AN ALTERNATIVE VIEW OF LOUDSPEAKER NONLINEARITIES USING THE HILBERT-HUANG TRANSFORM

Leela K. Gudupudi[1]*, Navin Chatlani*[2]*, Christophe Beaugeant*[3] *and Nicholas Evans*[1]

[1]EURECOM, Sophia-Antipolis, France
lastname@eurecom.fr
[2]INTEL, Allentown, U.S.A.
[3]INTEL, Sophia-Antipolis, France
firstname.lastname@intel.com

ABSTRACT

This paper presents a new approach to nonlinear loudspeaker characterization using the Hilbert-Huang transform (HHT). Based upon the empirical mode decomposition (EMD) and the Hilbert transform, the HHT decomposes nonlinear signals into adaptive bases which reveal nonlinear effects in greater and more reliable detail than current approaches. Conventional signal decomposition techniques such as Fourier and Wavelet techniques analyse nonlinear distortion using linear transform theory. This restricts the nonlinear distortion to harmonic distortion. This work shows that real nonlinear loudspeaker distortion is more complex. HHT offers an alternate view through the *cumulative* effects of harmonics and intra-wave amplitude-and-frequency modulation. The work calls into question the interpretation of nonlinear distortion through harmonics and points towards a link between physical sources of nonlinearity and amplitude-and-frequency modulation. The work furthermore questions the suitability of traditional signal analysis approaches while giving weight to the use HHT analysis in future work.

Index Terms— HHT, EMD, intra-wave frequency modulation, loudspeaker nonlinearities.

1. INTRODUCTION

Nonlinear acoustic echo cancellation (NAEC) has attracted growing attention over recent years [1, 2, 3, 4, 5, 6]. Miniature loudspeakers, often used for mobile devices, are generally nonlinear systems associated with several nonlinear effects including electronic, magnetic, mechanical and sound. Approaches to NAEC depend fundamentally upon a proper understanding and estimation of these nonlinear effects. Data analysis plays an integral role in scientific research and the understanding any system and/or signals.

Traditional Fourier-based data analysis methods such as the discrete Fourier transform (DFT) and the short-time Fourier transform (STFT) dominate the signal analysis field. These methods all assume linear, (short-term) stationary signals. Wavelet analysis designed to handle non-stationary data still assumes linearity. Accordingly, Fourier and wavelet methods may not be the most suitable approaches for the analysis of miniature loudspeakers.

Since they rely on a priori defined bases for data representation, Fourier-based approaches are ill-suited to the analysis of nonlinear signals; they assume the *linear* superposition of different signal components. As a consequence, the energy of a nonlinear signal is spread across a number of harmonics. Nonlinear distortion is then represented as harmonic distortion, even if the link to a physical source is questionable.

Huang et al. proposed a new approach to signal analysis referred to as the Hilbert-Huang Transform (HHT) [7]. Based on empirical mode decomposition (EMD), the approach is well-suited to the analysis of nonlinear, nonstationary signals. Unlike traditional approaches, EMD adapts the bases to the signal itself and can therefore yield more physically relevant results. HHT analysis leads to a new physical interpretation of nonlinear distortion. Huang et al. argue that a priori defined bases impose harmonics and that these are nothing more than a mathematical artifact, with no link to a physical source [7, 8]. In place of harmonic distortion is the concept of amplitude and frequency modulation. A brief description of HHT analysis is presented in 2. More detailed presentations can be found in [7, 8, 9].

Among other applications of HHT and/or EMD analysis is work in speech enhancement [10, 11], source separation [12], oceanography [13], vibration analysis [14], EEG/ECG enhancement [15, 16] and climate analysis [17]. Our own work investigated an EMD-based approach to NAEC [1]. This paper reports our first attempt to apply HHT to the analysis of nonlinear distortion produced by miniature loudspeakers. The work aims to provide an alternative approach to signal analysis which goes beyond Fourier-based approaches while avoiding the limitations to harmonic distortion. The work is our first steps to align the analysis of nonlinear distortion to its physical origins. The paper reports the application of HHT to the analysis of real mobile phone loudspeaker signals.

