New features in VTAR: A Matlab-based computer program for vocal tract acoustic modeling

INTRODUCTION
VTAR [1] is a Matlab-based computer program for vocal tract acoustic response calculation based on a frequency-domain vocal tract model [2]. It is able to model various sounds such as vowels, nasals, nasalized sounds, and liquids with area functions as input. It calculates the vocal tract acoustic response function, the formant frequencies and bandwidths. The user-friendly interface allows directed data input for defined categories: vowels, nasals, nasalized sounds, consonants, laterals, and rhotics. The program also provides an interface for input and modification of arbitrary vocal tract geometry configurations, which is ideal for research applications.

Several new features are included in VTAR 2.0, the latest version, which include: (1) acoustic sensitivity function calculation for formants, (2) area function modification for targeted formant pattern, (3) susceptibility plot calculation, which is useful particularly for nasalized vowel analysis, (4) speech sound synthesis with source model options, and (5) addition of a new set of area function data for liquid sounds extracted from MR (Magnetic Resonance) images. These new features along with the user-friendly interface significantly enhance the usability of VTAR for both teaching and research purposes.

MODELING METHOD

Cross-sectional area functions

- Frequency-domain formulation based on a transmission-line model and the chain matrix.
- Vocal tract modeled as a concatenation of various modules (such as single tube, branching, and lateral channels).
- The vocal tract transfer function defined as the area function at the lips divided by the volume velocity at the glottis.

MAIN FEATURES IN FIRST VERSION

- Acoustic response calculation for different kinds of sound and different vocal tract configurations.
- Formant and bandwidth calculation from acoustic response.
- Generic area function for different kinds of sounds.
- Arbitrary tube configuration allowed for vocal tract model.
- Fluid and wall property setting.
- User-friendly interface for area function input and manipulation.

NEW FEATURE 1: Acoustic sensitivity function calculation for formants

The acoustic sensitivity function for each formant shows how much that formant will change with the area function perturbation at different locations [3]. It is defined as the difference between the kinetic energy (KE) and the potential energy (PE) as a function of distance starting from the glottis, divided by the total energy (TE) (sum of kinetic and potential energy) in the system. As the sensitivity function for F2 shows in Fig. 2, mainly F2 will be changed by a decrease in the area of the lip opening. This result is reflected in the acoustic spectra shown in Fig. 3.

NEW FEATURE 2: Automatic area function modification based on targeted formant pattern

Modification of the vocal tract area function based on targeted formants is performed by using sum and difference combinations of acoustic sensitivity functions of formants to modify the initial vocal tract area function [4].

NEW FEATURE 3: Susceptance plot for vocal tract analysis

Susceptance plot of a nasal /m/

looking from the coupling location $B_{n}$: the susceptance of the pharyngeal cavity $B_{c}$: the susceptance of the oral cavity $B_{n}$: the susceptance of the nasal cavity

Poles of a nasal or nasalized vowel should occur at frequencies at which $B_{n}$ and $(B_{c}+B_{n})$. Thus, the frequencies of the poles can easily be seen from the intersections of the plots of $B_{n}$ and $(B_{c}+B_{n})$.

NEW FEATURE 4: Sound synthesis with source model options (Liljencrants-Fant model[5], Rosenberg model[6], and white noise model)

NEW FEATURE 5: A new set of area function data for rhotic sound from MR images in a database with 21 subjects [7]

FREE DOWNLOAD AVAILABLE
http://www.isr.umd.edu/labs/SCL/vtar

REFERENCES

ACKNOWLEDGMENTS
This work was supported by NIH Grant 1 R01 DC05250-01

http://www.isr.umd.edu/Labs/SCL/V TAR