SINGLE FREQUENCY FILTER BANK BASED LONG-TERM AVERAGE SPECTRA FOR HYPERNASALITY DETECTION AND ASSESSMENT IN CLEFT LIP AND PALATE SPEECH

Mohammad Hashim Javid, Krishna Gurugubelli, and Anil Kumar Vuppala
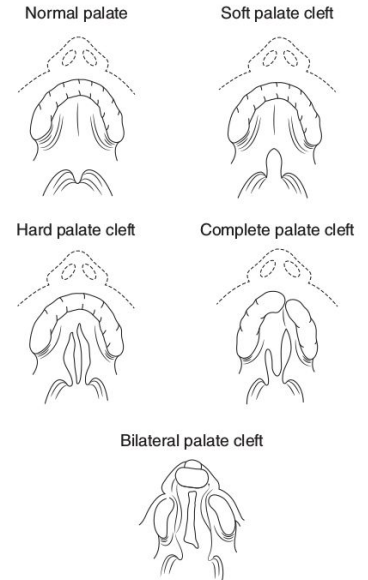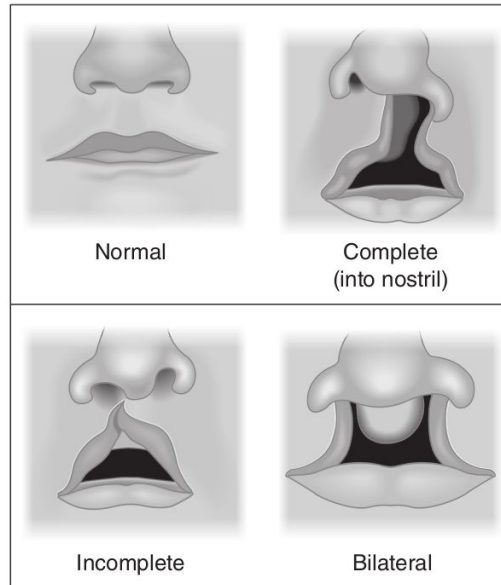
Hypernasality

Abnormal resonance in speech due to coupling of nasal tract.

Cleft Lip and Palate :

An abnormal opening or a fissure in lip and Palatal structures

[1]

Normal

Complete
(into nostril)

Incomplete

Bilateral

Normal palate

Soft palate cleft

Hard palate cleft

Complete palate cleft

Bilateral palate cleft

Hypernasality Detection

- Low frequency dominance

- High to Low energy ratio

- F1 bandwidth increase

- Formant around 250 Hz

Hypernasality Assessment

- Group delay spectrum

- High spectral resolution in low frequency region

- Distance between F1 and F2

- Formant around 250 Hz

Database

NMCPC-CLP

Severity	# speakers	# utterances
Normal	32	677
Mild	11	357
Moderate	14	373
Severe	16	406

- 73 speakers
  (41 CLP, 32 normal)

- 1813 sentences.

- 64 min duration ....?

- 76 different utterances.

- Average 25 utterances/ subject

Hypothesis

Information being present in the entire utterance.

- Suppress linguistic information across the sentence

- Nasal coupling is present throughout the utterance

Statistical Averaging - LTAS

Spectral cues

•Increase in band width of formants []

•Decrease in strength of formant []

•Steep valley between F1 and F2 []

•F1 and F2 drift apart [2]

High Spectral Resolution - SFFB

SFFB-LTAS

speech	SFFB	STAT

speech

Sampled at 8000Hz.

Normalized

SFFB

$$H(z) = \frac{1}{1 - az^{-1}}$$

a=0.985
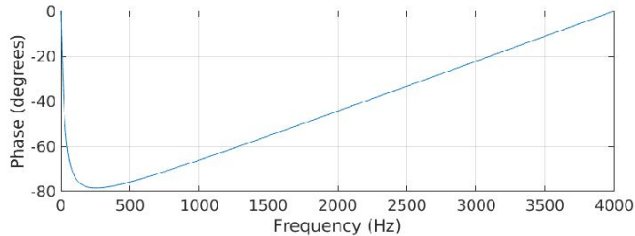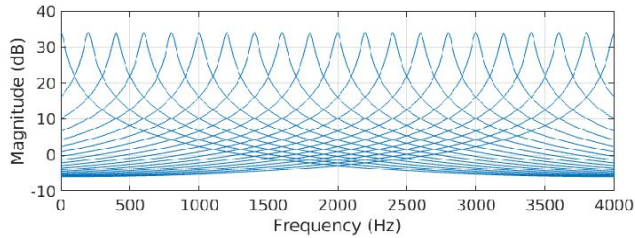
SFF frequency step 20HZ

