Science Arena Publications
Specialty journal of Engineering and Applied Science
Available online at www.sciarena.com
2016, Vol, 1 (2): 71-75

# A New Stimulating Algorithm for Cochlear Implant

**Marjan Mozaffari Legha[1], Arash Nowroozi[2], Mahsa Mirshekari[3]**
[1]*Medical Physics & Biomedical Engineering Department, Shahid Beheshti University of Medical Sciences, Tehran, Iran. Email:* Mozaffarilegha.m@gmail.com
[2]*Department of Power Engineering, Institute of Higher Education Javid, Jiroft, Iran.*
*Email:* arash59ns.ms@gmail.com
[3]*Department of Power Engineering, Institute of Higher Education Javid, Jiroft, Iran.*
*Email:* Mahsa.mirshekari70@gmail.com

***Abstract:*** *Cochlear implants are being widely used for the patients with severe to profound sensorineural hearing loss. Speech coding algorithms play an important role in improving the performance of cochlear implant. At recent years, the performance of CI has been improved for most users under silent environment. However, as the background noise level increases, speech recognition scores are degraded considerably. In this paper, the Empirical Mode Decomposition and Teager-Kaiser Energy Operator are applied as a speech enhancement method for cochlear implants. This algorithm is developed to extract features, called intrinsic mode functions, by a sifting process. Then, Frequency and amplitude of each IMF is extracted based on TKEO. Finally, performance of this algorithm in terms of correlation analysis was compared to continuous interleaved sampling (CIS), frequency amplitude modulation encoding (FAME) and Hilbert Huang Transform Stimulating (HHTS) strategies. The results showed the highest correlation coefficient between spectrum of synthesized signal and original speech with proposed method.*

***Keyword*** *: Algorithm, Cochlear, Server, performance*

## INTRODUCTION

Cochlear implant (CI) is an electronic prosthetic device surgically implanted into the inner ear for restoring some degree of hearing of profoundly deaf patients with sensory-neural origin [1]. It includes internal and external components. The external part consists of a microphone that picks up sound, a signal processor that converts sound into electric impulses, and a transmitter that is magnetically attached to the internal device to which it transmits the electric impulses via radio waves. The impulses are sent to an array of electrodes, which are surgically inserted into the cochlea. The electrodes stimulate the auditory nerves, providing auditory information to the brain [2]. Most CI users achieve 80% word recognition scores in quiet listening conditions [3]. However, speech recognition scores are degraded in noisy conditions [4]. Several studies have been proposed to develop speech processing techniques for CI. In CIS strategy, envelope characteristics of speech signal are extracted [5]. It utilizes a filter-bank for the frequency decomposition of incoming speech which is a simplification of frequency decomposition function of biological cochlea. Outputs from each channel of the filter-bank are used to modulate the amplitudes of electrical stimulation pulses. In FAME strategy, envelope, frequency and phase information are extracted. This algorithm provides too much indiscriminate information. These techniques are not successful in providing time and frequency resolutions at the same time. Wavelet Transform (WT) overcome the limitations of the previous methods by providing both time and frequency resolutions [6]. However, it suffers to analyze non-stationary signals like speech and depends on the basis wavelet. In the last decade, a new nonlinear technique, termed empirical mode decomposition

(EMD), has been introduced by N. E. Huang et al. [7] for adaptively representing non stationary signals. The most important characteristic of EMD is that the basis functions are directly derived from the speech signal itself. HHTS strategy [8] is used to encode both temporal envelope and instantaneous frequency (IF) of input speech signal for CI. While this strategy has allowed cochlear implant users to achieve good speech recognition in quiet, their performance in noisy condition is severely compromised. Also, an IF has a true meaning only for mono component signals, where there is only one frequency or at least a narrow range of frequencies varying as a function of time. This definition of IF is doubtful and will mislead analysis of instantaneous frequency, such as negative frequency [9]. In this paper, we propose a new algorithm based on the EMD technique and Teager-Kaiser Energy Operator (TEO) applied to modes. TEO can track the energy and detect the instantaneous frequency and instantaneous amplitude of mono-component AM-FM signal [10]. The basic idea is to reconstruct the signal with using TEO to extract amplitude and IF of each IMF. This paper is outlined as follows. Section 2 describes theoretical overview of EMD (Section 2.1), TEO (Section 2.2), and stimulation of new algorithm for CIs (Section 2.3) are described. Section 3 covers the results. Section 4 devotes to the conclusion.

