Collaborative Quantitative Analysis of RT-MRI

Nuno Almeida[1], Samuel Silva[1], António Teixeira[1], Conceição Cunha[2]

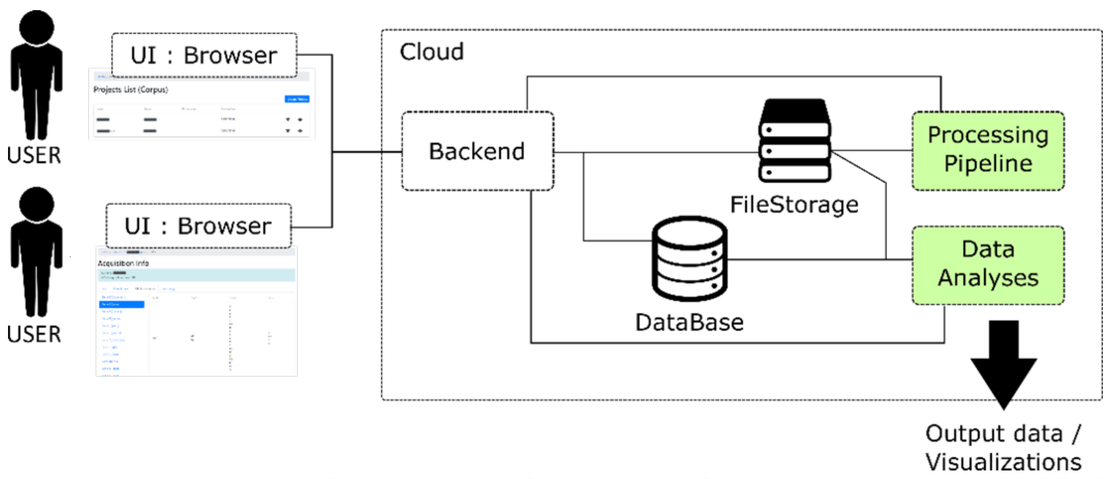
[1]Dep Electronics Telecom. & Informatics / IEETA, University of Aveiro, Portugal
[2]IPS, LMU Munich, Germany

The study of the dynamic aspects of speech production can profit from real-time magnetic resonance (RT-MRI).
For a non-specialist in image processing using RT-MRI data to test her/his hypothesis is not a simple task. While a few images might be manually processed and analyzed, it is far from profiting from the wide range of data available form recent RT-MRI technology, adding to the costs, complex setups and the higher demands to subjects, prevents processing of data from a large number of subjects.
The processing of RT-MRI acquisitions of many subjects at increasing frame rates (above 50Hz) is only possible by an automation and use of data-driven approaches, to avoid the time-consuming annotation, errors, and inconsistencies associated with manual processes.
Speech production researchers need tools to allow complex, state-of-the-art processing and analyzes, as they have for acoustical analyses.

Proposal

Processing of RT-MRI can be made accessible to non-specialists by tools that:
1) enable them to organize, annotatend process data from MRI (and other methods);
2) provide off-the-shelf processing components to create easily complex processing pipelines;
3) make possible collaboration and distribution of tasks by different researchers;
4) have ubiquitous access (anywhere, any computing platform).