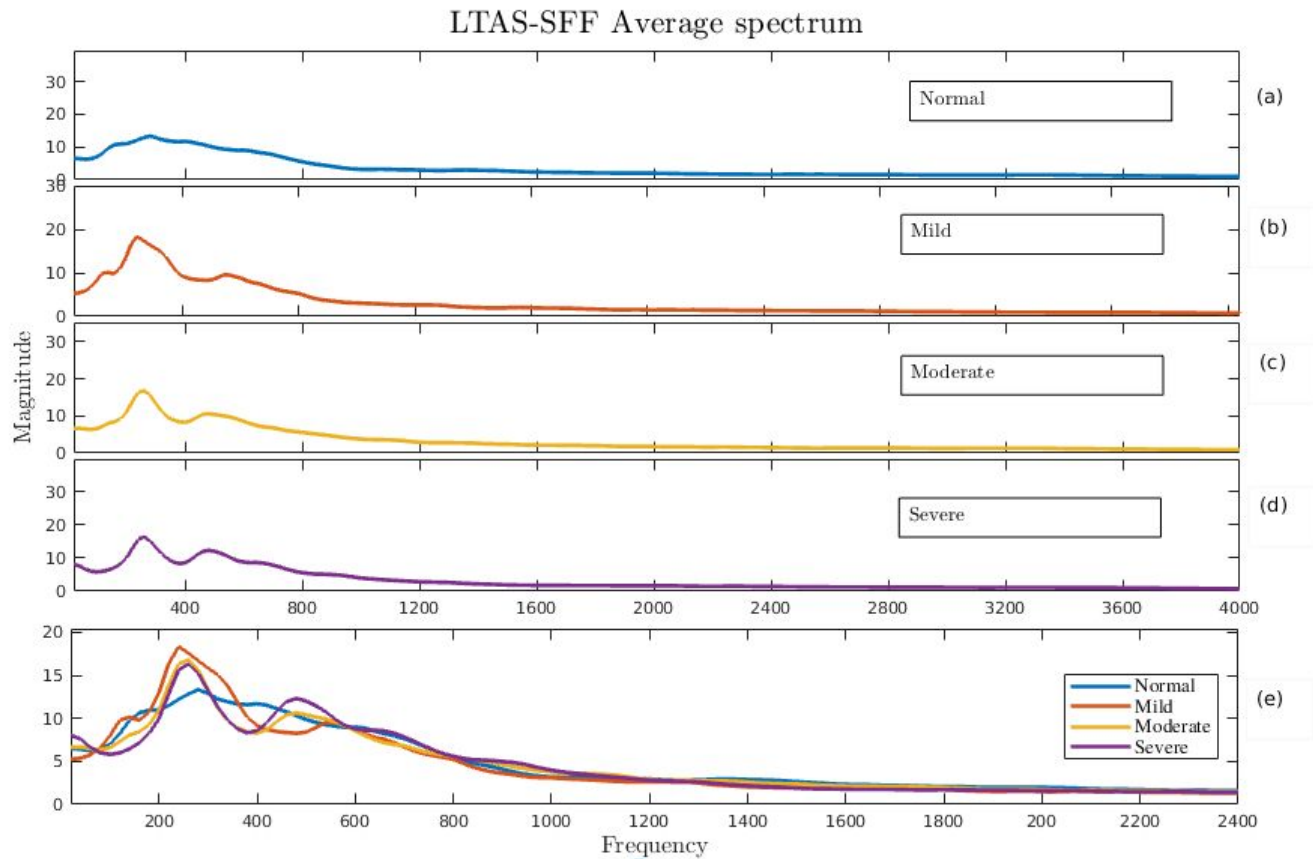**STAT**

Mean

Standard deviation

Kurtosis

Skewness

range

SFFB-LTAS

Average of SFFB-LTAS features from each severity.

The first 200 components corresponding to spectral average are shown here.