2. THE HILBERT-HUANG TRANSFORM

The Hilbert-Huang Transform (HHT) is a signal analysis approach which is well-suited to nonlinear, nonstationary signals [7, 8]. The application of HHT involves two steps. The first decomposes a discrete time-domain signal $y(n)$ into a set of M intrinsic mode functions (IMFs), $y_j(n)\,;\ j = 1, \cdots, M$, using empirical mode decomposition (EMD) such that:

$$y(n) = \sum_{j=1}^{M} y_j(n) + r(n) \tag{1}$$

where $r(n)$ is the residue. The second step determines the instantaneous frequency (IF) and instantaneous amplitude (IA) of each IMF y_j using the Hilbert Transform. From these, one can construct straightforwardly the time-frequency-energy distribution referred to as the Hilbert spectrum [7, 8].

2.1. The Hilbert transform

Reverting temporarily to continuous notation, for any arbitrary time series, $X(t)$, the Hilbert Transform (HT), $Y(t)$, is obtained as follows [18]:

$$Y(t) = \frac{1}{\pi} P \int_{-\infty}^{+\infty} \frac{X(\tau)}{t-\tau} d\tau \tag{2}$$

where P denotes the Cauchy principal value. With this definition, $X(t)$ and $Y(t)$ form a complex conjugate pair leading to an analytical signal:

$$Z(t) = X(t) + jY(t) = a(t)e^{j\theta(t)} \tag{3}$$

in which

$$\begin{aligned} a(t) &= \sqrt{(X^2(t) + Y^2(t))} \\ \theta(t) &= \arctan \frac{Y(t)}{X(t)} \end{aligned} \tag{4}$$

Here, $a(t)$ is the instantaneous amplitude (IA) and $\theta(t)$ is the instantaneous phase. The instantaneous frequency (IF) can be computed as:

$$\omega(t) = \frac{d\theta(t)}{dt} \tag{5}$$

This classical wave theory-styled definition of IF is computed through differentiation rather than integration. The IF is local, not global, and reflects intra-wave frequency modulation [7]. Cohen [19] showed that the HT produces meaningful IF only for *monocomponent* signals while the Bedrosian and Nuttall theorems [20, 21] impose further constraints, e.g. non-overlapping spectra of $a(t)$ and $cos(\theta(t))$.

These limitations are avoided through the use of IMFs which satisfy the following two requirements [7, 8]:

- the number of extrema in an IMF (the sum of the maxima and the minima) and the number of zero-crossings must either be equal or differ at most by one, and
- at any point, the mean value of the envelop defined by the local maxima and the envelop defined by the local minima shall be zero.

Unfortunately, most practical data do not meet these requirements. As a result, the full potential of the HT had to wait for the development of empirical mode decomposition (EMD).

2.2. Empirical Mode Decomposition

EMD is an approach to the analysis of non-linear, non-stationary signals [7, 9]. EMD decomposes any signal into a finite number of IMFs. IMFs are essentially basis functions, but are not predefined as is the case in Fourier and wavelet analysis. Instead, they are extracted adaptively from the input data.

The EMD algorithm examines the input signal between two consecutive extrema and iteratively extracts the highest frequency components between these two points [9]. The remaining local, lower frequency components can then be extracted through consecutive iterations.

