
IEEE Speech Coding Workshop

Sept 17–20, 2000

Lake Lawn Resort

Delavan, WI



Jean-Marc Valin, Roch Lefebvre
University of Sherbrooke

Bandwidth Extension of Narrowband Speech for Low Bit- Rate Wideband Coding

Outline

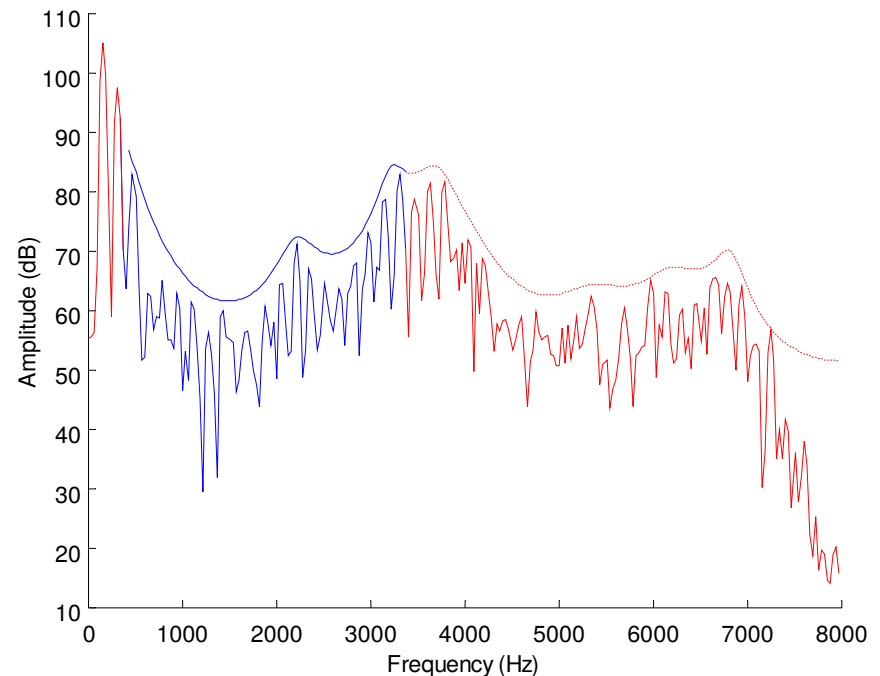
- Problem statement
- Proposed solution
- System performance
- Discussion

Problem Statement

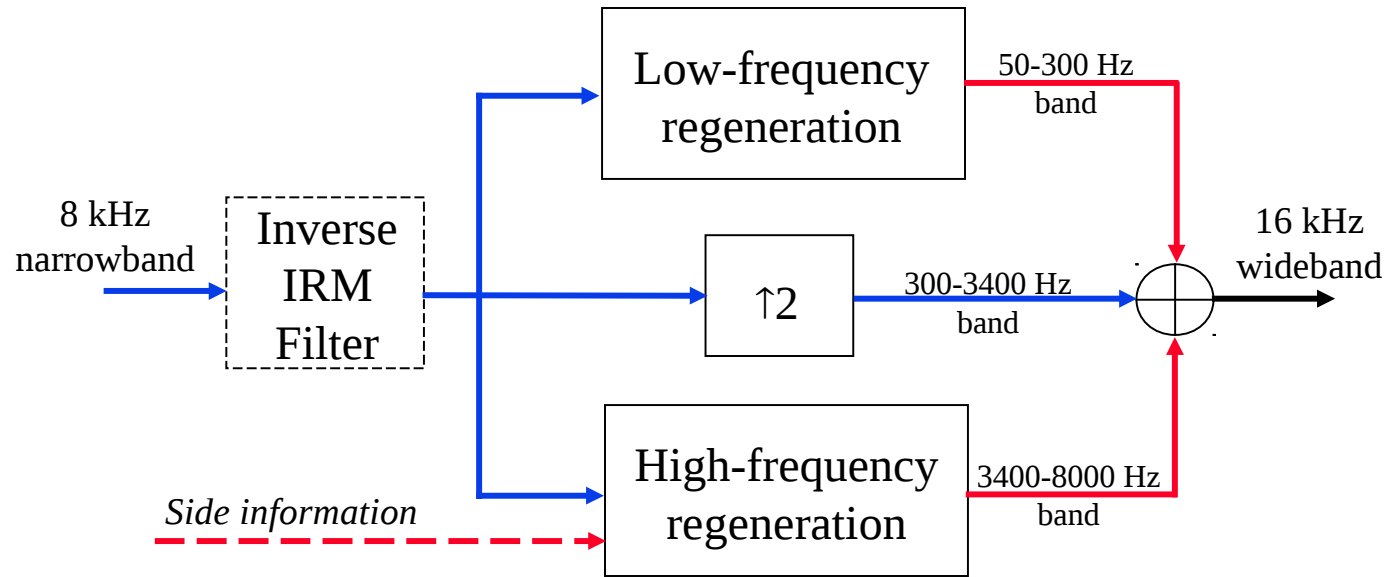
- Telephone Band: 300 - 3400 Hz
- AM Band: 50 - 7000 Hz
- How to make  sound like  with 500 bits/sec?
(G.729)
- We need to recover information from both low and high-frequency bands

