ICIC Express Letters ICIC International ©2018 ISSN 1881-803X
Volume **12**, Number **4**, April **2018** pp. **337–343**

ANALYSIS WITH HILBERT-HUANG TRANSFORM FOR POST-APNEIC SNORING SOUNDS IN PATIENTS WITH SEVERE OBSTRUCTIVE SLEEP APNEA

Hirotaka Takahashi[1], Tsuyoshi Mikami[2], Satoshi Ueki[1] and Kazuya Yonezawa[3]

[1]Department of Information and Management Systems Engineering
Nagaoka University of Technology
1603-1 Kamitomioka, Nagaoka, Niigata 940-2188, Japan
hirotaka@kjs.nagaokaut.ac.jp

[2]Department of Engineering for Innovation
National Institute of Technology, Tomakomai College
443 Nishikioka, Tomakomai, Hokkaido 059-1275, Japan

[3]National Hospital Organization Hakodate Hospital
18-16 Kawaharacho, Hakodate, Hokkaido 041-8512, Japan

Received October 2017; accepted January 2018

Abstract. *The acoustic properties of snoring sounds are known to be specific to the obstructive sleep apnea syndrome (OSAS). To clarify such specific properties may be useful for medical and physiological diagnosis of OSAS, and thus many researchers have analyzed them using traditional spectral analysis such as Fourier transform or linear predictive coding. In our research, Hilbert-Huang transform (HHT), a novel nonlinear and non-stationary spectral analysis method, is adopted to investigate difference of the acoustic properties of snoring sound waveforms. The snoring sounds analyzed in this paper are recorded from seven OSAS patients while sleeping, and the snoring sound waveforms in the recorded data are cut out semi-automatically. After that, all snoring sound waveforms are allocated to 3 categories: sounds produced by rebreathing within 10 seconds (hypopneic), snoring sounds produced by rebreathing after 10 seconds of an apnea event (post-apneic: class 0), and snoring sounds produced by normal breathing during sleep (non-apneic: class 1). It is important to understand the difference of the acoustic properties of these categories and investigate such difference from a physiological point of view. As the first step, we analyzed the class 0 and class 1 waveforms in this paper. We found that we could qualitatively understand the difference of the acoustic properties between class 0 and class 1 waveforms by using HHT.*

Keywords: Snoring sounds, Hilbert-Huang transform, Obstructive sleep apnea syndrome

1. **Introduction.** The snoring is known to be related to the obstructive sleep apnea syndrome (OSAS). Therefore, many medical and physiological researchers have focused on the biomechanism of snoring and the acoustic properties [1]. Beck et al. [2] pointed out two dominant patterns, called simple-waveform and complex-waveform, of snoring sounds based on the linear acoustic properties. The snore of simple-waveform means a quasi-sinusoidal waveform whose spectrum consists of a single prominent peak at the fundamental frequency and two or three harmonics. On the other hand, the snore of complex-waveform is characterized by multiple and equally spaced peaks of power, such as comb-like spectrum. Quinn et al. [3] found two distinct patterns of waveforms and spectra in palatal and tongue base snoring sounds. The palatal snores have a prominent peak corresponding to their fluttering mechanism. The tongue base snores have more higher frequency components and they look noise-like waveforms. Fiz et al. [4] found that the presence of a fundamental frequency and several harmonics in snoring sounds

of many simple snorers and a low frequency peak with the second energy scattered on a narrower band and without clearly identified harmonics in those of OSAS patients. Many other researchers have also analyzed the acoustic properties of snoring sounds. Especially, the formant-like spectral peaks have been focused for the purpose of classifying OSAS patients and simple snorers [5, 6]. Most of these conventional studies have used some linear analysis methods such as fast Fourier transform (FFT) and linear predictive coding (LPC). On the other hand, the snoring sounds have been considered to consist of the mixture of nonlinear oscillation sounds of the oropharyngeal soft tissues and the airflow noises during inhalation. In addition, the dynamics is changing gradually or suddenly; the waveform is suddenly changing and distorted from a sinusoidal wave. Thus, it is natural to consider that the snoring has strong nonlinear and non-stationary properties in its sound structure.