As described in [7, 10], a signal $y(n)$ is decomposed into a set of M IMFs according to the following procedure known as *sifting*:

1. Identify all extrema (local maxima and minima) of the signal, $y(n)$.
2. Obtain the upper envelope $e_{max}(n)$ and the lower envelope $e_{min}(n)$ by interpolating the local maxima and minima, respectively.
3. Compute the local mean $m(n) = \frac{e_{min}(n)+e_{max}(n)}{2}$.
4. Extract the detail signal $d(n) = y(n) - m(n)$.
5. $d(n)$ is an IMF if it has zero mean and all its local maxima and minima are positive and negative respectively. If not, steps 1–4 are repeated with $d(n)$ in place of $y(n)$.
6. For the next IMF, the entire process is applied to the residual $r_1(n) = y(n) - d(n)$.
7. Iterate on the residual until the number of extrema in the residual is smaller than 2 or until a maximum number of iterations is reached. Assign the last residual as $r(n)$.

The above *sifting* process decomposes any signal $y(n)$ into a set of successively lower frequency IMF components $y_j(n)$; $j = 1, \cdots, M$. Together they represent $y(n)$ according to Eq. 1. A full description of EMD is available in [7].

2.3. Hilbert-Huang Spectrum

The HT is readily applied to each IMF in order to determine the IA ($a_j(n)$; $j = 1, \cdots, M$) and IF ($\omega_j(n)$; $j = 1, \cdots, M$) according to Eqs. 4 and 5 respectively. The analytic representation of the input signal may then be expressed as:

$$y'(n) = \sum_{j=1}^{M} a_j(n) e^{i \int \omega_j(n) dn} \tag{6}$$

where, since it is constant, the residue $r(n)$ is omitted. The original input signal, $y(n)$, is the real part of the analytic signal. The IAs ($a_j(n)$; $j = 1, \cdots, M$) and IFs ($\omega_j(n)$; $j = 1, \cdots, M$) then give a time-frequency-amplitude representation of the signal, termed the Hilbert-Huang Spectrum [7, 8]. A plot of the time-frequency distribution of IA^2 (square the amplitude) illustrates the energy density in similar fashion to a conventional spectrogram.

2.4. Relation to Fourier techniques

Expressed as a sum of sinusoids, the input signal is given by:

$$y'(n) = \sum_{j=1}^{\infty} a_j e^{i\omega_j n} \tag{7}$$

where a_j and ω_j are constant amplitude and frequency terms respectively. Because the frequency of each sinusoidal function is time-independent, Fourier analysis is able to construct stationary data only. Also, since the sine waves used to describe a signal are infinite in extent, Fourier analysis is considered a global analysis tool. The accuracy thus depends critically on data length and stationarity, yet practical data is generally short in existence and of arbitrary duration.

The comparison of Eqs. 6 and 7 show that the HHT is a generalised Fourier expansion but with time-varying amplitude and frequency which accommodate nonlinear, nonstationary data. The Fourier representation implies constant energy at a given frequency, i.e. a regular harmonic wave which persists unchanged throughout the full data record. HHT analysis, in contrast, reflects the *local* likelihood of energy at a given frequency.

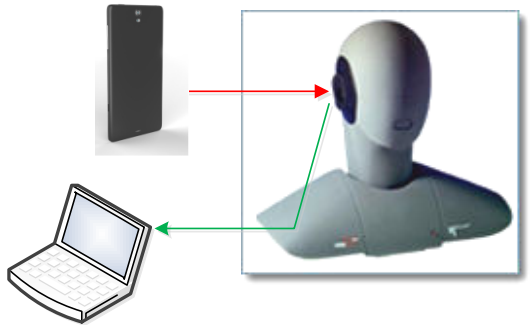

Figure 1: Experimental setup in an anechoic chamber to measure loudspeaker outputs.

3. LOUDSPEAKER DISTORTION ANALYSIS

This section reports the application of HHT to the analysis of nonlinear loudspeaker distortion. This work was performed using real mobile phone loudspeaker recordings. Before that, the physical sources of nonlinear distortion in a typical loudspeaker enclosure microphone system (LEMS) and recording set-up are described.

