Human Communication Technologies

ULTRA-ARTI-SYNTH

Articulatory Vowel Synthesis from Ultrasound Tongue

Interdisciplinary Speech Research Lab

Pramit Saha[1], Yadong Liu[2], Bryan Gick[2,3], Sidney Fels[1]
[1]Electrical and Computer Engineering Department, University of British Columbia, Canada
[2]Department of Linguistics, University of British Columbia, Canada
[3]Haskins Laboratories, New Haven, Connecticut, USA

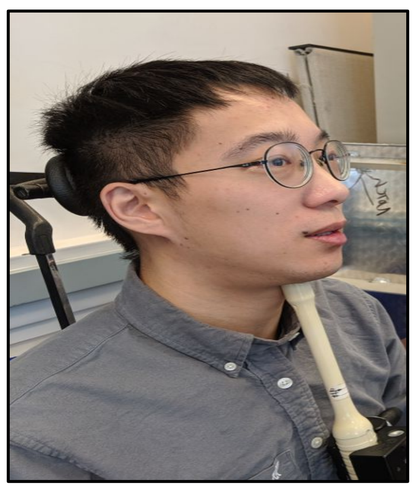
INTRODUCTION

- Ultrasound-based silent speech interface
- Area-function based articulatory speech synthesis
- Articulatory-to-acoustic forward mapping
- Potential to help post-laryngectomy patients
- Using Vocal Tract Model in ArtiSynth

DATA COLLECTION

- Midsagittal tongue ultrasound Video of single male participant
- Participant seated with head stabilized against a headset
- Ultrasound transducer placed beneath chin
- Participant asked to make open Vocal tract sounds (vowels)

