USC UNIVERSITY OF SOUTHERN CALIFORNIA | IMSC Integrated Media Systems Center

Expressive Speech Synthesis

Murtaza Bulut Shrikanth S. Narayanan

Research Goals

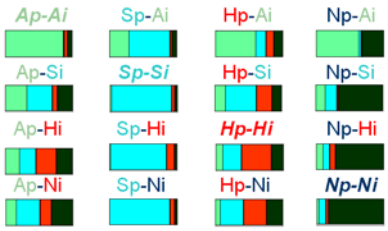
>> To build a speech synthesizer that will imitate the expressive human speech.

>> To develop emotion conversion algorithms that will enable transformation of one emotion into another.

>> To build a large database of natural emotional speech.

Accomplishments

>> Basic emotions synthesized with high recognition accuracy: Anger 86.1%, Sadness 89.1%, Happiness 44.2%, Neutral 81.8%.



>> Large emotional inventory has been collected (2 female and 3 male speakers)

>> Paper *"Expressive Speech Synthesis Using a Concatenative Synthesizer"* presented in ICSLP 2002.

Role in IMSC

>> Development of the natural human-machine communication systems based on speech.

>> Contributing to the improvement of the systems involving human-machine interactions.

Research Approach

>> Data driven approach using concatenative speech synthesis methods and waveform modification techniques.

>> Emotional inventories are recorded for each of the target emotions: Anger, Happiness, Sadness, Frustration and Neutral.

>> Prosodic and spectral modification is performed.

>> Basic Software: Festival, HMM-based HTS

>> Main Methods: TD-PSOLA, LP-PSOLA, HNM

>> GMM based emotion transformation.

Uniqueness & Related Work

>> Modification and concatenation of emotional speech units.

>> Development of emotion specific analysis techniques.

>> Parameter driven approaches: Formant & Articulatory synthesis

>> Emotion Recognition

Five-year Plan

>> 2002-2004: Creation of preliminary database and initial analysis and synthesis for canonical emotions.

>> 2004-2006: Development and testing of algorithms for expressive speech synthesis and transformation. Updated databases with new context and additional emotions.

>> 2006-2008: Implementation and demonstration in human-machine interaction systems for a range of emotions.