3.1. Sources of nonlinearity

It is widely accepted that the LEMS down-link path is the most significant contributor to nonlinear distortion [4, 6, 22]. It encompasses a digital-to-analog converter (DAC), an analog power amplifier and the loudspeaker. Principle sources of nonlinear distortion include:

- high-resolution DACs: fabrication variabilities cause amplitude, pulse shape and timing errors in the DAC output which amount to nonlinear distortion;
- power amplifiers operating close to saturation in order to produce high output signal levels from low electrical power: the result is nonlinear clipping distortion;
- the miniature loudspeaker itself is a major source of nonlinearities.

As described in [23], the latter is highly complex with multiple causes:

- When the loudspeaker is driven with a constant current, the force on the voice coil $F_v = Bl(x_{po})$ varies and depends on its position x_{po}. If the voice coil moves toward the air gap then the magnetic flux density B increases. If it moves away from the air gap, then B decreases. The non-uniform magnetic flux density affects the driving force on the voice coil, hence causing non-linear distortion. Since the length of the voice coil l is constant, the non-linear function between B and F_v is static and can hence be modeled as a power series expansion.
- The self-inductance of the voice coil L_e also depends on its position causing non-linear distortion.
- Loudspeakers use a suspension system, comprising a spider and a surround to center the voice coil in the air gap. The suspension behaves like a normal spring and may be characterized by the force-displacement curve, which often exhibits hysteresis. This is due to the non-linear stiffness of the spider which is not constant but a function of voice coil displacement (x_{po}).

3.2. Experimental set-up

The nonlinear response of a loudspeaker is observed from its output to a single sinusoidal excitation signal. This approach was used to characterize a real mobile phone loudspeaker placed before a head and torso mannequin at a distance of 30cm in an anechoic chamber. The experimental set-up used is illustrated in Fig. 1. The device is configured to operate in hands-free mode and at maximum volume at which nonlinear distortion is greatest. Input signals sampled at 48kHz are pure sinusoids with frequencies between 100Hz and 3800Hz in 100Hz intervals. They are stored in mobile phone memory and played back using a pre-installed VLC player. Loudspeaker outputs are recorded with a high-quality (linear) microphone mounted in the mannequin ear. Recorded signals are stored on a PC at the same 48kHz sampling frequency.

3.3. HHT Analysis

As an example we consider a real mobile phone loudspeaker subjected to a single sinusoidal excitation of frequency 1kHz. Fig. 3(a) shows the results of STFT analysis. Several high-order harmonics are visible, representing the traditional view of nonlinear distortion.

Fig. 3(b) illustrates the four (out of eight) IMFs which result from decomposition of the loudspeaker signal using EMD and the routines available in [24]. Since EMD extracts the highest-frequency IMF first, IMF-1 is the distorted harmonic caused by loudspeaker nonlinearities. IMF-2 is the distorted natural frequency at 1kHz whereas the other IMFs have negligible energy.

Fig. 3(c) illustrates the IA profiles of the four IMF components which exhibit intra-wave amplitude modulation, namely variation in amplitude across time. Fig. 3(d) illustrates the corresponding IF profiles which exhibit intra-wave frequency modulation. This is due to the displacement of the loudspeaker diaphragm which is no longer a pure sinusoidal function on account of nonlinear distortion. A relatively strong third-order harmonic is also generated as a result of asymmetrical loudspeaker nonlinearities.

The wave-profile deformation caused by the nonlinear distortion is the result of accumulated harmonic content and intra-wave amplitude-and-frequency modulation. This *cumulative* effect is observed in the time domain response of the loudspeaker shown in Fig. 2. The waveform deformation is not constant, but varies from high to low and vice versa in accordance with the IA profile in Fig. 3(c). The extent of the deformation depends on the magnitude of the additional harmonics and the strength of the intra-wave amplitude-and-frequency modulation. Close observation of IA and IF profiles in Figs. 3(c) & (d) respectively shows that the frequency variation of the IMF components increases when their amplitude decreases and vice versa. This is indicative of *softening* nonlinearity [25].