Proposed Solution

- 1) Do our best to recover the wideband information from narrowband speech
- 2) Use coding for the information that cannot be recovered
 - Recovered information :
 - Low-frequency band
 - High-frequency excitation
 - Coded information :
 - High-frequency spectral envelope



System Overview

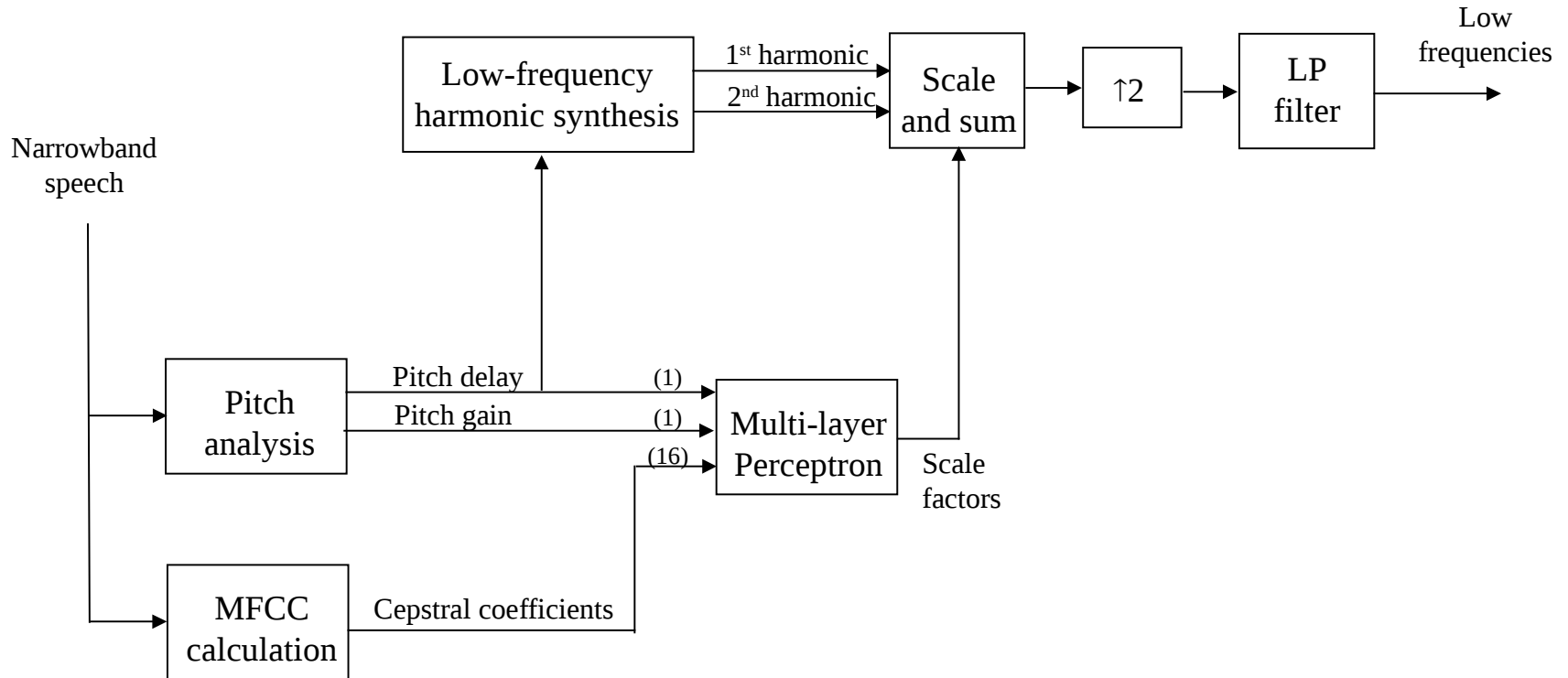


- Inverse IRM filter is optional
 - produces a flat response from 200-3500 Hz

Low-Frequency Regeneration (1/2)

