Relative Spectral Analysis-RASTA

By Davies Odu & Shengxin Qian

Introduction

- RASTA algorithm use auditory masking principle in reducing the perception of noise.
- This method uses frequency domain masking to maximize noise reduction while minimizing speech distortion
- Masking threshold is the spectral level (determined from the speech spectrum) below which non-speech components are masked by speech components in frequency.
- Masking threshold for complex signals is determined based on known information about input signals

Algorithm

- Compute the short time power spectrum of windowed signal.
- Transform spectral amplitude through a compressing static nonlinear transformation.
- Filter the time trajectory of each transformed spectral component.
- Transform the filtered speech representation through expanding static nonlinear transformation.
- Perform the overlap add synthesis and reconstruct the signal.

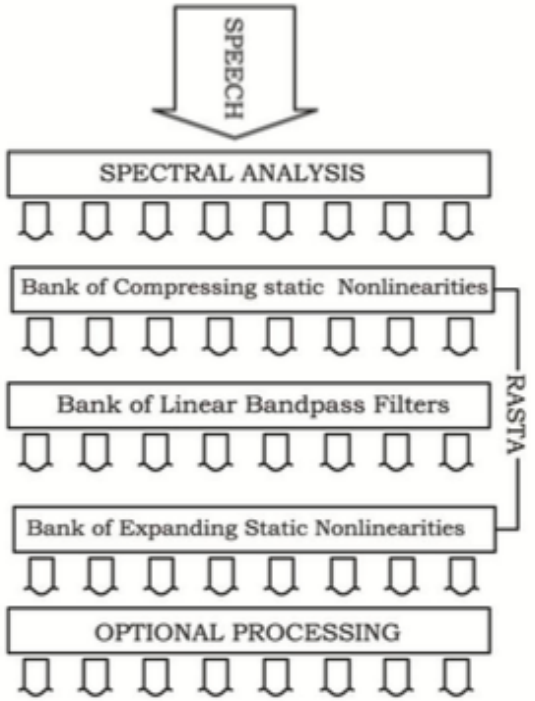


Fig. 5.2 Block diagram: RASTA processing system

One RASTA filter example

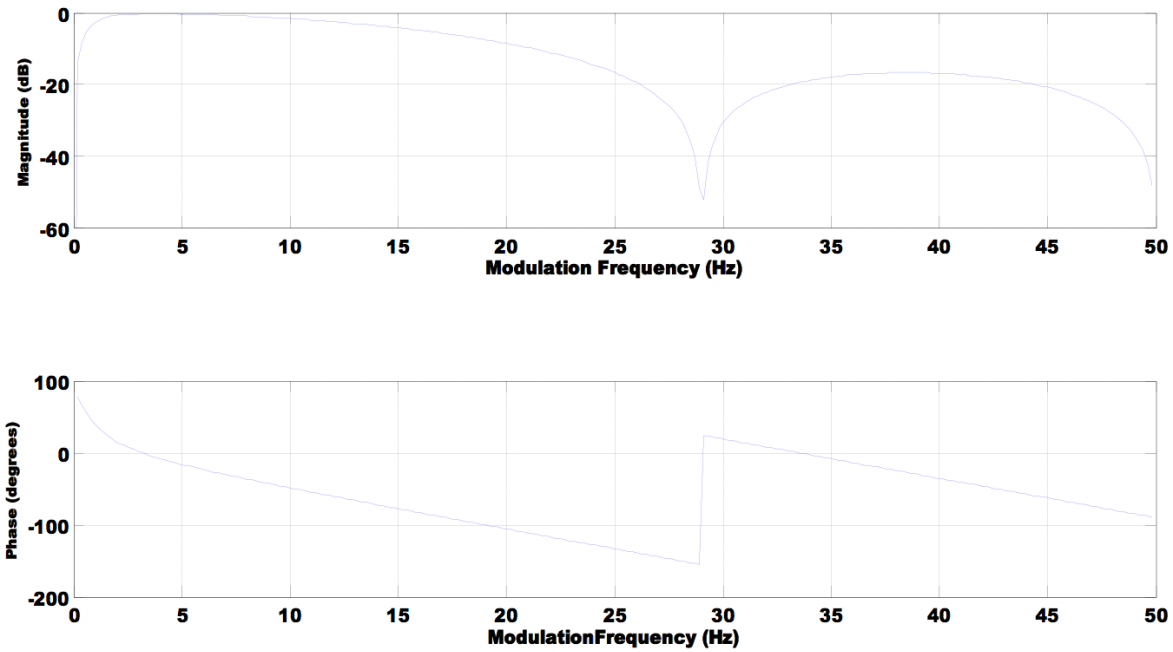
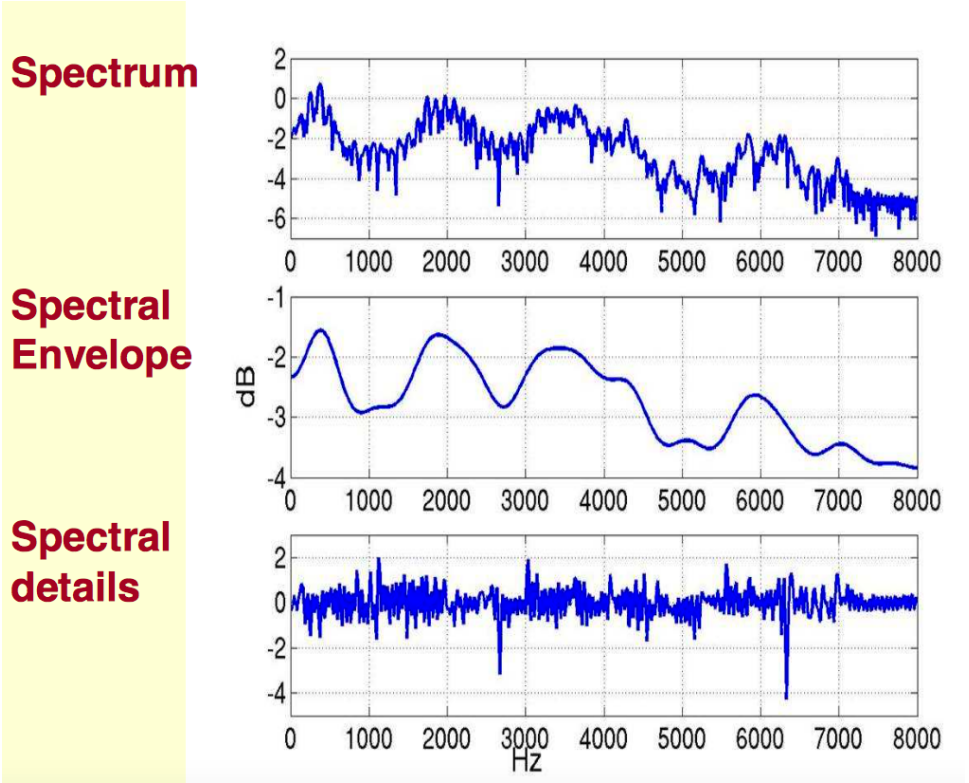


