Title	経験的モード分解を用いた音信号分析に関する基礎研究
Author(s)	澤口，知希
Citation	
Issue Date	2010-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/8946
Rights	
Description	Supervisor:鵜木祐史，情報科学研究科，修士

A study on a method of speech signal analysis using the Empirical Mode Decomposition

Tomoki Sawaguchi (0810030)

School of Information Science,
Japan Advanced Institute of Science and Technology

February 9, 2010

Keywords: empirical mode decomposition, intrinsic mode function, noise reduction, comodulation, modulation index.

Recently, the Fourier transform and wavelet transform have been generally used to analyze the signal as the technique for time-frequency analysis. These methods can analyze the temporal spectral fluctuations of the signal in the time-frequency domains, if the analytical signal is assumed to be stationary. However, realistic signals (e.g., electroencephalogram (EEG), seismic waves, and speech signal, etc.) are non-stationary signal so that these methods cannot precisely analyze non-stationary fluctuations of the instantaneous amplitude and phase of the signal. In recent years, the empirical mode decomposition (EMD) technique has been used for analyzing non-stationary signal. This method was originally proposed by Huang *et al.* The EMD has been used for analyzing EEG and for exploring the source from seismic waves. Currently, this technique has been applied for speech signal processing. In particular, EMD-based noise reduction methods have been proposed to reduce the musical noise from the restored speech and to robust voiced/unvoiced decision in noisy environment.

Speech signal is generally non-stationary signal. Therefore, speech representation based on the EMD seems to be suitable for representing speech features such as non-stationary fluctuations, in comparison with the traditional methods. However, it is unclear how noisy speech can be represented as suitable forms (separately, speech and noise) based on the EMD. Moreover, it is also unclear whether these components (speech and noise) can be completely separated in these representations. Since their methods for noise reduction use particular IMFs corresponding to noise components to be noise-reduced speech, they may remove significantly IMFs of non-stationary speech by reducing noise components on these representations.

In this study, we investigate the properties of the analysis method for non-stationary signal using the EMD and characteristics of the decomposed IMF (intrinsic mode function). We then examine the possibility of analysis method, by utilizing these advantages of the EMD.

According to the EMD algorithm, we investigate the properties of the EMD and characteristics of the decomposed IMFs. With this advantage, we consider a separation of non-stationary speech signal and stationary white noise as an application.

As the results, we found that the essence of signal analysis based on the EMD is to separate non-stationary signal from stationary signal in the signal representation, by a common-envelope-based decomposition.

We thus proposed a noise reduction method based on the EMD, according to our considerations. we consider a separation of non-stationary speech signal and stationary white noise. We carried out simulations for verifying effectively of the proposed method. 30 speech signals (each three words from five males and five females) of ATR database a-set were used in these simulations. White noise was added into original speech signals to obtain noisy speech signals in which SNRs are -5, 0, 5, 10, 15, and 20 [dB]. We compared the proposed method with Molla & Hirose' method. Improved SNR, Improved LSD and PESQ were used to evaluate verifying effectively of the proposed method. As the results, both SNR and LSD of the restored speech are improved. But PESQ is not improved. We compared the proposed method with Molla & Hirose's method. Improvements in the proposed method are greater than those in Molla & Hirose's method in the case of lower SNR condition. Moreover, in their method, speech signal was over-filtered out in the case of higher SNR so that the restored signal was corrupted due to over subtraction. In contrast, the proposed method can reduce the noise components from noisy speech without distortion in the case of higher SNR.

We proposed a noise reduction method based on the EMD, according to our considerations. It was shown that that non-stationary speech and stationary noise can be separated by using the EMD-based noise reduction method.