Tracking Larynx Movement in Real–Time MRI Data

Miran Oh[1], Asterios Toutios[2], Dani Byrd[1], Louis Goldstein[1], Shrikanth S. Narayanan[2]

1. Department of Linguistics, University of Southern California, USA; 2. Department of Electrical Engineering, University of Southern California, USA

miranoh@usc.edu   http://sail.usc.edu/span

Introduction

- Provide techniques for the quantification of larynx movement from dynamic vocal-tract Magnetic Resonance Imaging (MRI) data

- Data: USC's real-time MRI IPA database produced by phoneticians (Toutios et al., 2016)

- Methods: Principal Component Analysis (Bresch & Narayanan, 2009; Toutios et al., 2015) and Centroid Tracking Analysis

Principal Component Analysis

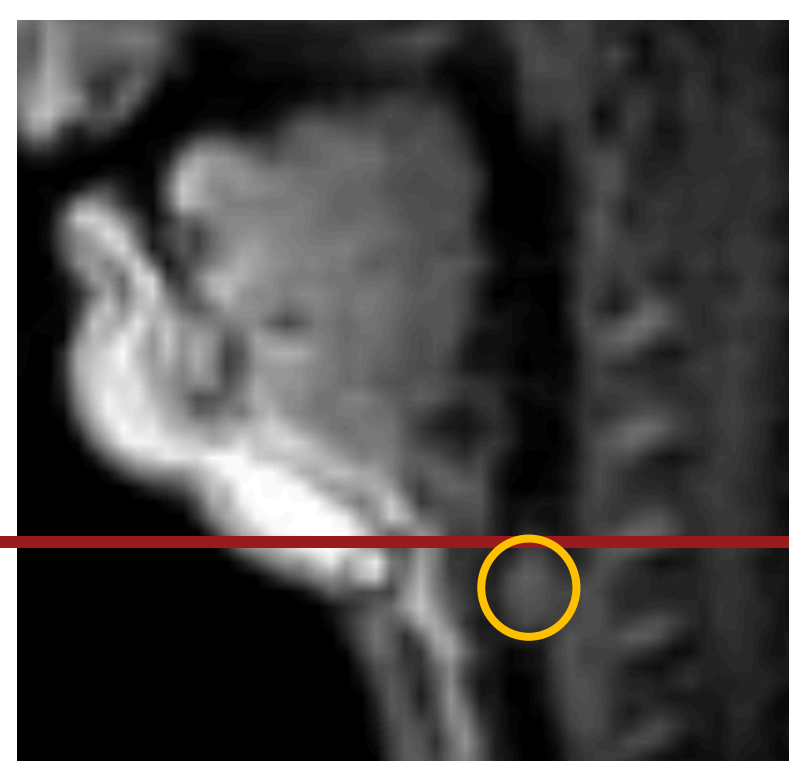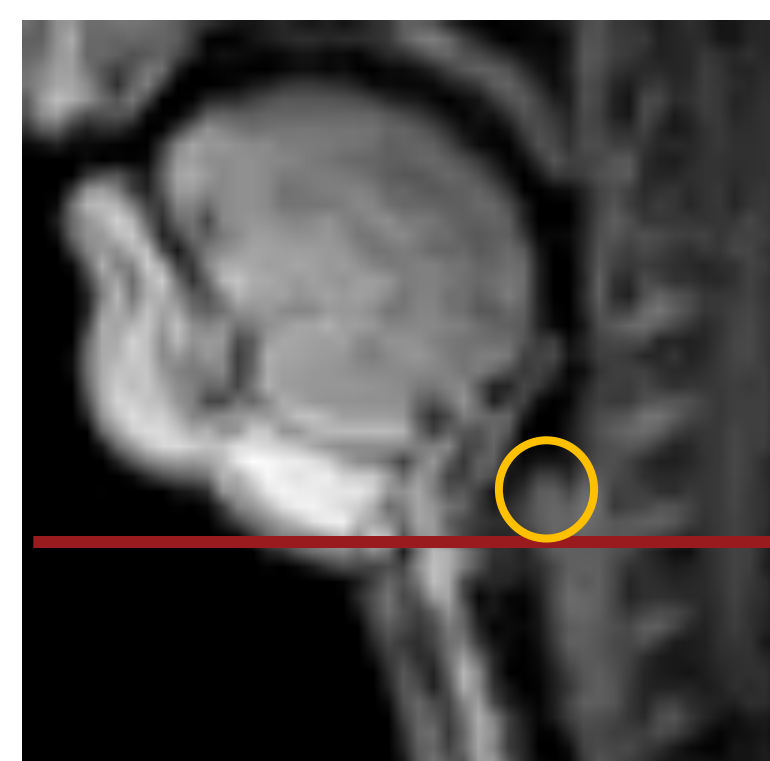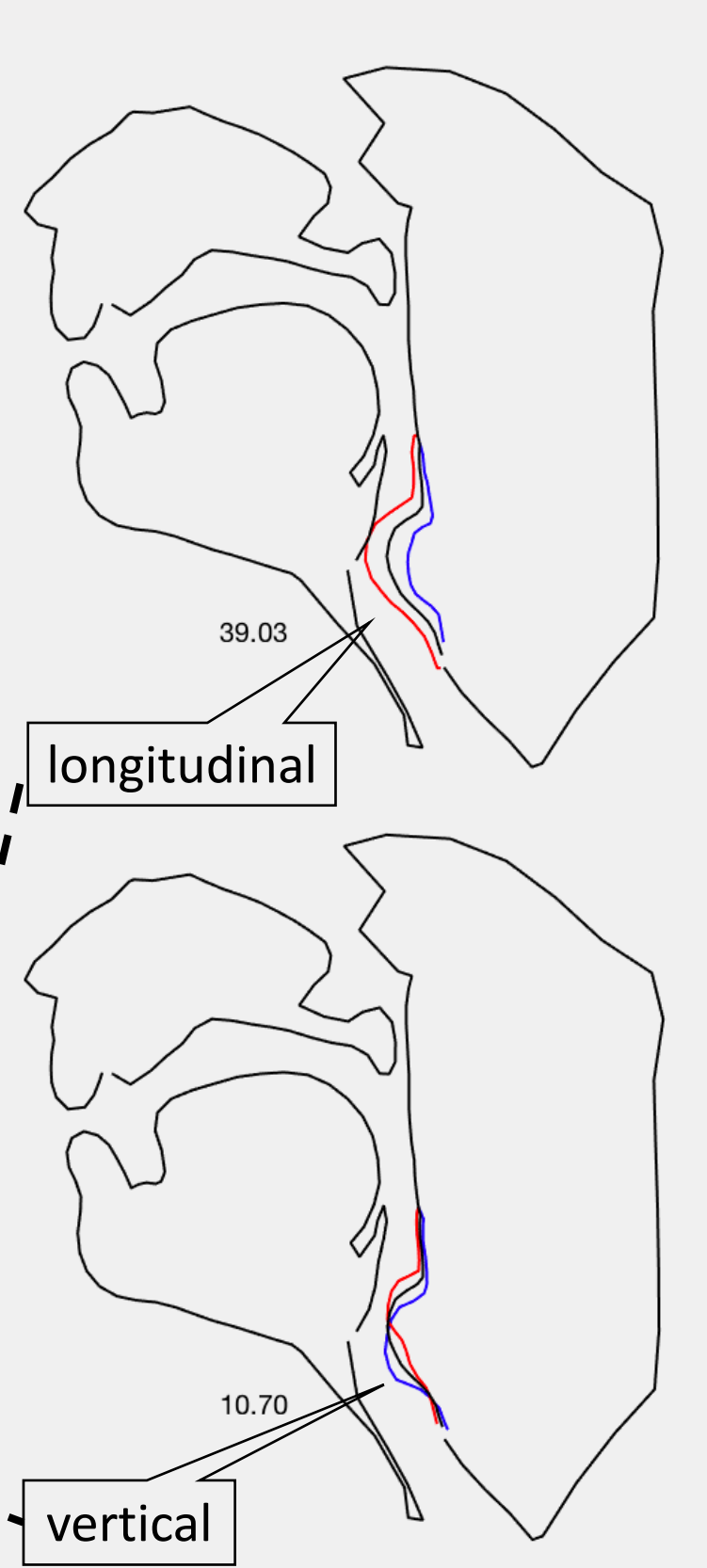### Articulator tracking