Fig. 5.3 Frequency response of fixed IIR RASTA filter

- The filter is designed with lower cut-off frequency of 0.26Hz. The filter slope decline 6dB/octave from 12.8Hz with sharp zeroes at 28.9Hz and at 50Hz.
- The low-pass filtering helps to smooth some of the fast frame-to-frame spectral changes present in the short-term spectral estimate due to analysis artifacts.
- The high-pass portion of the equivalent band pass filter is expected to alleviate the effect of convolutional noise introduced in the channel.

Intuition

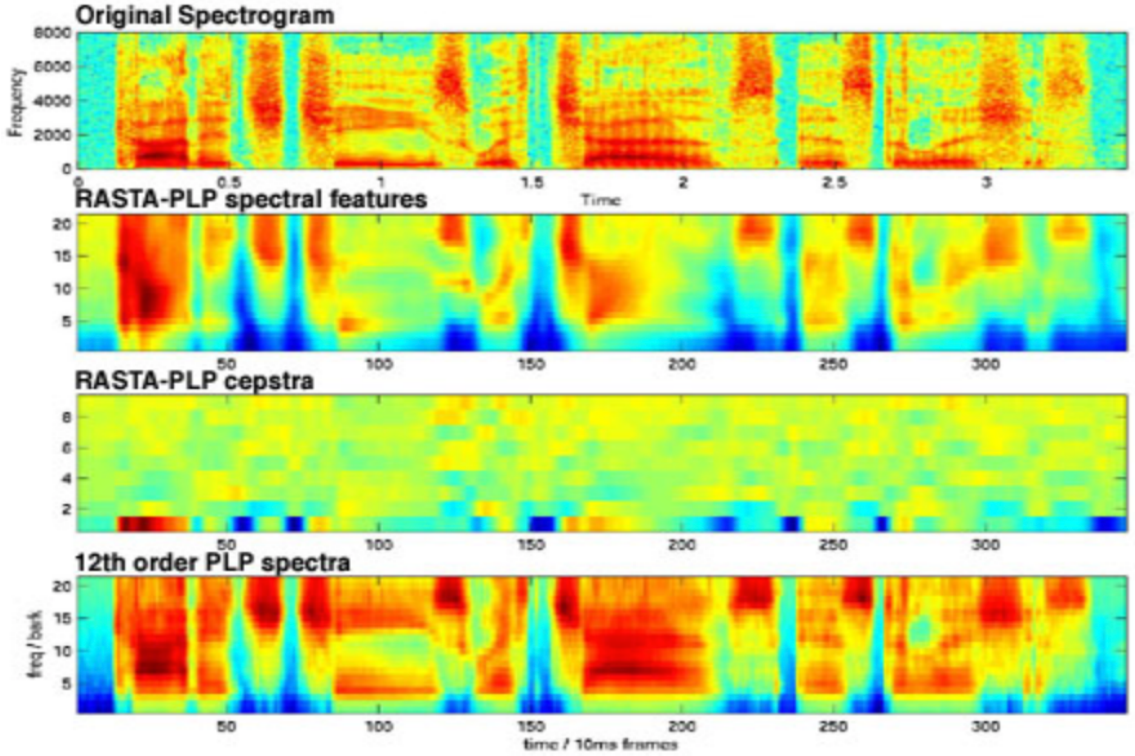
- For a wide range of frequency bands, the modulation spectrum of speech exhibits a maximum at about 4Hz, the average syllabic rate. RASTA filter suppresses the spectral components that changes more slowly than this frequency.
- The RASTA filter suppresses high modulation frequencies to account for the human's preference for signal change at a 4Hz rate. Disturbances such as additive noise may have different modulation spectrum properties than speech and often have modulation frequency components outside the speech range.



Comparison with other method

- Auditory masking VS Classical spectral subtraction approaches:
- Auditory based algorithm like RASTA outperforms other subtractive type residual noise suppression algorithms with respect to human perception with less distortion.

Demo



- Notice the auditory warping of the frequency axis to give more space to low frequencies and the way that RASTA filtering emphasizes the onsets of static sounds like vowels.

- Notice the greater level of temporal detail compared to the RASTA-filtered version. There is also greater spectral detail because our PLP model order is larger than the default of 8