Complex Wavelet Modulation Sub-Bands and Speech

Jean-Marc Luneau¹, Jérôme Lebrun² and Søren Holdt Jensen¹

¹ Department of Electronic Systems - MISP, Aalborg University, DK-9220 Aalborg, Denmark
² CNRS - I3S, UMR-6070, FR-06903 Sophia Antipolis, France

Modulation Transform

Modulation freq. 2-16Hz are crucial to speech (Greenberg’97).
Usual modulation transform: 2 main stages, 3-D representation

\[ X_k[n] \rightarrow Y_k[n] \]

- Human hearing is sensitive to phase (Lindemann et al. ’99): Influence on mapping in cochlea and IHC firing rate.

⇒ need to preserve phase at second stage of transform.

- Speech/voiced time-freq. signal envelopes are of “polynomial” type with magnitude and phase like the model:

\[ X_k(t) = |w(t)|e^{j2\pi(\omega_1 t + \omega_0)} \]

- Limitations of the standard Hilbert envelope:

The standard Hilbert approach:

\[ X_k[n] = X_k[n] \exp(j\hat{\phi}_k[n]) \]

with \( \hat{\phi}_k[n] = \arg(X_k[n]) \)

- Processing the sole envelope creates artifacts

⇒ need for joint magnitude-phase processing.

Experiments on the model buried in white noise:

- Sparsity of representation: denoising by (soft) thresholding coeff. \( Y_k[n] \).

- Processing the sole envelope creates artifacts

⇒ need for joint magnitude-phase processing.

Experiments on the model buried in white noise:

Original envelope recovered \( \Delta \) Phase information denoised

⇒ Ideal second stage for the modulation transform.

Where we stand

- From modulation frequencies to modulation sub-bands.
- Promising speech denoising results.
- Sound files available at: www.luneau.info/ITRW

Funding

The work of Jean-Marc Luneau is supported by the EU by a Marie-Curie Fellowship (EST-SIGNAL program: http://est-signal.i3s.unice.fr) under contract No MEST-CT-2005-021175

Complex wavelet sub-bands

Wavelets: log-scale resolution for low frequencies: 2-16Hz.
Complex Waves: joint amplitude/phase processing and good polynomial approximation.

⇒ Decomposition of \( \hat{X}_k[n] \) into modulation sub-spaces.

- Analytic and orthogonal filter-banks:

- Bandpass orthogonal filters condition on \( q[n] \) for analyticity (Blu et al.’03)

\[ q[n] := j^n u[n] \]

\[ U^*(1/z)U(z) + U^*(-1/z)U(-z) = 2 \]

with

\[ u[n] = \sqrt{\frac{5}{16}}[-1,0,5,5,0,-1] + j \sqrt{\frac{5}{16}}[0,1,3,3,1,0] \]

⇒ Joint magnitude-phase denoising

by soft thresholding inside orthogonal modulation sub-spaces:

Original envelope recovered \( \Delta \) Phase information denoised

⇒ Ideal second stage for the modulation transform.