Larynx components
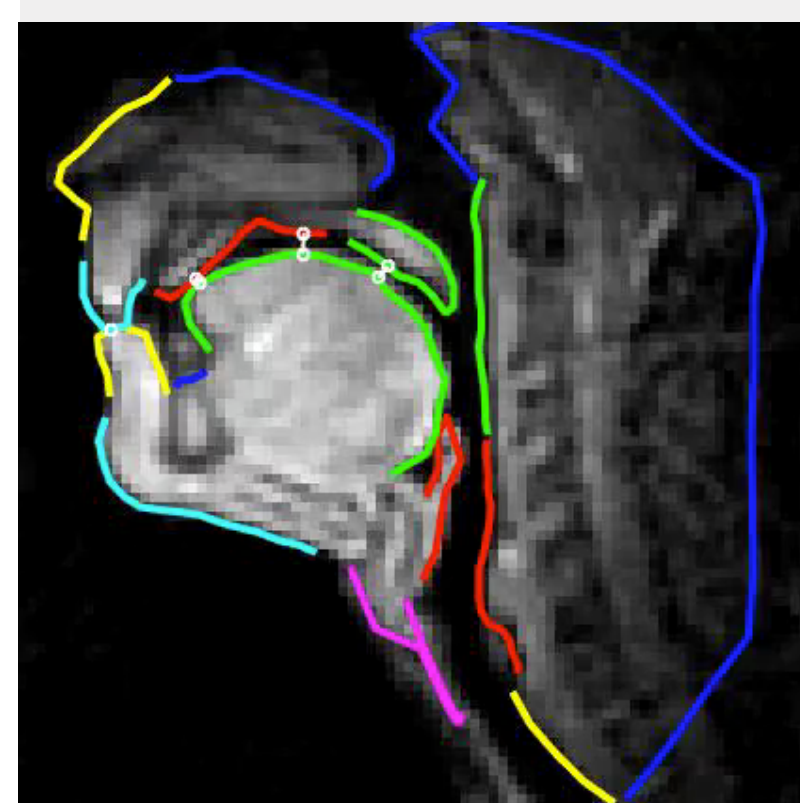
longitudinal

1) outline the vocal tract for each speaker
2) obtain gestural components and contour tracks from the outlines (semi-automatic)

Longitudinal larynx movement
voicing

Vertical larynx movement
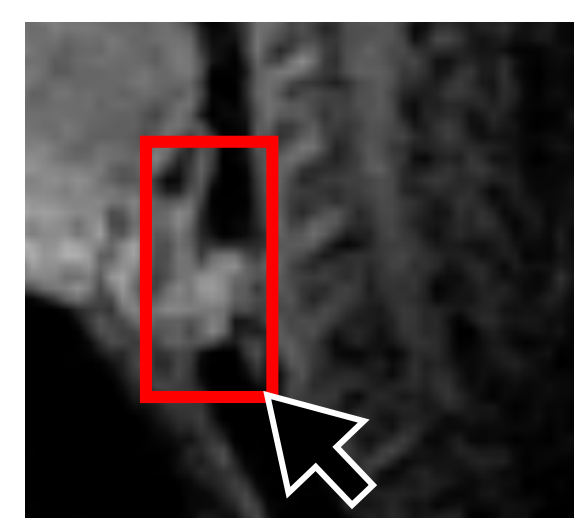ejective/implosives
pitch changes/tone

vertical

Centroid Tracking Analysis

Centroid:
Intensity-weighted average spatial position of an object
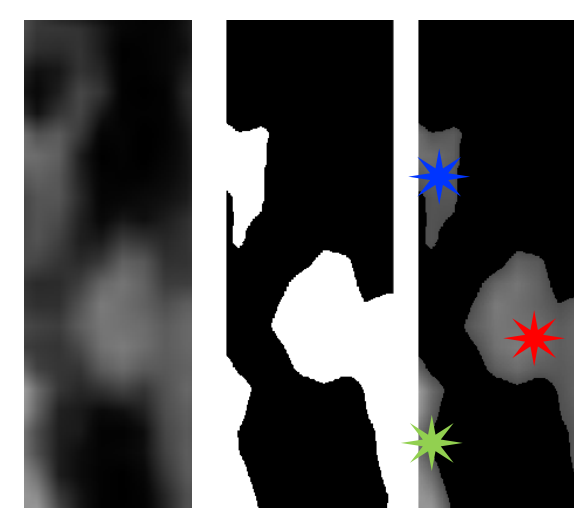
Pre-Processing Steps