The effects described above are not reflected in the traditional STFT spectrogram which instead shows spurious harmonics. HHT-derived estimates may thus reflect more reliably nonlinear behavior than STFT-derived estimates. The real question, however, is whether the frequency and amplitude modulation illustrated in Fig. 3 have a real, physical source or whether they are simply an artifact of the HHT.

Despite perfect reconstruction, HHT still suffers from so-called *end effect* artifact's [7]. The cubic spline fitting to local extrema in the EMD process is error prone, especially due to discontinuities at signal extremities. As a result, Gibbs phenomenon is induced upon the application of the HT to each IMF [7, 8]. This effect is observed in Figs. 3 and 4. Some solutions to avoid these artifacts

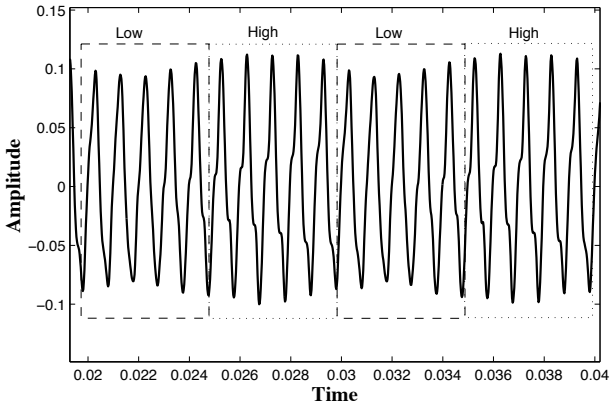

Figure 2: A real mobile phone loudspeaker response to 1kHz pure sine tone. The wave-profile deformation caused by the nonlinear distortion is not constant throughout the time.

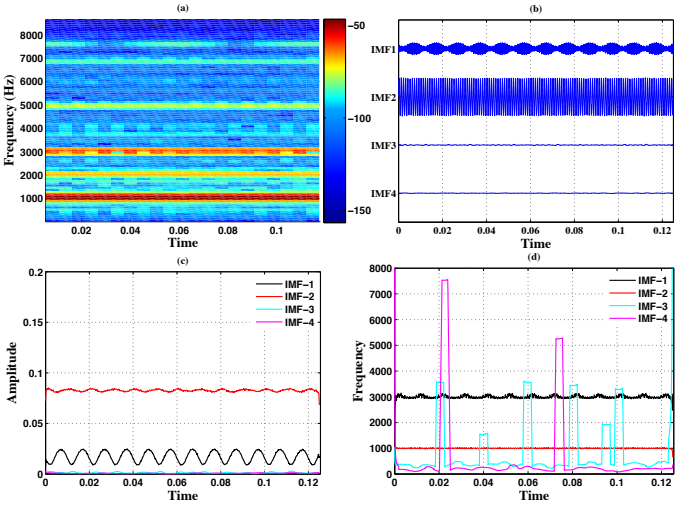

Figure 3: (a) STFT spectrogram of a mobile phone loudspeaker response to a pure sinusoidal input at 1kHz, sampled at 48kHz; (b) Loudspeaker response to 1kHz sine tone is decomposed by the EMD, resulting in the 8 IMFs, first 4 IMFs are listed above and others are not displayed since they are almost zero; (c) IA profiles of the IMFs obtained by HHT; (d) IF profiles of the IMFs obtained by HHT

are proposed in the literature, e.g. [26], but are beyond the scope of the present paper.

4. VALIDATION OF HHT

The HHT is thoroughly validated in [7] with analytical examples. This section aims to validate the HHT technique as a means of characterizing nonlinear loudspeaker behavior. Figs. 4(a) and 4(b) illustrate the spectrogram and IF profiles of a mobile phone loudspeaker response to a pure sinusoidal input at 500Hz. The IF profiles show *cumulative* harmonic and modulation nonlinear distortion. There is only a weak third-order harmonic and significant intra-wave amplitude-and-frequency modulation, whereas the spectrogram shows a strong third order harmonic and several, weak harmonics. Fig. 4(c) shows the corresponding IF profiles when the mobile phone loudspeaker is replaced with a high-quality (linear) loudspeaker and shows a total absence of amplitude-and-frequency modulation. Fig. 4(d) shows the IF profiles of the (high-quality) loudspeaker output when a simulated, 3rd order harmonic distortion is added to the input. Once again, there is no amplitude and frequency modulation indicating that the distortion observed in (b) has physical origins and is not simply an artifact of HHT processing.

