

SGER: Exploring Emotional Vocal Productions through the use of Real-time Magnetic Resonance Imaging

Shrikanth Narayanan, University of Southern California

Humans use intricate vocal and visual orchestrations to encode and communicate intent and emotions. Understanding and utilizing these expressive emotional elements hence is key to facilitating any human experience, including for developing human-centered communication technologies. The vocal instrument is central to the expressive human communication capability. While its use in speaking, singing and other forms of vocalizations have been studied well, the actual expressive mechanisms of vocal productions are less completely understood. For example, what vocal tract mechanisms do humans use to produce speech sounds conveying anger versus happiness? How do singers produce different sounds with different emotions? New technology tools, such as fast magnetic resonance imaging, combined with novel computational capabilities, such as statistical machine learning, offer ways for gaining insights into, and measuring and modeling, these processes in ways that were not possible before.

The project focuses on investigating the articulatory vocal production mechanisms of expressive human communication. It has two specific near term goals. The first goal aims at experimental data collection of emotional speech production using real-time magnetic resonance imaging. The goal focuses on pilot analysis of the collected data, both image and audio, and provide insights into the emotional modulation of speech mechanisms and its consequences on the audio signal. The work will provide the necessary foundation for a detailed research study on emotional human speech production.

The intellectual merit of the project lies in the use of novel methods for examining and modeling emotional speech production. It aims to discover details of how the human vocal process is modulated to encode emotional expressions and how such knowledge can be incorporated in the design of emotional speech processing and generation by the machine. Expressive aspects of information however have been largely ignored in the technology realm with the primary focus thus far put on content than style; human machine interfaces are limited in their emotional cognizance. The project on emotional human speech promises new directions in human language communication research and its applications.

The interdisciplinary approach to the problem leads to its broad impact along several dimensions including the innovative experimental and computational approaches that can impact several disciplines including engineering, computer science, psychology and linguistics, the use of the project as a vehicle for graduate and undergraduate training, and the dissemination of novel imaging data that is hitherto not available in the scientific community.