Manual selection of a Region of Interest (ROI) for the larynx

(e.g., from the midline of the 2nd cervical vertebra [C-2] to the bottom line of C-4)
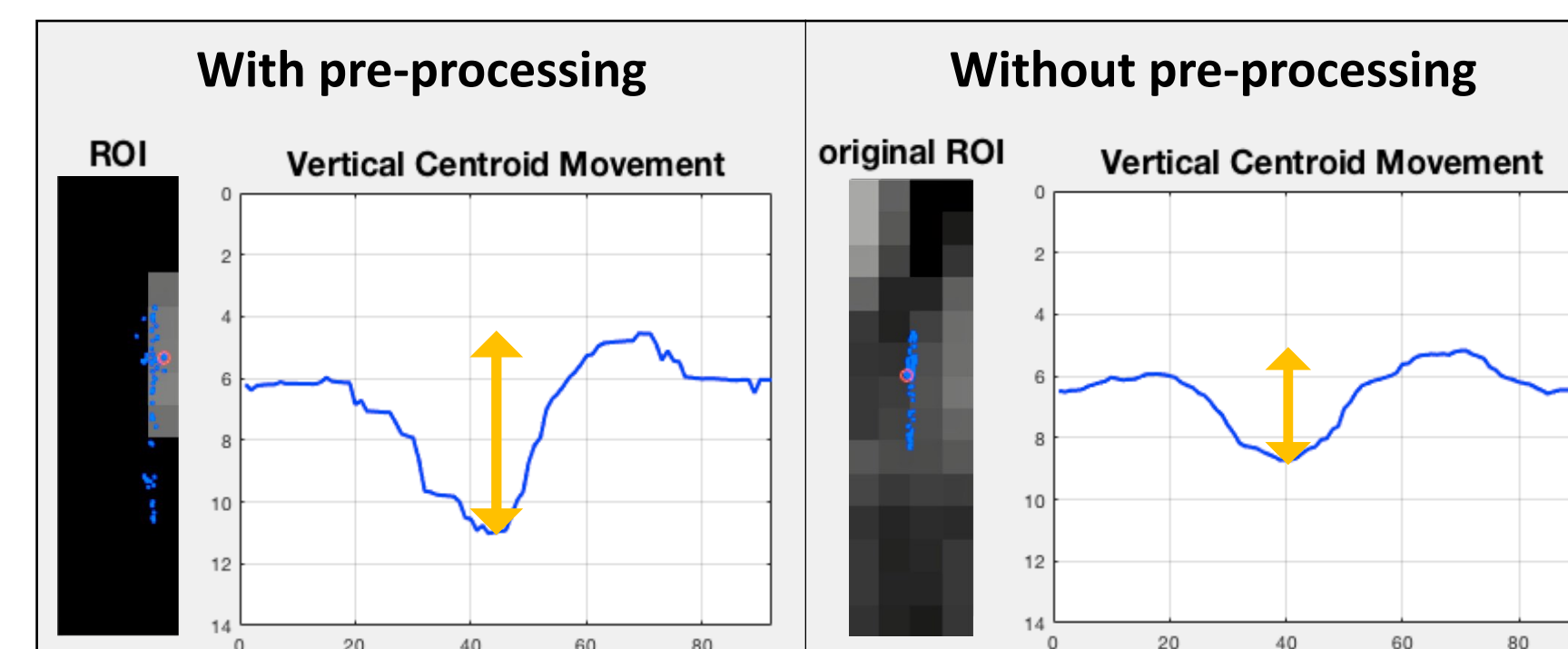

Seed-selection

to capture only the larynx object and neglect the other objects in the ROI (e.g., tongue root)

