Empirical Mode Decomposition For Noise-Robust Automatic Speech Recognition
Kuo-Hao Wu and Chia-Ping Chen

Introduction

- empirical mode decomposition
- a double iterative algorithm to extract intrinsic mode functions
- EMD-based feature post-processing
- illustrative examples
- evaluation on Aurora database

Empirical Mode Decomposition (EMD)

- generalization of the Fourier analysis
- Hilbert transform pairs of the real and imaginary parts
- allow time-varying amplitude and frequency
- better suited for non-stationary signals and non-linear systems

Extracting the Intrinsic Mode Functions (IMF)

- an outer loop to control the number of IMFs to be extracted
- an inner loop to enforce the required conditions of IMF
- the upper envelope \( u(t) \) and lower envelope \( l(t) \) average to 0
- the extremal points alternate around 0
- block diagram

Proposed Post-Processing

Property of IMF

- As \( k \) increases, the remainder \( r(t) \) is more and more smooth, as well as the extracted IMF \( c_k(t) \)

\[ \hat{c}(t) = c(t) - \sum_{i=1}^{K} c_i(t) \quad (1) \]

- The noisier the signal \( x(t) \), the more oscillations
- So we simply subtract the lower-order IMFs to reduce the spurious spikes introduced by the noises

Using a dynamic number of IMFs for subtraction

Using 1 or 2 IMFs for subtraction

Conclusion

- generalization of the Fourier analysis
- generalization of the filtering methods
- subtracting not certain frequency components, but functions with time-varying amplitude and frequency
- effective on MVN that is known to be noise-robust
- verify that the noisier the signal, the more post-processing is needed