GFM-Voc

A real-time voice quality modification system

PERROTIN Olivier, GIPSA-lab, *olivier.perrotin@gipsa-lab.fr*
MCLOUGHLIN Ian, University of Kent, *I.V.McLouglin@kent.ac.uk*

Towards expressive speech synthesis

- Speech expressivity partly conveyed by **glottis-related timbre variations** (e.g., effort, tenseness) called voice quality [2]
- Need for a full control of these variations over time in speech synthesis (e.g., voice substitution applications)

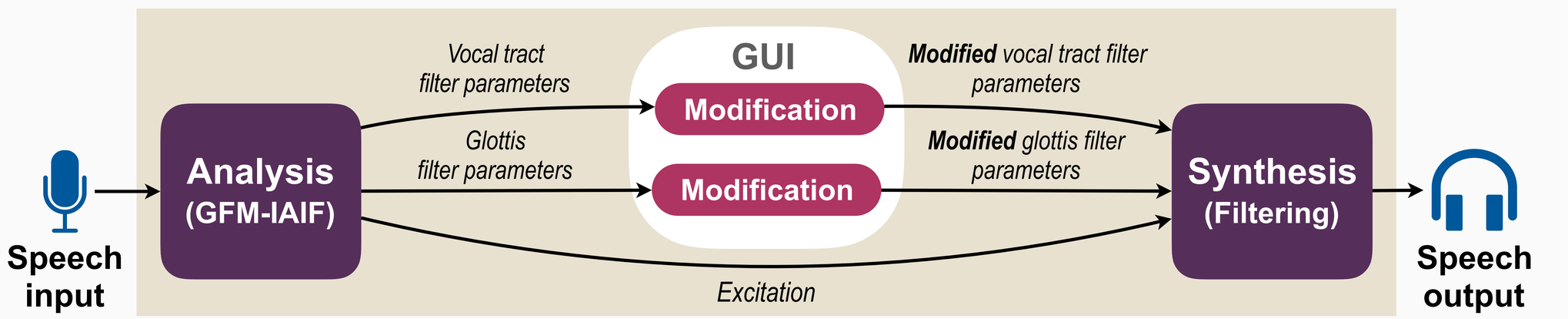
➤ Real-time modification of voice quality parameters with the **Glottal Flow Model-based Vocoder (GFM-Voc)**

Spectral envelope modification

- Existing systems (e.g., *Audapter* [3]) modify the full signal spectral (LP) envelope
- Need to disentangle vocal tract and glottis spectral envelopes

➤ Real-time implementation of the **GFM-IAIF** glottal inverse filtering method [1]