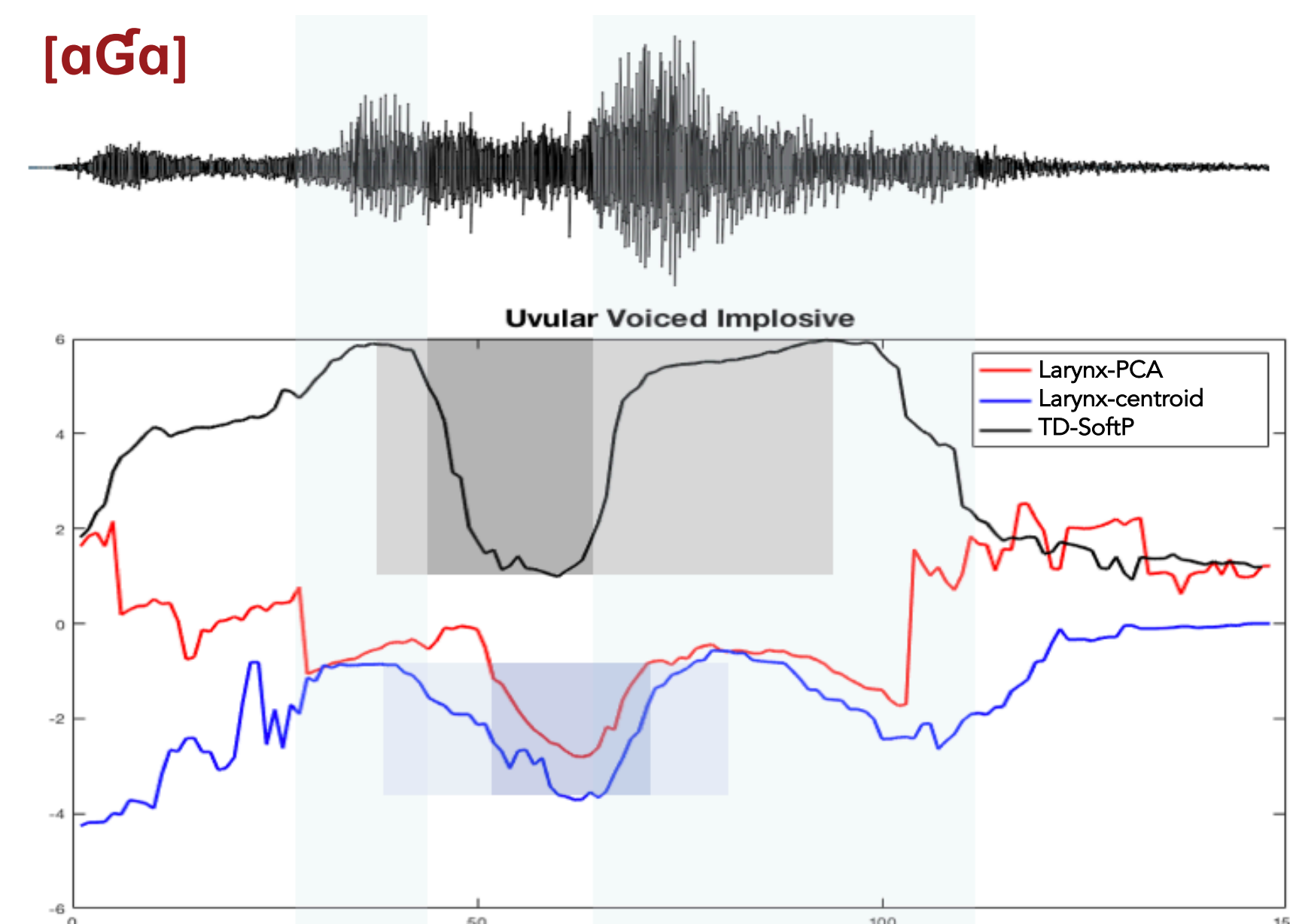Calculation of **the connected components (CC)** and **intensity-weighted centroids** for each CC

Original × CC = Centroid

The closest centroid from the previous frame's centroid is selected for each frame

With pre-processing — ROI — Vertical Centroid Movement

Without pre-processing — original ROI — Vertical Centroid Movement

Larynx Height Trajectories

[ɑɠɑ]

Uvular Voiced Implosive

Larynx-PCA
Larynx-centroid
TD-SoftP

Larynx height trajectories comparing **PCA** and **Centroid tracking analysis**, and the supralaryngeal constriction degrees (CD) of **tongue dorsum and soft palate** for uvular voiced implosive [ɑɠɑ]

- Centroid tracking analysis directly captures the spatial movement of the articulator of interest, unlike other ROI analyses (e.g., mean-pixel intensity analysis).

Conclusion

- The centroid tracking method out-performs PCA, as it is *much faster* and *completely automatic.*

- The centroid tracking analysis enables investigating the temporal coordination of the laryngeal raising/lowering gesture with the supralaryngeal constriction gestures.

- Moreover, this tool can be applied to velum raising/lowering and other articulatory movements by tracking vertical and horizontal centroids.

References

Bresch, E., & Narayanan, S. (2009). Region segmentation in the frequency domain applied to upper airway real-time magnetic resonance images. *IEEE transactions on medical imaging, 28*(3), 323-338. **Toutios, A., & Narayanan, S. S.** (2015, August). Factor analysis of vocaltract outlines derived from real-time magnetic resonance imaging data. In *International Congress of Phonetic Sciences (ICPhS), Glasgow, UK.* **Toutios, A., Lingala, S. G., Vaz, C., Kim, J., Esling, J., Keating, P. A., ... & Narayanan, S. S.** (2016). Illustrating the Production of the International Phonetic Alphabet Sounds Using Fast Real-Time Magnetic Resonance Imaging. In *INTERSPEECH* (pp. 2428-2432).