The Hilbert-Huang transform (HHT), which consists of an empirical mode decomposition followed by the Hilbert spectral analysis, has been developed by Huang et al. [7]. Compared with the Fourier decomposition and wavelet decomposition, the empirical mode decomposition approach has an advantage for analyzing the non-stationary and nonlinear data since it decomposes the signal into intrinsic mode functions based on the time scale of the signal itself with adaptive nature. The HHT is not limited by time-frequency uncertainty and thus it provides a high resolution time-frequency analysis. Therefore, the HHT is a powerful tool to extract the information from the snoring sounds.

HHT has also been applied to the airway pressure signals related to OSAS [8, 9]. The data analyzed in these papers are the airway pressure signals obtained from nasal breath [8] and oronasal breath [9], respectively. In these studies, the histogram of HHT spectra in a specific frequency range is calculated for 300 seconds and used to discriminate between OSAS and non-OSAS persons. These studies did not focus on the time structure since they calculated the histogram of HHT spectra. One of our hypothesis is that some useful information about OSAS would also be involved in the time structure. Mikami et al. [10] adopted HHT to clarify the nonlinear and non-stationary properties in a nasal snoring sound. They reported that two types of frequency fluctuation were found in the time-frequency-amplitude map (HHT map).

Azarbarzin and Moussavi [11] focused on the snoring sound segments from the respiratory recordings. They divided them into three categories: non-apneic, hypopneic, and post-apneic and pointed out that the properties of these three classes varied significantly within a subject depending on the level of obstruction and this variability was associated with the severity of OSAS.

Based on these categories of Azarbarzin and Moussavi [11], it is important to understand the difference of the acoustic properties of these classes and investigate such difference from a physiological point of view. Thus, as the first step, HHT is adopted to investigate the difference of the acoustic properties of snoring sounds between non-apneic (class 1) and post-apneic (class 0) in this paper.

The paper is organized as follows. In Section 2, we give an overview of the experiment of data recording and data processing. Moreover, we briefly review the Hilbert-Huang transform. In Section 3, the analysis results and discussion are presented. Section 4 is devoted to a summary.

2. Method.

2.1. Data recording and processing. Seven individuals, who were referred for full-night polysomnography (PSG) research at the National Hospital Organization Hakodate Hospital, gave written consent to participate in this research. The research was approved by National Hospital Organization Hakodate Hospital.

The used instrument is a portable linear pulse code modulation (PCM) recorder, Olympus LS-11, whose setting is 44.1kHz sampling rate and 16-bit resolution. The external microphone, Olympus ME30W, is set in the prop of head side of the bed.

Overnight sleep sounds of seven OSAS patients during sleep were recorded. The recording time was approximately 7-10 hours. These data include the snoring sounds and the other environmental sounds. Although we do not discuss the details in this paper, the snore sound waveforms in the recorded data were identified and cut out semi-automatically [12]. About a few thousands sound waveforms per night were obtained from one subject. After that, all the sound waveforms were also manually allocated to 3 categories: sound waveforms produced by rebreathing within 10 seconds (hypopneic), snoring sound waveforms produced by rebreathing after 10 seconds of an apnea event (post-apneic: class 0), and snoring sound waveforms produced by normal breathing during sleep (non-apneic: class 1). The number of class 0 waveforms was around twenty per subject. On the other hand, the number of class 1 waveforms was a few hundred per subject. Thus, we picked 50 of waveforms out randomly from class 1 waveforms set, and analyzed these 50 waveforms as the class 1 waveforms set for each subject.

2.2. **Hilbert-Huang transform.** Since the Hilbert-Huang transform (HHT) is not limited by the time-frequency uncertainty relation, unlike the short-time Fourier transform (STFT) and the wavelet transform (WT), it provides high time-frequency resolution, and therefore enables us to investigate phenomena that have even rapid changes in frequency as well as slow changes, while the STFT and the WT can detect only the latter.