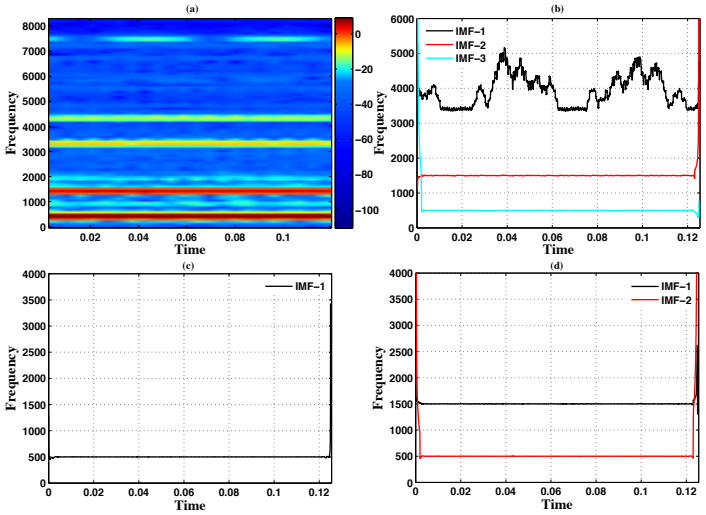

Figure 4: Time-frequency-energy distributions: (a) STFT spectrogram of a mobile phone loudspeaker response to a pure sinusoidal input at 500Hz, sampled at 48kHz; (b) the IF profiles obtained by HHT; (c) IF profiles for a high-quality loudspeaker response to the same input; (d) the IF profiles of the same loudspeaker subject to an input excitation comprised of pure sinusoidal at 500Hz and its third harmonic.

5. CONCLUSIONS

This paper investigates the application of the Hilbert-Huang Transform (HHT) to the characterization of nonlinear loudspeakers. This approach gives an alternative view of loudspeaker nonlinear behavior. The waveform deformation caused by the nonlinear distortion is the result of a *cumulative* effect, namely that of harmonics and intra-wave amplitude-and-frequency modulation, instead of the pure harmonic interpretation which results from Fourier treatments.

The application of this technique to analyse loudspeaker nonlinearities provides an alternative view which supports the exploration of different nonlinear phenomena: quadratic, cubic or higher-order, softening and hardening effects, intra-wave amplitude-and-frequency modulation and distorted harmonic responses etc. However, this work is in its infancy and hence the associated properties of the HHT technique and their impact on the analysis requires further exploration. Also, the discontinuity-induced Gibbs phenomenon at data extremities needs further study in order to fully understand the new approach. This is the subject of ongoing work.

6. REFERENCES

[1] L. K. Gudupudi, N. Chatlani, C. Beaugeant, and N. W. D. Evans, "Non-linear acoustic echo cancellation using empirical mode decomposition," in *Proc. ICASSP*, Apr. 2015.

[2] M. Z. Ikram, "Non-linear acoustic echo cancellation using cascaded Kalman filtering," in *Proc. ICASSP*, May 2014.

[3] C. Huemmer, C. Hofmann, R. Maas, A. Schwarz, and W. Kellermann, "The elitist particle filter based on evolutionary strategies as novel approach for nonlinear acoustic echo cancellation," in *Proc. ICASSP*, May 2014.

[4] J. M. Gil-Cacho, M. Signoretto, T. van Waterschoot, M. Moonen, and S. H. Jensen, "Nonlinear acoustic echo cancellation based on a sliding-window leaky kernel affine projection algorithm," *IEEE Transactions on Audio, Speech, and Language Processing*, Sept. 2013.