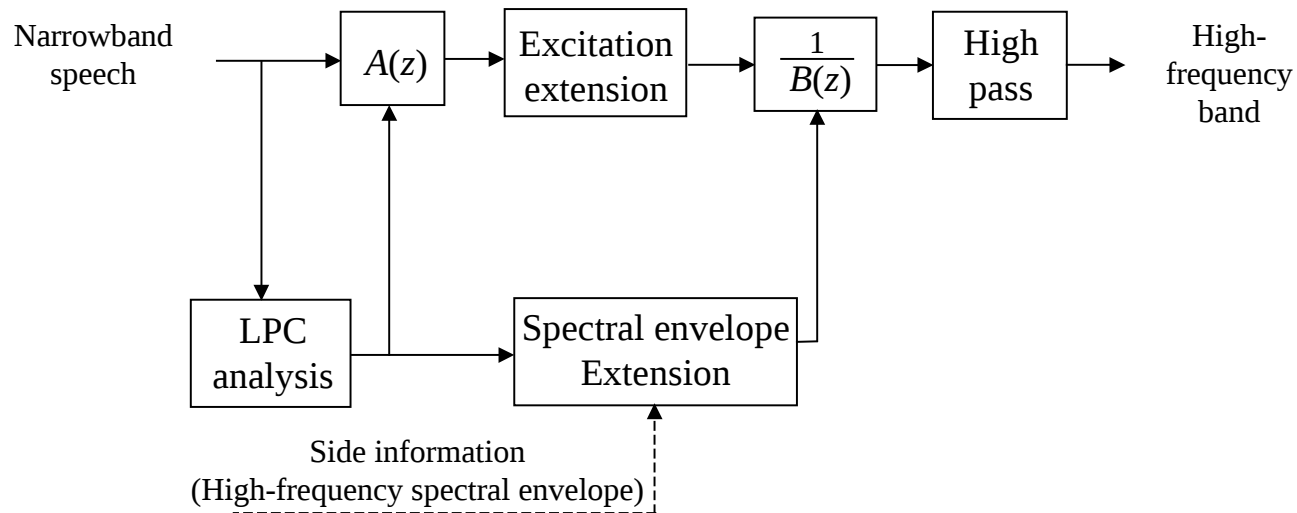
- Assumptions :
 - Only pitch harmonics need to be recovered
 - In general, no more than two pitch harmonics below 200 Hz
 - Absolute phase is not perceptually relevant
- Frequency of harmonics determined from pitch analysis
- Amplitudes determined from feed-forward multi-layer perceptron (output in log domain)

Low-Frequency Regeneration (2/2)