The HHT consists of two steps: a mode decomposition step and a spectral analysis step. At the mode decomposition step, input data will be decomposed into some intrinsic mode functions (IMFs). Then, at the spectral analysis step, the time evolution of the amplitude and the phase of each IMF are calculated by means of Hilbert transform. This spectral analysis is called Hilbert spectral analysis (HSA). Hence, we can express input data $s(t)$ as

$$s(t) = \sum_{i=1}^{N_{\mathrm{IMF}}} c_i(t) + r(t) = \sum_{i=1}^{N_{\mathrm{IMF}}} a_i(t)\cos(\phi_i(t)) + r(t), \tag{1}$$

where N_{IMF} is the number of IMFs, and $c_i(t)$, $a_i(t)$, $\phi_i(t)$ are the ith IMF, the instantaneous amplitude (IA), and the instantaneous phase (IP), respectively, while $r(t)$ is the residual or trend, which is the non-oscillatory part mode of the data. Instantaneous frequency (IF) $f_i(t)$ can be defined as

$$f_i(t) = \frac{1}{2\pi}\frac{d\phi_i(t)}{dt}. \tag{2}$$

Through this procedure, the HHT allows us to extract instantaneous time-frequency and time-amplitude trajectories. This time-frequency-amplitude map (HHT map) allows for a high resolution time-frequency analysis of waveforms with strong frequency modulation.

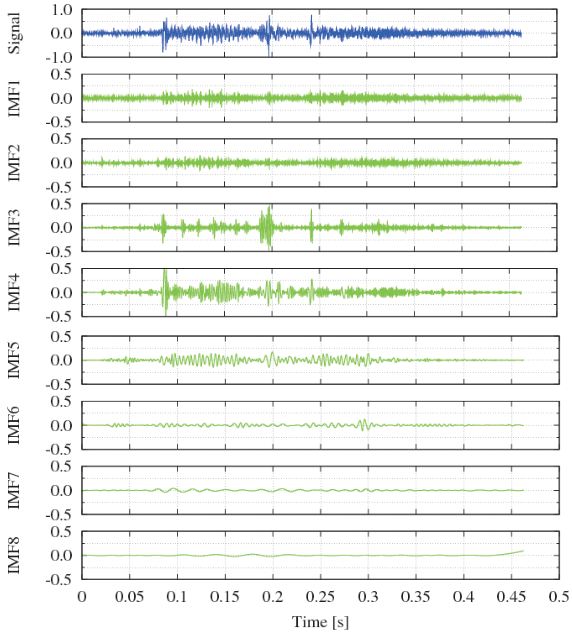
The original HHT proposed in [7] uses the empirical mode decomposition (EMD) as decomposition method. It is known, however, that the EMD has many drawbacks, such as mode-mixing [13, 14], mode-splitting [15], and lack of mathematical foundation [16], and therefore many improved methods of the EMD have been proposed [15, 16, 17, 18]. We will adopt the ensemble EMD (EEMD), which is proposed by Wu and Huang [19] to overcome the mode-mixing problem.

More detailed review of the algorithm of the HHT including the EEMD is found in [20, 21, 22].