[5] P. Shah, I. Lewis, S. Grant, and S. Angrignon, "Nonlinear acoustic echo cancellation using feedback," in *Proc. ICASSP*, May 2013.

[6] M. I. Mossi, C. Yemdji, N. W. D. Evans, C. Beaugeant, and P. Degry, "Robust and low-cost cascaded non-linear acoustic echo cancellation," in *Proc. ICASSP*, May 2011.

[7] N. E. Huang, Z. Shen, *et al.*, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 454, Mar. 1998.

[8] P. Flandrin, P. Gonalves, and G. Rilling, *Hilbert-Huang Transform and its Applications*. World Scientific, 2005.

[9] G. Rilling, P. Flandrin, and P. Gonçalves, "On empirical mode decomposition and its algorithms," in *Proc. NSIP*, June 2003.

[10] L. Zao, R. Coelho, and P. Flandrin, "Speech Enhancement with EMD and Hurst-Based Mode Selection," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, no. 5, May 2014.

[11] N. Chatlani and J. J. Soraghan, "EMD-Based Filtering (EMDF) of Low-Frequency Noise for Speech Enhancement," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, May 2012.

[12] B. Mijovic, M. De Vos, I. Gligorijevic, J. Taelman, and S. Van Huffel, "Source separation from single-channel recordings by combining empirical-mode decomposition and independent component analysis," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 9, Sept. 2010.

[13] P. A. Hwang, N. E. Huang, and D. W. Wang, "A note on analyzing nonlinear and nonstationary ocean wave data," *Applied Ocean Research*, vol. 25, no. 4, Aug. 2003.

[14] Z. K. Peng, W. T. Peter, and F. L. Chu, "An improved hilberthuang transform and its application in vibration signal analysis," *Journal of Sound and Vibration*, vol. 286, no. 12, Aug. 2005.

[15] X. Navarro, F. Poree, and G. Carrault, "Ecg removal in preterm eeg combining empirical mode decomposition and adaptive filtering," in *Proc. ICASSP*, Mar. 2012.

[16] C. Gaochao, Y. Yunchao, T. Tanaka, and C. Jianting, "Eeg energy analysis for evaluating consciousness level using dynamic memd," in *Proc. IJCNN*, July 2014.

[17] H. En-Ching, "Hilbert -huang transform analysis of hydrological and climatic time series," Ph.D. dissertation, Purdue University, 2006.

[18] F. W. King, *Hilbert Transforms*. Cambridge University Press, 2009.

[19] L. Cohen, *Time-frequency analysis*. Prentice-Hall, 1995.

[20] E. Bedrosian, "A product theorem for hilbert transforms," *Proc. IEEE*, vol. 51, no. 5, May 1963.

[21] A. H. Nuttall, "On the quadrature approximation to the hilbert transform of modulated signals," *Proc. IEEE*, vol. 54, no. 10, Oct. 1966.

[22] F. Kuech, A. Mitnacht, and W. Kellermann, "Nonlinear acoustic echo cancellation using adaptive orthogonalized power filters," in *Proc. ICASSP*, Mar. 2005.

[23] W. Klippel, "Loudspeaker nonlinearities - causes, parameters, symptoms," in *Audio Engineering Society Convention 119*, Oct. 2005.

[24] P. Flandrin *et al.*, "Matlab/C codes for EMD and EEMD with examples," 2007, [Online]. [Online]. Available: http://perso.ens-lyon.fr/patrick.flandrin/emd.html

[25] A. Nayfeh, *Introduction to perturbation techniques*. Wiley, 1981.

[26] L. Da-Chao, L. Zhang, P. Feng, and L. Fan, "Elimination of end effects in empirical mode decomposition by mirror image coupled with support vector regression," *Mechanical Systems and Signal Processing*, vol. 31, 2012.