Collaborative RT-MRI Processing Framework

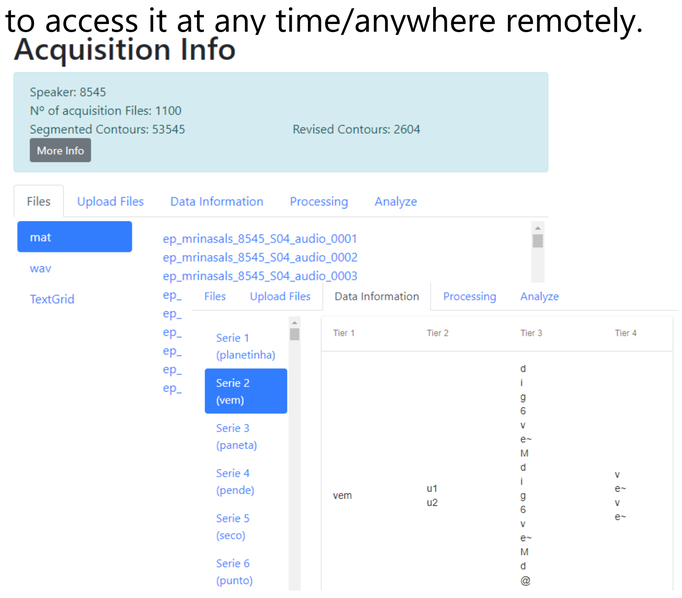


General Architecture of the Framework for Collaborative Quantitative Processing of RT-MRI

Backend - manages the interactions with users and system, can invoke the processing pipleline and data analyses
Processing pipeline - executes intensive operations in the users' data.
Data Analyses - Applies analysis methods to the data, stores results and provides visual interactive tools for data exploration and visualization
Framework - is cloud based, users to access it at any time/anywhere remotely.

Overall platform features

- Create a project
- Add an aquisition
- Upload data
- Segment images
- Revise segmented images
- Analyse data

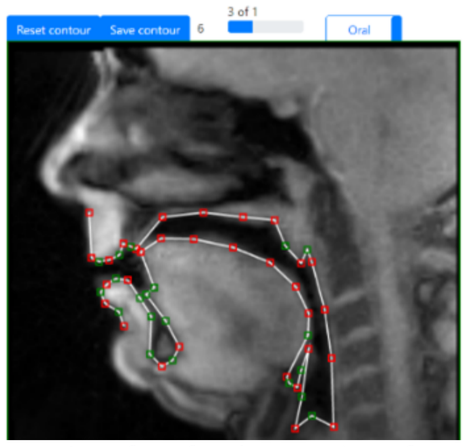


Data segmentation
- Current, vocal tract segmentation as in Silva and Teixeira 2015

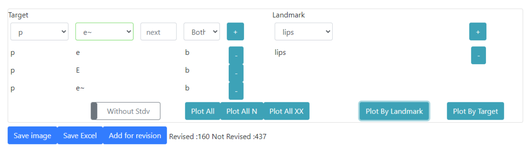
Collaborative data revision
- Current, tract segmentation with multi-annotator management

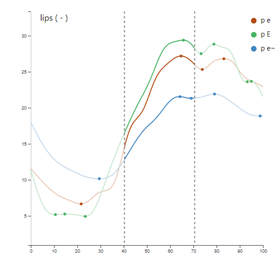
Data exploration and analysis
- Current tract variables over time for chosen sounds/contexts