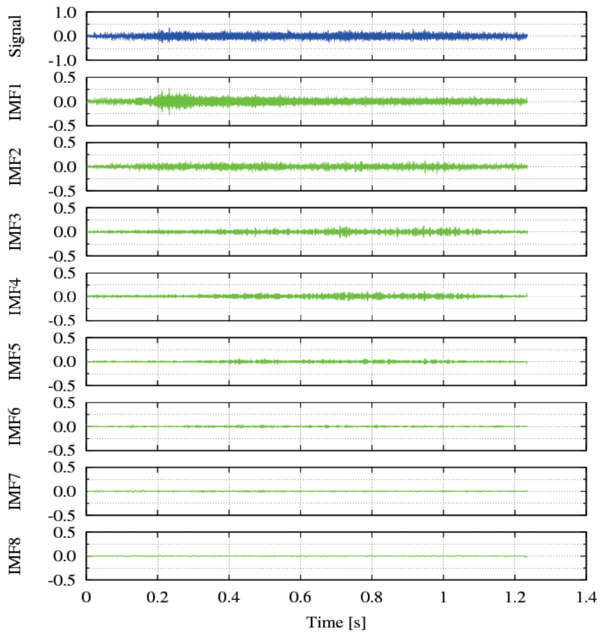
There are some parameters to be fixed in the EEMD. In this paper, we choose the parameters for the EEMD as follows: the stoppage criterion $\epsilon = 10^{-3}$, the standard deviation of the Gaussian noise in EEMD $\sigma_{\mathrm{e}} = 10^{-7}$ and the size of ensemble $N_{\mathrm{e}} = 200$.

As for N_e, we verified that the results hardly change even with $N_e > 100$ but the value $N_e \approx 50$ is too small [20]. We specify $N_{IMF} = 8$ in this paper.

3. **Results and Discussion.** Before analyzing the snoring sound waveforms, the time series for one snoring sound waveform are clipped, and all sound waveforms are categorized in class 0, class 1 or hypopneic, which are explained in Section 2.1. Since the amplitude and length of one snoring sound waveform are different from each waveform and each subject, etc., the clipped time series are standardized. The top panels of Figures 1(a) and 1(b) show the typical examples of class 0 and class 1 snoring sounds waveforms from one subject, respectively. Note that the time scale between Figures 1(a) and 1(b) is different.



(a) Class 0: Analyzed signal and its IMFs



(b) Class 1: Analyzed signal and its IMFs

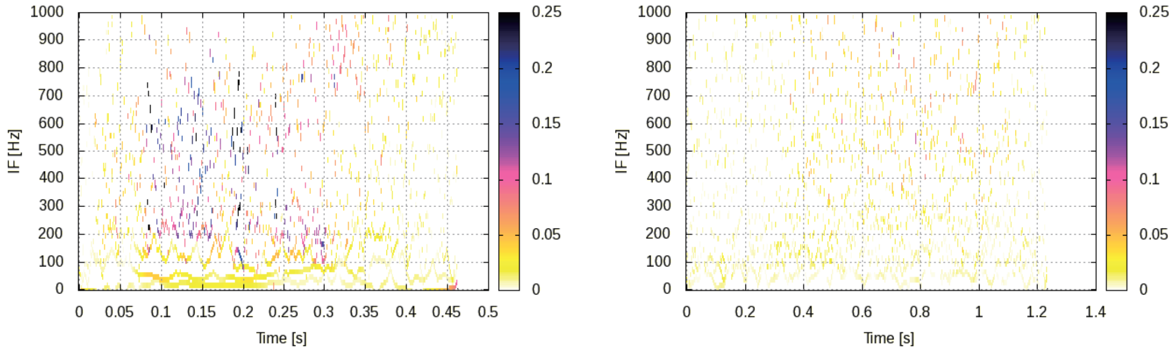
FIGURE 1. Typical examples of the decomposed results. The analyzed snoring sound waveforms from one subject and estimated 8 IMFs, respectively. Note that the time scale between (a) and (b) is different, and the scale of vertical axis between signal (top panel) and 8 IMFs is also different.

First, we apply the EEMD to each snoring sound waveform. In Figures 1(a) and 1(b), we show the decomposed results for snoring sound waveforms. The upper panel in each figure box in Figure 1 shows the analyzed snoring sound waveforms, and each snoring sound waveform is decomposed into 8 IMFs in this case. Note that the scale of vertical axis between signal (top panels) and 8 IMFs is also different.

