

Comparison of Acoustic Parameters and MFCCs for Speaker Identification

Alec Colvin

Srikanth Vishnubhotla

Daniel Garcia-Romero

Advisor : *Prof. Carol Y Espy-Wilson*

Contents

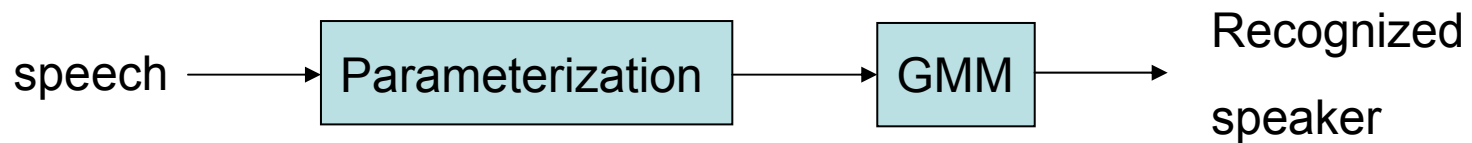
Feature Extraction

- MFCCs
- Voice Quality
- Formants

Motivation and Design

Results & Future Work

Feature Extraction in a Nutshell



- MFCCs (state of the art)
- Knowledge Based Acoustic Parameters (APs)
 - Glottal source
 - Vocal tract

Contents

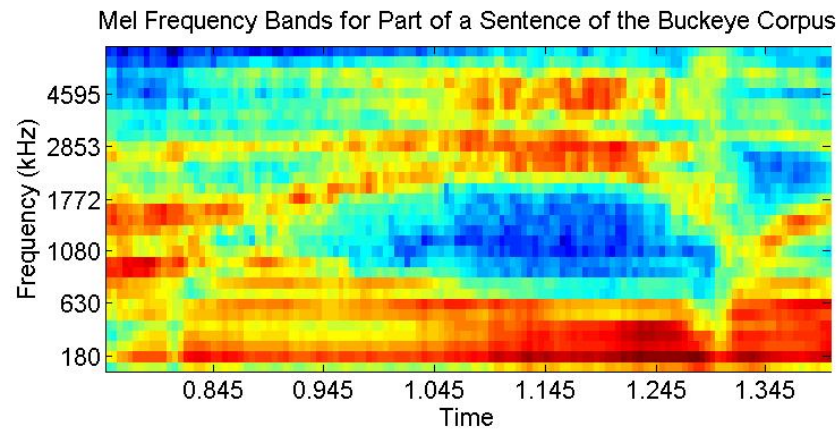
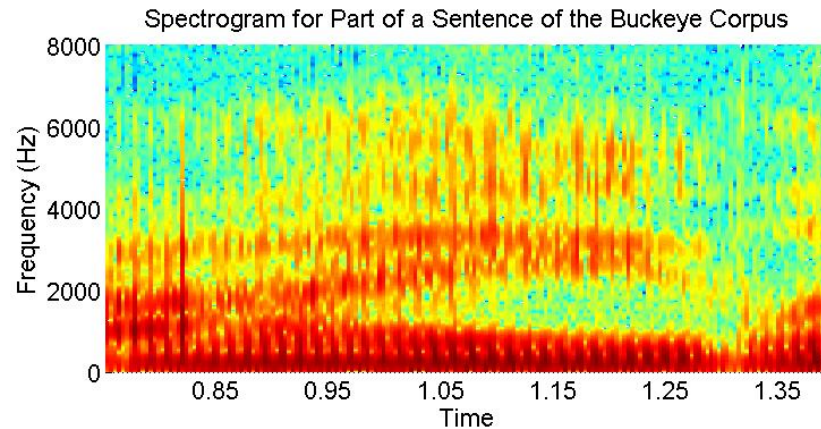
Feature Extraction

- MFCCs
- Voice Quality
- Formants

Motivation and Design

Results & Future Work

MFCCs: Mel Frequency Cepstral Coefficients



MFCCs quantify the relative energy in different frequency bands and project it into a DCT basis.

Above: a comparison of the energy distribution in the wide-band spectrogram and Mel-frequency bands.

