Development of a deep learning approach based glottal flow model using high-fidelity numerical simulations on universal vocal fold kinematics models

Yang Zhang[1], Xudong Zheng[1], Qian Xue[1]*, Simeon Smith[2], Ingo Titze[2]

[1]Department of Mechanical Engineering, University of Maine, Orono, ME, USA
[2]The National Center for Voice and Speech, University of Utah, Salt Lake City, UT, USA

Keywords: Glottal Flow; Machine Learning; Deep Neural Network; Reduced-order Modeling

Introduction

An accurate prediction of vocal fold vibration and sound source relies on an accurate prediction of intraglottal pressure distribution and glottal flow rate in flow-structure interactions (FSI). Current computer models of glottal flow during voice production is based on either Navier-Stokes equation, which can provide accurate and detailed flow information but requires extremely high computational cost, or Bernoulli equation, which is simple and fast but has low accuracy. In this research, we aim to develop a deep-learning based glottal flow model that can provide fast and accurate predictions of the intraglottal pressure distribution and flow rate in any glottal shapes during vocal fold vibration. The flow model will be further coupled with a continuum vocal fold model to realize nearly real-time FSI simulations with high accuracy.

Methods

The underlying assumption of the approach is that the vocal fold kinematics can be approximated by a few vibration modes as described by the surface –wave approach in [1]. For normal phonation, the vocal fold vibration is dominated by two modes. The displacement of the glottis over time can be represented by a linear combination of the modal displacement of these two modes as follows:

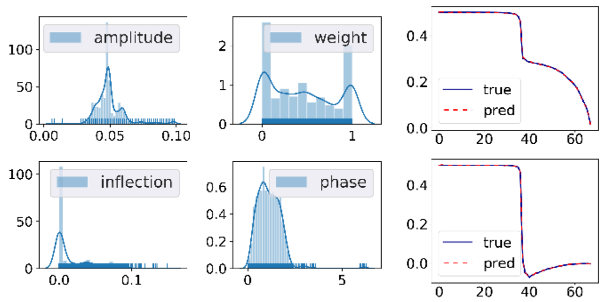
$$\xi(\mathrm{y},\mathrm{z},\mathrm{t}) = \xi_{\mathrm{m}} \sin\left(\frac{\pi z}{L}\right)\left[sin\omega t cos\frac{\omega(y-y_m)}{c} - (1-\alpha) cos\omega t sin\frac{\omega(y-y_m)}{c}\right] \quad (1)$$

where $\xi_m, \alpha, \beta, \phi$ are the governing parameters to be determined for each shape. A generalized glottal shape library is built through systematic variations of these governing parameters in broad ranges. For each shape in the library, the groundtruth values of the flow rate and pressure distribution are obtained by solving the Navier-Stokes equation. Then, the mapping relationship between the governing parameters (input features) and the corresponding groundtruth values of the flow rate and pressure distribution (output targets) can be established by a fully connected deep neural network (DNN). With this trained DNN, the flow rate and pressure distribution along any glottis shape generated by Eq. (1) can be predicted.

Results and Discussion

To verify that Eq. (1) can be used as a generalized equation to represent any glottal shape during normal phonation, FSI simulations of a three-layer vocal fold model under different subglottal pressures and material properties are conducted. Eq (1) is then used to fit the obtained glottal shapes. The results show a very good agreement. *Figure 1(a)* shows the probability density function (PDF) of the fitted governing parameters. The DNN flow model is then used to predict the pressure distribution and flow rate in these glottal shapes. The prediction is compared with the truth values obtained from the Navier-Stokes flow solver. The preliminary results show good prediction accuracy. *Figure 1(b)* shows the comparison of the pressure distribution in two typical convergent and divergent glottal shapes.

The coupling of the DNN flow model with a continuum vocal fold model is currently untaken. The accuracy of this FSI model in predicting vocal fold vibrations and glottal flow waveform will be compared with Navier-Stokes equation based high-fidelity FSI simulations.



(a) (b)

Figure 1: (a) PDF of fitted governing parameters; (b) Performance of DNN pressure prediction

Acknowledgements

The project was supported by Grant Number 5R21DC016428 from the National Institute on Deafness and Other Communication Disorders (NIDCD).

References

[1] Simeon Smith and Ingo Titze. Journal of Biomechanics, 73:177–184,2018.

**corresponding author e-mail: qian.xue@maine.edu*