMSSD = 49.66 pixel. Figure 1 shows that the red contour is significantly shorter than the blue contour, and thus the MAD and the RMSD disregard the non-overlapping segments. In MSD and RMSSD the non-overlapping segments are also involved in the calculation (see Fig. 2 left). Furthermore, RMSSD gives more weight to this kind of visually clear difference; this is why it is the largest.

More detailed analysis (including data from the five speakers) will be carried out by the time of the conference. Based on the analysis we also hope to draw further conclusions on the types of research questions the particular distance measures may be suitable for.

## References

Csapó, T. G., Deme, A., Gráczi, T. E., Markó, A., & Varjasi, G. (2017). Synchronized speech, tongue ultrasound and lip movement video recordings with the "Micro" system. In Challenges in analysis and processing of spontaneous speech.

Li, M., Kambhamettu, C., & Stone, M. (2005). Automatic contour tracking in ultrasound images. Clin Ling Phon, 19(6–7), 545–554.

Stone, M. (2005). A guide to analysing tongue motion from ultrasound images. Clin Ling Phon, 19(6–7), 455–501.

Zharkova, N., & Hewlett, N. (2009). Measuring lingual coarticulation from midsagittal tongue contours: Description and example calculations using English /t/ and /a/. J Phon, 37(2), 248–256.

**Keywords: methodological research, distance measures, contours**