LTAS-SFF Average spectrum

Results

Feature	HND	HNA
MFCC	77.96	53.48
PLP	78.10	52.68
MGD	71.45	51.99
CQCC	*81.42*	*67.48*
MT-MFCC	75.92	54.18
PE-SFCC	82.13	65.14
Duration	45.87	38.89
Intonation	84.60	70.60
AFB-LTAS	85.90	72.20
SFFB-LTAS	**89.00**	**82.10**

Feature	HND	HNA
MFCC-STAT	84.07	62.64
PLP-STAT	85.44	63.19
MGD-STAT	76.37	59.07
CQCC-STAT	*84.07*	*70.05*
MT-MFCC-STAT	78.30	59.89
PE-SFCC-STAT	83.24	65.93
AFB-LTAS	85.90	72.20
SFFB-LTAS	**89.00**	**82.10**

LTAS-SFF Results
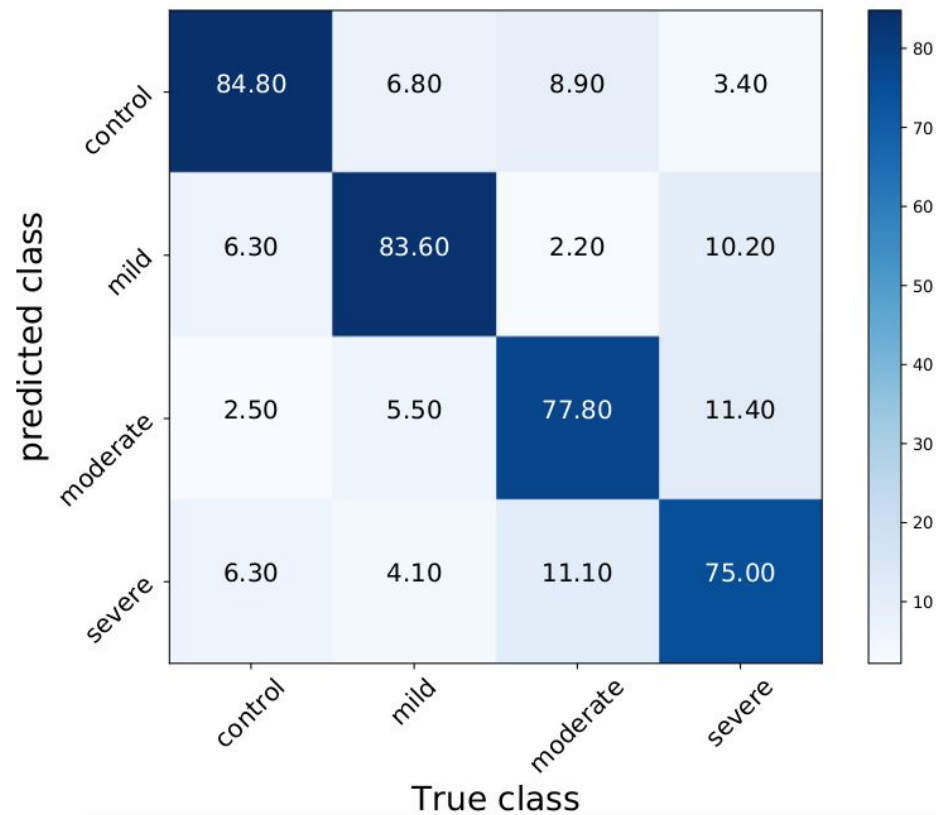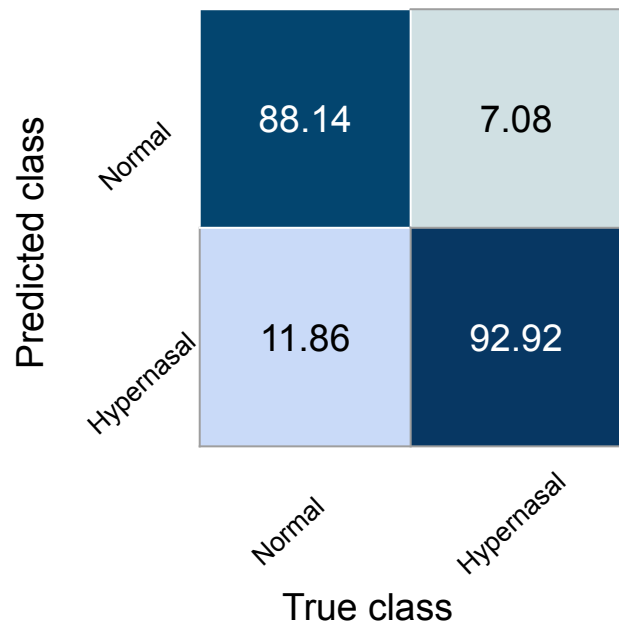
Conclusion

Long term spectral features are better for utterance level hypernasality evaluation .

Spectral resolution is vital for Assessment of hypernasality.

References

1.

2. S. Maeda. *The role of the sinus cavities in the production of nasal vowels*. In Proc. ICASSP, volume 7, pages 911–914, 1982.

3.

4.

5.