Integration with external tools
- As in article Silva 2020

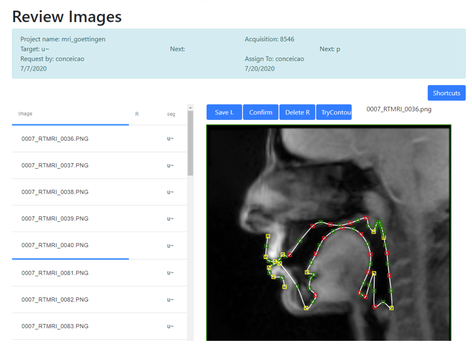


Interface for the creation of initial segmentations





Interface for the data analyses visualization



Interface for the revision

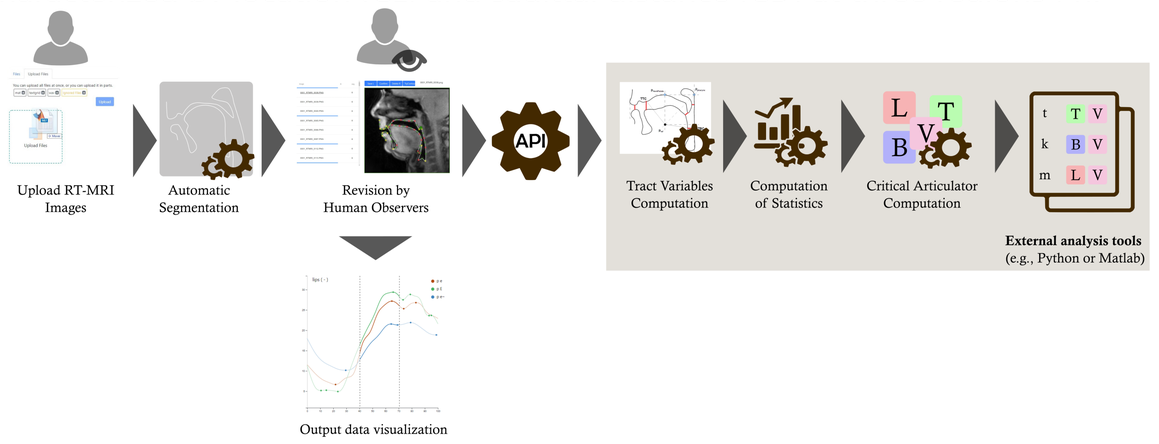
Reviews Requests List

Request By	Project (Acq)	Creation Date	Target	Prev/Next	Series	Assigned To		
conceicao	mri_goettingen (8546)	7/6/2020	p	/ 6	0..21	user3		REVIEW
conceicao	mri_goettingen (8546)	7/6/2020	p	/ 6	50..300	user3		REVIEW
conceicao	mri_goettingen (8546)	7/6/2020	a	p /	0..175	conceicao		REVIEW
conceicao	mri_goettingen (8546)	7/6/2020	p	/ a	0..175	user4		REVIEW
conceicao	mri_goettingen (8546)	7/6/2020	e	p /	0..300	user3		REVIEW

Interface for the request of revision assignment

Pipeline

It constitutes the main processing component of the architecture. The vocal tract outlines are extracted adopting the method proposed by Silva et al. [3], resulting in contours identifying the different regions of interest. The comparison among vocal tract configurations was performed adopting and extending a previously proposed framework [3] enabling normalized quantification of differences, for different articulators/regions of the vocal tract and their visualization: velum (VEL), tongue dorsum (TD), tongue back (TB), tongue tip (TT), lip protrusion (LP), lip aperture (LA) and pharynx (Ph). Additionally, the work presented here also considered constrictions, characterized by location (CL) and sagittal distance (CD) at three regions [4].



Conclusion

The need for a new generation of tools to process increasing amounts of RT-MRI data by speech production researchers in general is defended and a Framework is proposed. Concrete examples of application of the framework in Phonetic research will be presented at the Symposium (if submission accepted). Examples will include the extension to a larger set of speakers of authors' recent study on the role of oral configurations in European Portuguese nasal vowels [4].

References

[1] C. Lammert, M. I. Proctor and S. S. Narayanan, "Data-driven analysis of realtime vocal tract MRI using correlated image regions," in InterSpeech, 2010.
[2] S. Silva and A. Teixeira, "Unsupervised segmentation of the vocal tract from real-time MRI sequences," Computer Speech and Language, p. 25–46, 2015.
[3] S. Silva and A. Teixeira, "Quantitative systematic analysis of vocal tract data," Computer Speech and Language, vol. 36, pp. 307-329, 2016.
[4] Cunha, S. Silva, A. Teixeira, C. Oliveira, P. Martins, A. Joseph and J. Frahm, "On the role of oral configurations in European Portuguese nasal vowels," in InterSpeech, Graz, 2019.
[5] S. Silva, N. Almeida, C. Cunha, A. Joseph, J. Frahm, A. Teixeira. Data-Driven Critical Tract Variable Determination for European Portuguese. Information 2020, 11, 491.

Acknowledgments: This work is partially funded by IEETA Research Unit funding (UID/CEC/00127/2019), by Portugal 2020 under the Competitiveness and Internationalization Operational Program, and the European Regional Development Fund through project SOCA – Smart Open Campus (CENTRO-01-0145-FEDER-000010), project MEMNON (POCI-01-0145-FEDER-028976) and by the German Federal Ministry of Education and Research (BMBF, with the project 'Synchonic variability and change in European Portuguese').