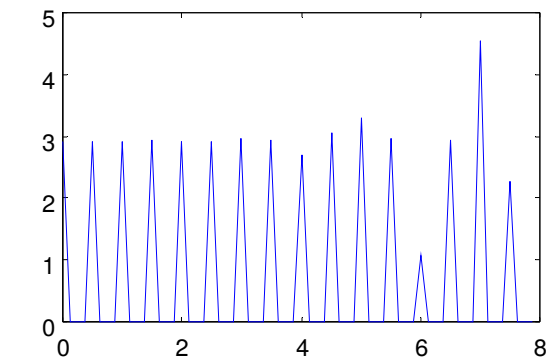
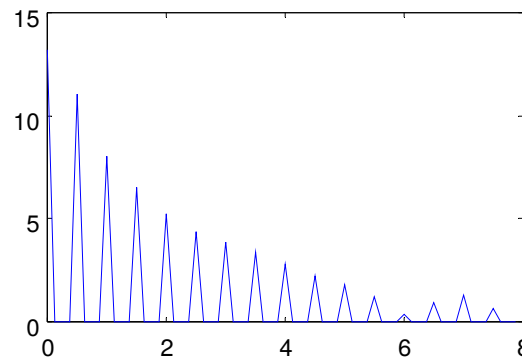
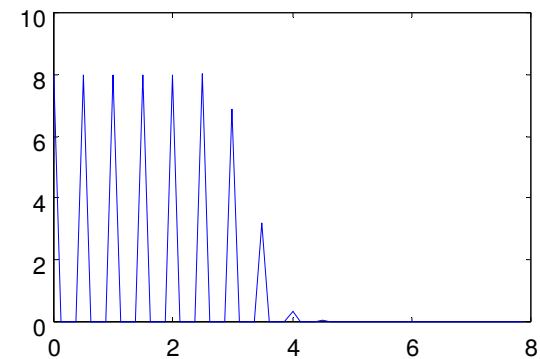
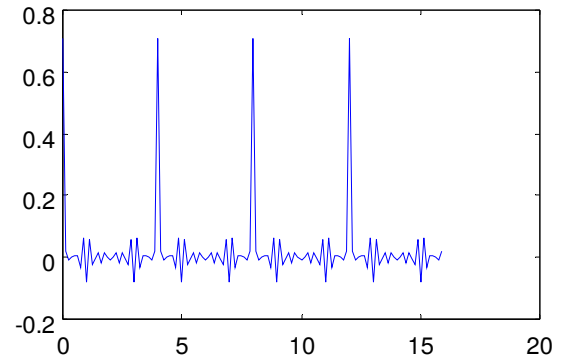
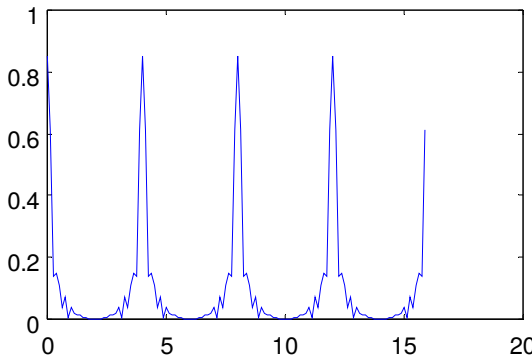
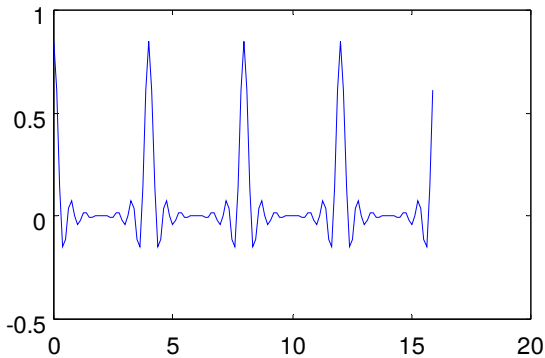
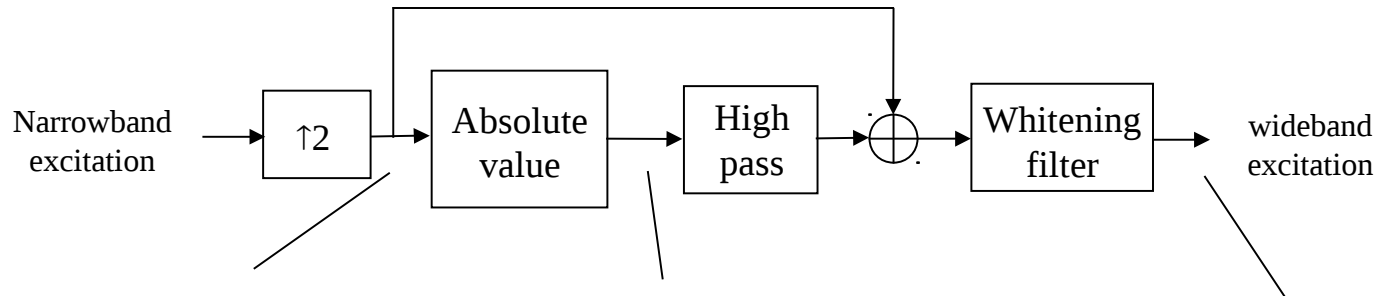


High-Frequency Extension

- Excitation-filter model (16 ms frames)
- Problem is separated in two parts
 - Excitation extension (no side information)
 - Spectral envelope coding (side information)



Excitation Extension



Spectral Envelope Coding

- Spectral envelope calculated from the wideband LPC coefficients
- Quantization of the 3000-8000 Hz range (40 points)
 - Log domain
 - 8-bit Vector Quantization (500 bits/s side information, using 16 ms frames)
- Concatenation with envelope obtained from LPC analysis on narrowband speech

Objective results

- Low-frequency band
 - 3 dB RMS error on harmonic amplitude
- High-frequency band
 - 3.6 dB RMS error on envelope
 - No objective measure for excitation extension (perceptually very close to original)

Subjective Results

female

male



Original wideband



Recovered from
original IRM-filtered speech



Recovered from
G.729 coded speech



Discussion

- Highlights
 - Expand IRM-filtered telephone-band speech to AM band
 - Very low side information rate (500 bits/s)
- Areas of improvement
 - Use high-band spectral estimation before coding
 - Use residual low-frequency information (below 300 Hz)
 - Noise robustness
 - Post-filtering