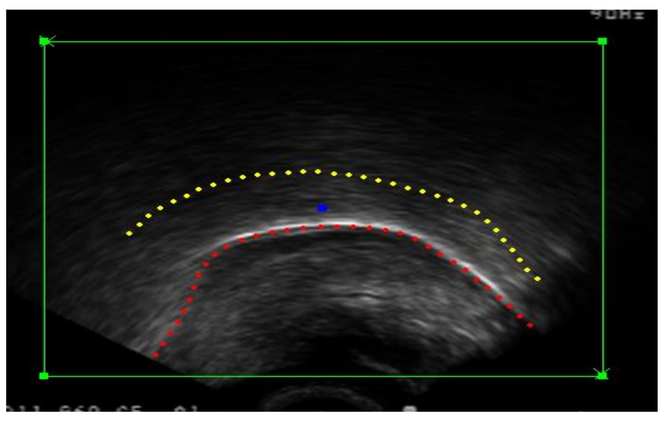
METHODOLOGY



Mid-sagittal ultrasound data acquisition

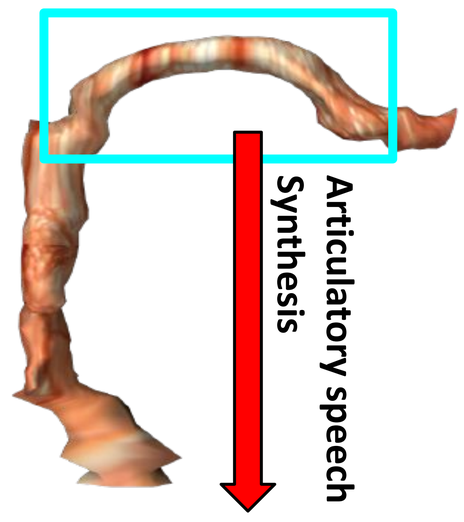
Palatal tracing

Tongue Tracking with SLURP

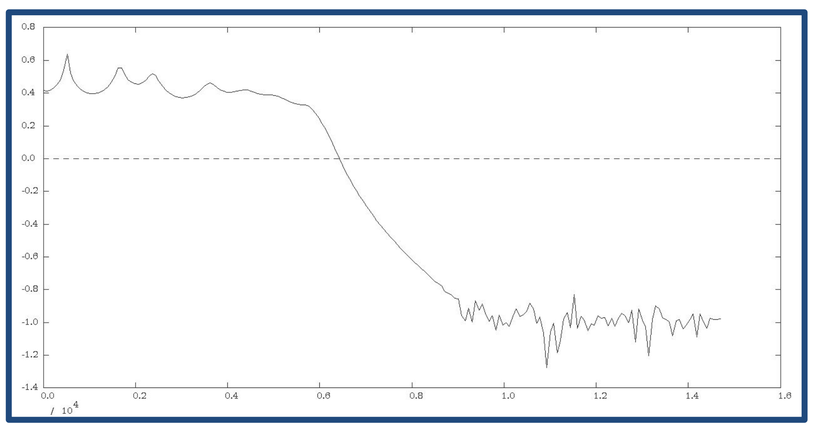


Single frame of ultrasound video

Vocal Tract C/S Area computation



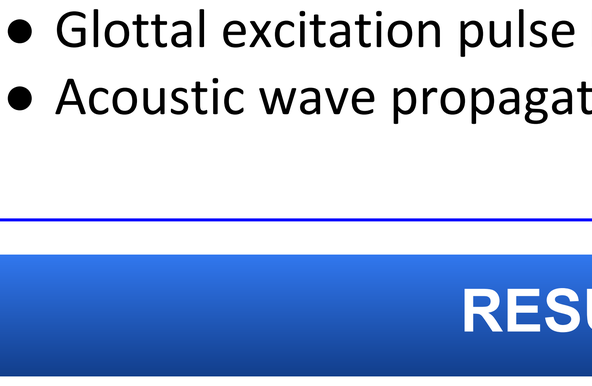
- Semi-automatic tongue tracking using active-contour segmentation
- Cross-sectional area computation with semi-polar grid strategy
- Use of area-functional values only between tongue and palate
- Other area-functional values kept constant at .8 sq cm
- Glottal excitation pulse based on Rosenberg's model
- Acoustic wave propagation using pressure-velocity wave equations

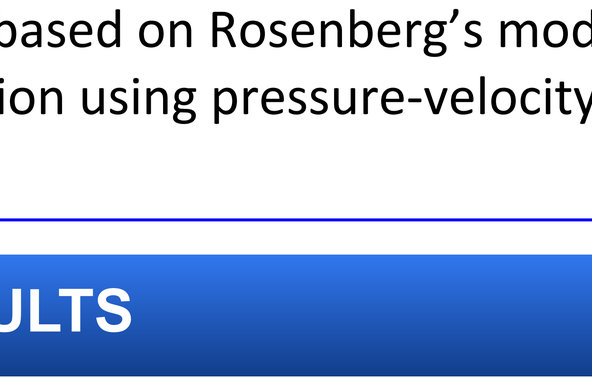


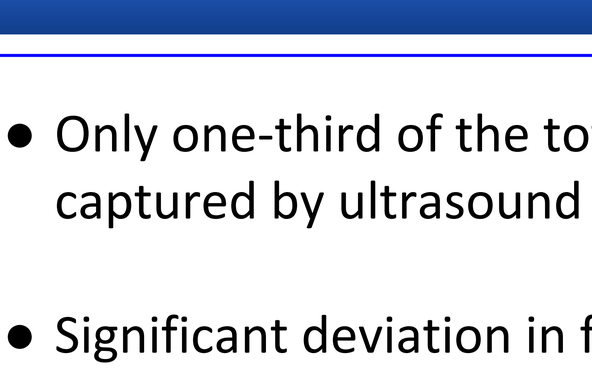
Synthesized speech in frequency domain

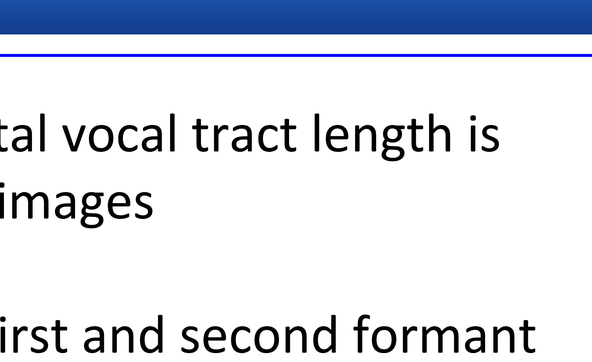
RESULTS

- Only one-third of the total vocal tract length is captured by ultrasound images
- Significant deviation in first and second formant frequencies of the vowels









DISCUSSION

- Incapable of capturing the entire vocal tract geometry like lips, subglottal tract, etc
- Achievable in real-time unlike MRI, CT etc
- Need automatic tongue tracking for real-time implementation
- Need to automatically generate the incomplete vocal tract geometry
- Synthesized speech sounds instrumental. Better articulatory speech synthesis engine will lead to high quality sounds
- New direction in silent speech interfaces

NSERC CRSNG CIHR IRSC Canadian Institutes of Health Research / Instituts de recherche en santé du Canada