Next, for each of IMFs, we compute the IA and IF with HSA. Based on the obtained IFs and IAs, we plot the time-frequency-amplitude map, which we call the HHT map, in Figure 2. In these figures, we use the time resolution $\Delta t = 1/44100$sec and frequency resolution $\Delta f = 10$Hz.

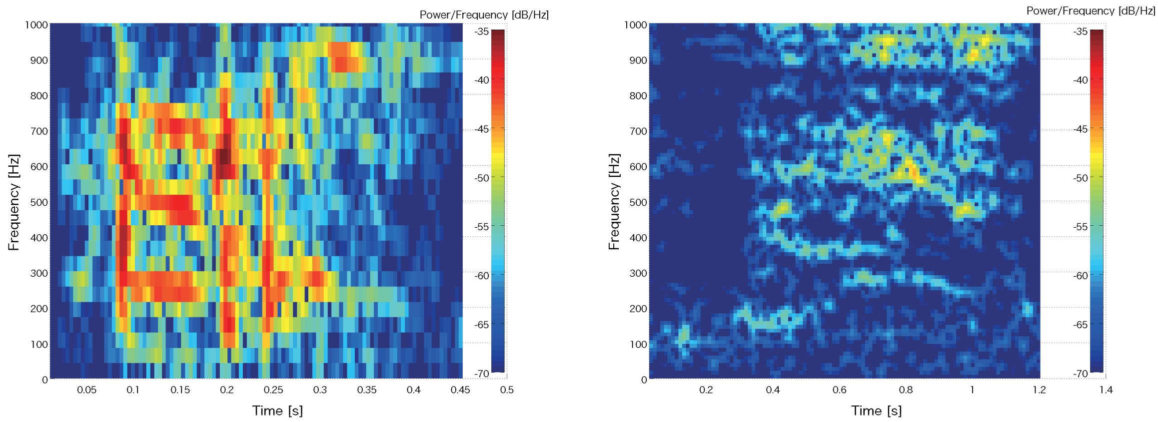
For comparison, we also analyze the same waveforms with short time Fourier transform (STFT) [23] and compare the feature of time-frequency representation. For window function of STFT, a Hamming window is chosen in this analysis. The window size is set to 4.4% of data length and 80% overlap. The spectrograms are shown in Figure 3.

From the time-frequency map of Figures 2 and 3, we can find that there are almost same properties. However, the frequency evolution in Figure 2 can be seen more clearly than that in Figure 3.



(a) Class 0: HHT map from Figure 1(a) (b) Class 1: HHT map from Figure 1(b)

Figure 2. Time-frequency-amplitude (HHT) maps obtained from IMF1-8. The color bar shows the value of IA. Note that the time scale between (a) and (b) is different.



(a) Class 0 (b) Class 1

Figure 3. Spectrograms obtained from waveforms of top panel of Figures 1(a) and 1(b). Note that the time scale between (a) and (b) is different.

Snoring sounds are the mixture of the nonlinear oscillation sounds of the oropharyngeal soft tissues and the airflow noises during inhalation. In addition, the dynamics is changing gradually or suddenly as time passes. This phenomenon can easily be understood by seeing the top panels of Figure 1, where the waveform is suddenly changing and distorted from a sinusoidal wave. From the top panels of Figure 1, the class 0 waveforms have more nonlinear and non-stationary properties than that of class 1 waveforms. And the features of class 0 waveforms are more rapidly changing than that of class 1 waveforms. In class 1 waveforms, although we can confirm the gradual change of amplitude (volume), the drastic change of waveforms cannot be confirmed. On the other hand, in the class 0 waveforms, we can confirm the drastic change of waveforms, e.g., in the top panel of Figure 1(a) at 0.08sec, 0.1∼0.18sec, 0.18∼0.20sec and 0.24sec. These non-stationary properties can be clearly confirmed in IMF3 $c_3(t)$ and IMF4 $c_4(t)$ of Figure 1(a). Thus, we find that the snoring sound waveform of class 0 is decomposed into some IMFs, which has the time localized non-stationary properties. These properties can also be seen in HHT map of Figure 2. In the case of class 0 (Figure 2(a)), we can see that the value of IF is changing in non-stationary. In Figure 3(a), we can also confirm these structures. However, we can not clearly judge the change of frequency because of the less of resolution.