Contents

Feature Extraction

- MFCCs
- Voice Quality
- Formants

Motivation and Design

Results & Future Work

Voice Quality: H1-H2 and Slope

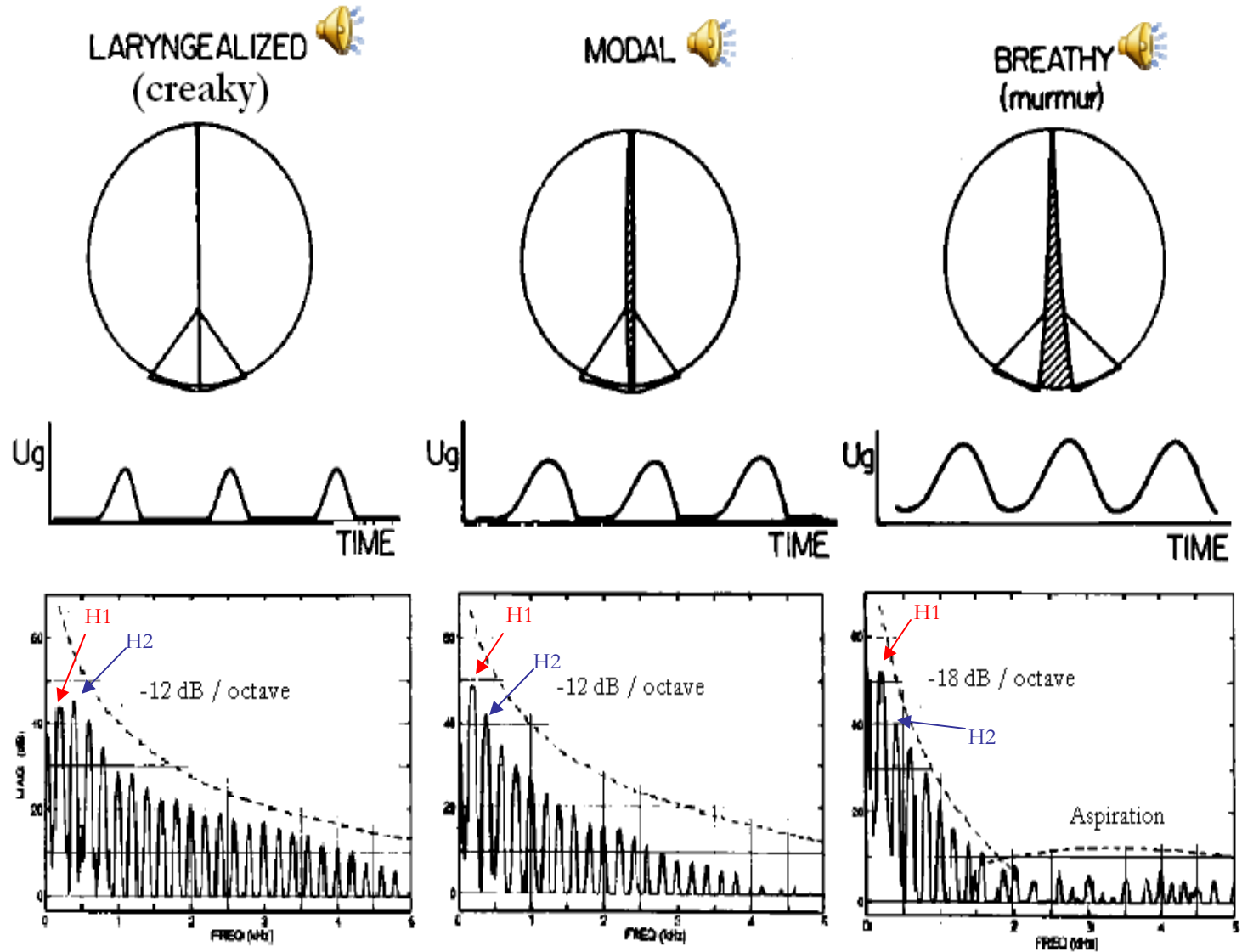


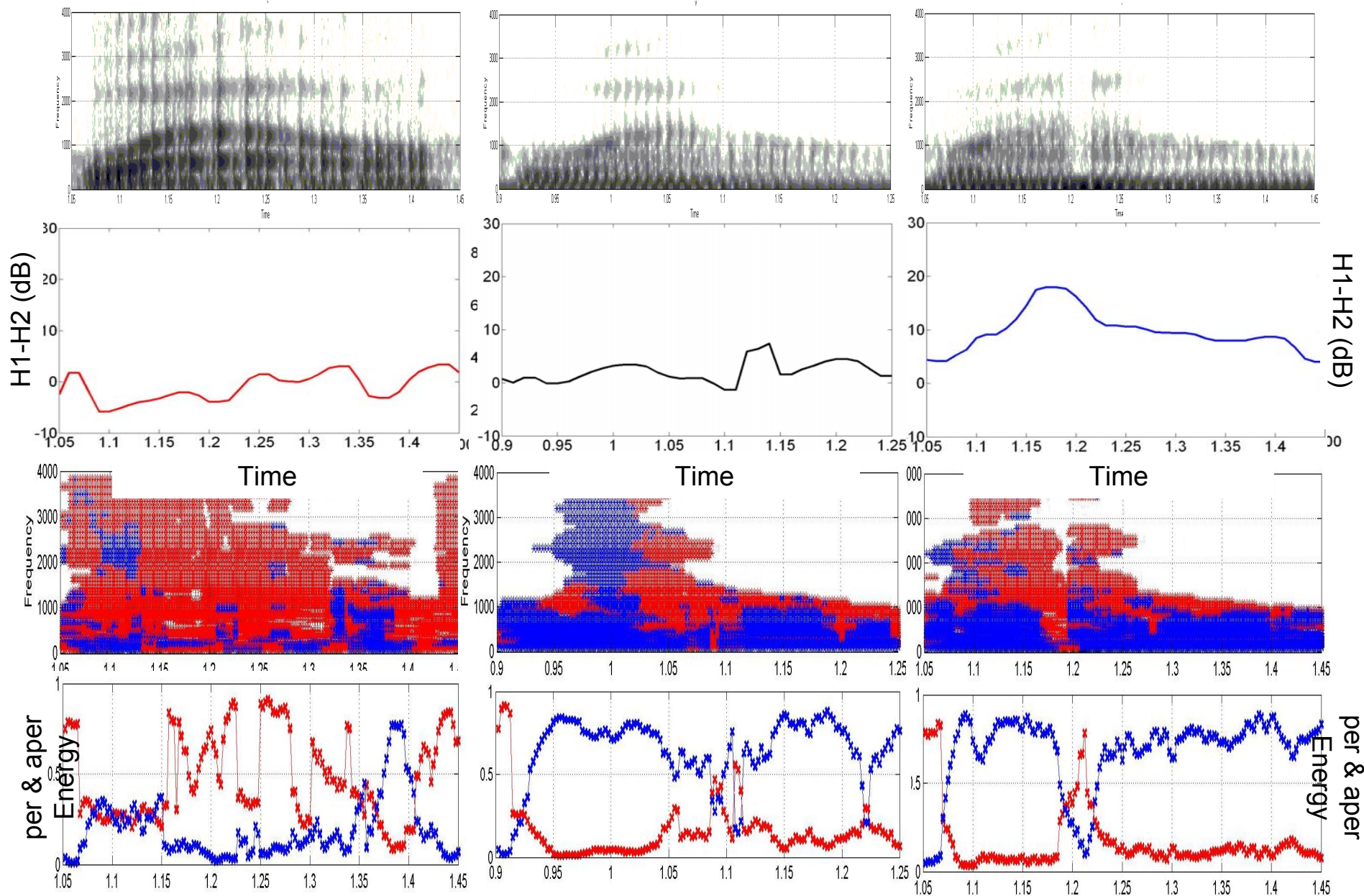
Figure taken from Klatt & Klatt, 1990



Creaky Voice

Modal Voice

Breathy Voice



Contents

Feature Extraction

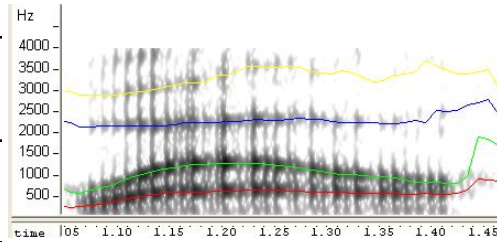
- MFCCs
- Voice Quality
- **Formants**

Motivation and Design

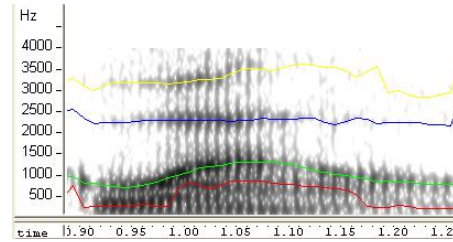
Results & Future Work

Formants

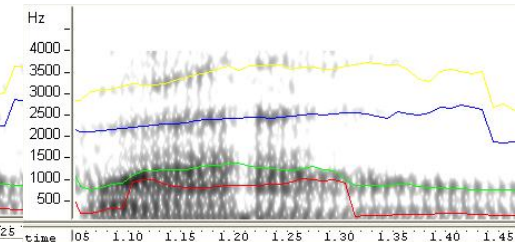
Creaky Voice



Modal Voice



Breathy Voice



- They describe the shape of the vocal tract.
- They don't vary considerably with the voice quality.



Contents

Feature Extraction

- MFCCs
- Voice Quality
- Formants

Motivation and Design

Results & Future Work

Motivation and Design

- There are 8 acoustical parameters (APs):
 1. The 4 formants
 2. The 2 energies: periodic and aperiodic
 3. H1-H2
 4. The Spectral Slope
- They were compared against the MFCCs last year and did well.
- Old Database: NIST '98 Evaluation Database
 - Problem: telephone filtered speech (200 Hz – 3400 Hz), which invalidated H1-H2
- New Database: Buckeye Corpus, from Ohio State University
 - Integrity of the low frequency information has not been distorted



Contents

Feature Extraction

- MFCCs
- Voice Quality
- Formants

Motivation and Design

Results & Future Work

Results

% Accuracy	Male	Female
60 MFCCs	99.67	100.00
7 APs (no H1-H2)	95.72	99.23
8 APs	94.17	98.21

- H1-H2 hindered overall performance.
- Overall performance of the APs was comparable to the MFCCs.

Future Work

- We should improve creakiness detection to improve H1-H2.
- We should improve harmonic detection.