GFM-Voc framework

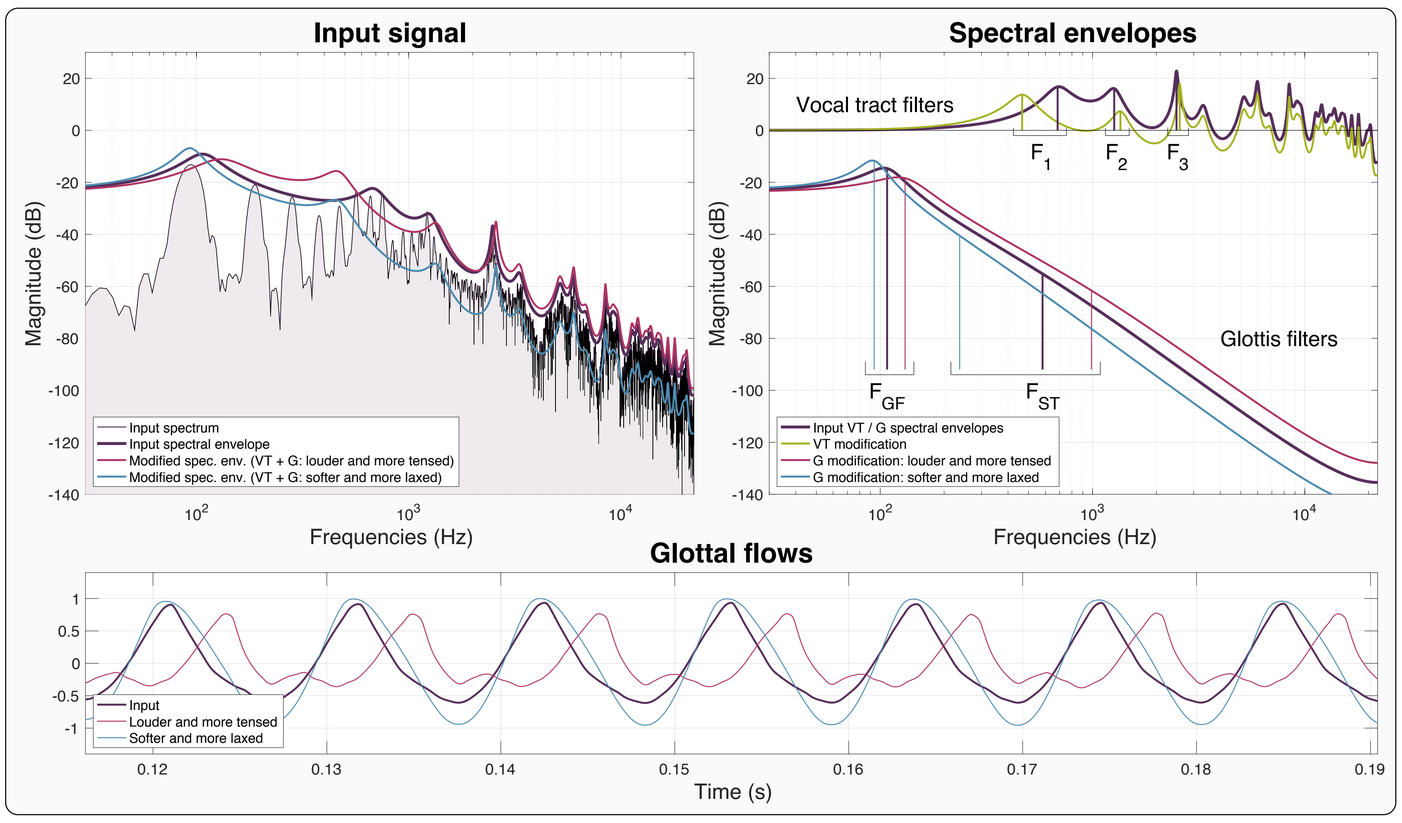


Implementation details

- Matlab Audiotoolbox
- Standalone C-based application to come

GFM-IAIF code here

Modification of filter parameters



- Extraction of vocal tract (VT) and glottis (G) spectral envelopes with GFM-IAIF [1]
- Modification of filter parameters using analogue filters [4]
- Mapping from effort / tenseness to glottis parameters [5]

Acoustic correlates

	Vocal Tract				*Glottis*	
Jaw	Closing	⟷	Opening	**Effort**	Softer	⟷ Louder
F_1	Lower		Higher	F_{GF}	Lower	Higher
Tongue	Backwards	⟷	Forwards	F_{ST}	Lower	Higher
F_2	Lower		Higher	**Tenseness**	More laxed	⟷ More tensed
Lip	Rounding	⟷	Stretching	F_{GF}	Lower	Higher
F_3	Lower		Higher	Q_{GF}	Narrower	Wider

User Interface

Conclusions

➤ Real-time voice quality modification system

➤ **First system** with independent modification of vocal tract and glottis parameters (voice quality)

➤ **No loss of information** ☛ high-quality modification

➤ Applications to expressive speech synthesis, auditory feedback perturbation, and speech therapy

References

[1] O. Perrotin and I. V. McLoughlin (2019), “A spectral glottal flow model for source-filter separation of speech,” in *IEEE ICASSP,* Brighton, UK, May 12-17, pp. 7160–7164.
[2] B. Doval, C. d’Alessandro, and N. Henrich (2006), “The spectrum of glottal flow models,” *Acta Acustica united with Acustica*, 92(6), pp. 1026–1046.
[3] J. A. Tourville, S. Cai, and F. Guenther (2013), “Exploring auditory-motor interactions in normal and disordered speech,” in *Proc. of Meetings on Acoustics*, Montreal, Canada, June 2-7, pp. 1–8.
[4] R. Bristow-Johnson (2001). Audio-eq-cookbook [online]. *https://music.columbia.edu/pipermail/music-dsp/2001-March/041752.html*
[5] L. Feugère, C. d’Alessandro, B. Doval, and O. Perrotin (2017), “Cantor digitalis: Chironomic parametric synthesis of singing,” *EURASIP Journal on Audio, Speech, and Music Processing*, 2017(2).

Grenoble Images Parole Signal Automatique
UMR CNRS 5216 – Grenoble Campus
38400 Saint Martin d’Hères - FRANCE