From a physiological point of view, soft palate and tongue base are generally known to be the source of snoring sounds and the location of upper airway obstruction as well. Post-apneic inhalation with an explosive air inflow rapidly widens the upper airway closed by the soft palate and tongue base. The sounds which correspond to this phenomenon may indicate the drastic change of waveforms in class 0 (IMF3 $c_3(t)$ and IMF4 $c_4(t)$) at around 0.1s and 0.2s respectively. In the future, this physiological mechanism should be considered more objectively by analyzing many snoring sounds of OSAS patients.

On the other hand, IMFs of the snoring sound waveform of class 1 have almost same properties of the superposition of sinusoidal waves plus white Gaussian noise. We can also confirm these properties through Figures 1(b)-3(b).

We present the analysis results of the snoring sound waveforms shown in the top panels of Figures 1(a) and 1(b) for the most part. Although there was a difference in the degrees, we confirmed the properties discussed in this section in other class 0 and class 1 waveforms in all essentials.

4. **Summary.** In this paper, we investigated the difference of the acoustic properties of snoring sounds between non-apneic (class 1) and post-apneic (class 0) by using the Hilbert-Huang transform (HHT).

From the plot results of intrinsic mode functions (IMFs) which were decomposed from the snoring sound waveforms through EEMD process and the time-frequency-amplitude map (HHT map), we qualitatively found that the class 0 sound waveforms had more nonlinear and non-stationary properties than that of class 1 sound waveforms.

In the future, based on this research, we will investigate the method of quantitatively understanding the difference of the acoustic properties. The method and result will be discussed elsewhere. Moreover, it is necessary to develop a theoretical model to explain such phenomena from a physiological point of view.

Acknowledgments. This work was supported in part by JSPS Grant-in-Aid for Scientific Research (C) (Grant No. 17K05437; H. Takahashi and Grant No. 26330338; T. Mikami) and by JSPS Grant-in-Aid for Young Scientists (B) (Grant No. 26800129; H. Takahashi).

REFERENCES

[1] D. Pevernagie, R. M. Aarts and M. D. Meyer, The acoustics of snoring, *Sleep Medicine Reviews*, vol.14, no.2, pp.131-144, 2010.

[2] R. Beck, M. Odeh, A. Oliven and N. Gavriely, The acoustic properties of snores, *European Respiratory Journal*, vol.8, pp.2120-2128, 1995.

[3] S. J. Quinn, L. Huang, P. D. M. Ellis and J. E. Ffowcs Willams, The differentiation of snoring mechanisms using sound analysis, *Clinical Otolaryngology & Allied Sciences*, vol.21, no.2, pp.119-123, 1996.

[4] J. A. Fiz, J. Abad, R. Jane, M. Riera, M. A. Mananas, P. Caminal, D. Rodenstein and J. Morera, Acoustic analysis of snoring sound in patients with simple snoring and obstructive sleep apnoea, *European Respiratory Journal*, vol.9, pp.2365-2370, 1996.

[5] T. Emoto, U. R. Abeyratne, T. Kusumoto, M. Akutagawa, E. Kondo, I. Kawata, T. Azuma, S. Konaka and Y. Kinouchi, Discriminating apneic snorers and benign snorers based on snoring formant extracted via a noise-robust linear prediction technique, *Transactions of Japanese Society for Medical and Biological Engineering*, vol.48, no.1, pp.115-121, 2010.

[6] A. K. Ng, T. S. Koh, E. Baey, T. H. Lee, U. R. Abeyratne and K. Puvanendran, Could formant frequencies of snore signals be an alternative means for the diagnosis of obstructive sleep apnea?, *Sleep Medicine*, vol.9, no.8, pp.894-898, 2008.

[7] N. Huang, Z. Shen, S. Long, M. Wu, H. Shih, Q. Zheng, N. Yen, C. Tung and H. Liu, The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, *Proc. of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol.454, pp.903-995, 1998.

[8] J. I. Salisbury and Y. Sun, Rapid screening test for sleep apnea using a nonlinear and nonstationary signal processing technique, *Medical Engineering & Physics*, vol.29, pp.336-343, 2007.

[9] P. Caseiro, R. Fonseca-Pinto and A. Andrade, Screening of obstructive sleep apnea using Hilbert-Huang decomposition of oronasal airway pressure recordings, *Medical Engineering & Physics*, vol.32, pp.561-568, 2010.

[10] T. Mikami, S. Ueki, H. Takahashi and K. Yonezawa, Detecting nonlinear acoustic properties of snoring sounds using Hilbert-Huang transform, *Proc. of BIOSIGNALS 2015 – International Conference on Bio-Inspired Systems and Signal Processing*, Lisbon, Portugal, pp.306-310, 2015.

[11] A. Azarbarzin and Z. Moussavi, Snoring sounds variability as a signature of obstructive sleep apnea, *Medical Engineering & Physics*, vol.35, pp.479-485, 2013.

[12] T. Mikami, H. Takahashi and K. Yonezawa, Detecting nonstationary acoustic properties of post-apneic snoring sounds using the Hilbert-Huang transform, submitted to *Medical Engineering & Physics*.

[13] M. Dätig and T. Schlurmann, Performance and limitations of the Hilbert-Huang transformation (HHT) with an application to irregular water waves, *Ocean Engineering*, vol.31, pp.1783-1834, 2004.

[14] G. Rilling and P. Flandrin, One or two frequencies? The empirical mode decomposition answers, *IEEE Trans. Signal Processing*, vol.56, no.1, pp.85-95, 2008.

[15] J. Yeh, J. Shieh and N. E. Huang, Complementary ensemble empirical mode decomposition: A novel noise enhanced data analysis method, *Advances in Adaptive Data Analysis*, vol.2, no.2, pp.135-156, 2010.

[16] M. A. Colominas, G. Schlotthauer and M. E. Torres, An unconstrained optimization approach to empirical mode decomposition, *Digital Signal Processing*, vol.40, pp.164-175, 2015.

[17] B. Huang and A. Kunoth, An optimization based empirical mode decomposition scheme, *Journal of Computational and Applied Mathematics*, vol.240, pp.174-183, 2013.

[18] M. A. Colominas, G. Schlotthauer and M. E. Torres, Improved complete ensemble EMD: A suitable tool for biomedical signal processing, *Biomedical Signal Processing and Control*, vol.14, pp.19-29, 2014.

[19] Z. Wu and N. E. Huang, Ensemble empirical mode decomposition: A noise-assisted data analysis method, *Advances in Adaptive Data Analysis*, vol.1, pp.1-42, 2009.

[20] H. Takahashi, K. Oohara, M. Kaneyama, Y. Hiranuma and J. B. Camp, On investigating EMD parameters to search for gravitational waves, *Advances in Adaptive Data Analysis*, vol.5, no.2, 2013.

[21] M. Kaneyama, K. Oohara, H. Takahashi, Y. Hiranuma, T. Wakamatsu and J. B. Camp, Towards constructing an alert system with the Hilbert-Huang transform – Search for signals in noisy data, *ICIC Express Letters, Part B: Applications*, vol.5, no.1, pp.285-292, 2014.

[22] M. Kaneyama, K. Oohara, H. Takahashi, Y. Sekiguchi, H. Tagoshi and M. Shibata, Analysis of gravitational waves from binary neutron star merger by Hilbert-Huang transform, *Physical Review D*, vol.93, pp.1-11, 2016.

[23] L. Cohen, *Time-Frequency Analysis: Theory and Applications*, Prentice Hall, New York, 1995.