## 1. Material and Methods

### 2.1. Theoretical overview of EMD

The Empirical Mode Decomposition (EMD) has been proposed as an adaptive time-frequency data analysis method [7]. This adaptive technique is derived from the simple assumption that any signal consists of different intrinsic mode functions (IMFs) each of them representing an embedded distinctive oscillation on a separated time-scale. An IMF is defined by two criteria: i) the number of extrema and of zero crossings must either equal or differ at most by one, and, ii) at any instant in time, the mean value of the envelope defined by the local maxima and the envelope of the local minima is zero. The following plan offers an idea about the principle algorithm of the EMD:

1. **Initialize** $r_0(t) = x(t); j = 1$

2. **Extract** the $j\_th$ IMF:

(a) **Initialize** $h_0(t) = r_j(t); k = 1$

(b) Locate local **maxima** and **minima** of $h_{k-1}(t)$

(c) Cubic **spline interpolation** to define upper and lower envelope of $h_{k-1}(t)$

(d) Calculate **mean** $m_{k-1}(t)$ from **upper and lower envelope** of $h_{k-1}(t)$

(e) **Define** $h_k(t) = h_{k-1}(t) - m_{k-1}(t)$

(f) If **stopping criteria** are satisfied then $h_j(t) = h_k(t)$ else go to 2. (b) with $k = k + 1$

3. **Define** $r_j(t) = r_{j-1}(t) - h_j(t)$

4. If $r_j(t)$ still has at least two extrema then go to 2. (a) with $j = j + 1$ else the EMD is finished

5. $r_j(t)$ is the residue of $x(t)$

At the end of this numerical sifting process the signal $x(t)$ can be expressed:

$$x(t) = \sum_{j=1}^{n} h_j(t) + r_n(t)$$

Where $h_j(t)$ indicates the $j\_th$ IMF, $n$ as the number of sifted IMF and $r_n(t)$ denotes a residue which can be understood as the trend of the signal.

### 1.2. Teager Kaiser Energy Operator (TKEO)

It is shown that the TKEO can track the energy and identify the instantaneous frequency (IF) and the amplitude of a signal [10]. The TKEO, ψ (.) is defined for continuous-time signal:

$$\psi[\mathrm{x(t)}] = [\dot{x}(t)]^2 - x(t)\ddot{x}(t)$$

where $\dot{x}(t)$ and $\ddot{x}(t)$ are the first and the second time derivatives of $x(t)$ respectively. In the discrete case, the time derivatives may be approximated by time differences. The discrete-time counterpart of TKEO becomes:

$$\psi[\mathrm{x(t)}] = x^2(n) - x(n+1).x(n-1)$$

An essential characteristic of TKEO is that it is approximately instantaneous. This is because only three samples are required for the energy computation at each time instant. This time resolution provides us with ability to capture the energy fluctuations. Furthermore, implementation of this operator is very easy. The Energy separation algorithm (ESA) [11] uses the TKEO to separate $x(t)$ into its amplitude envelope $a(t)$ and IF signal $f(t)$ to achieve mono-component AM-FM signal demodulation:

$$f(t) \approx \frac{1}{2\pi}\sqrt{\frac{\Psi[\dot{x}(t)]}{\psi[x(t)]}}$$

$$|a(t)| \approx \frac{\psi[x(t)]}{\sqrt{\psi[\dot{x}(t)]}}$$

The ESA is less computationally complex and has better time resolution than other demodulation methods such as the Hilbert transform.

### 1.3. Stimulation of New Algorithm for CIs

Figure 1 is a block diagram representing acoustic synthesis of proposed algorithm. One of the first processing steps in cochlear implants is to apply pre-emphasis to the signal. The pre-emphasis filter attenuates low frequencies and amplifies high frequencies, to compensate for the typical 6 dB/octave spectral roll-off of speech signals. It makes the low-energy, high-frequency consonants to stand out better against the high-energy, low-frequency vowels.

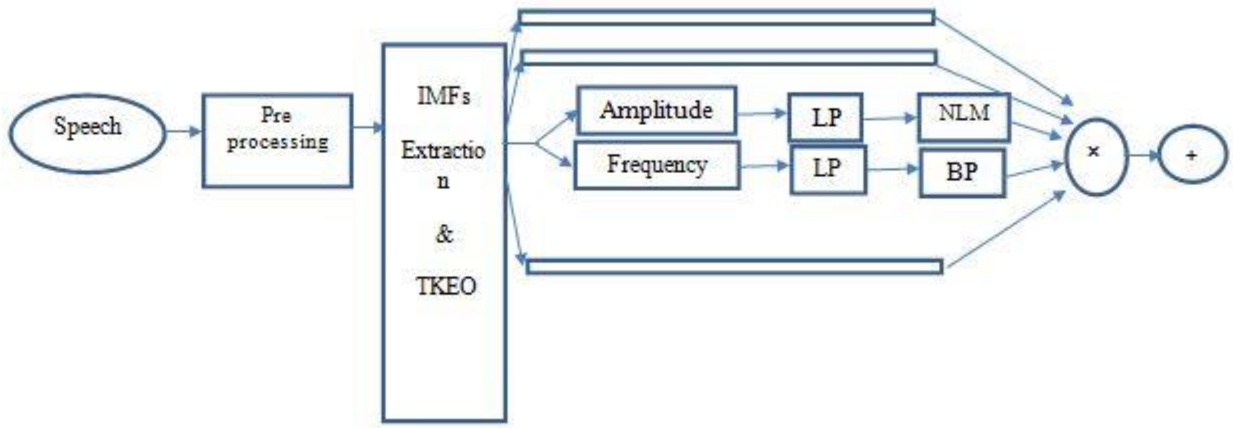


Figure 1- Block diagram of New Algorithm for CIs

Then the signal is processed through empirical mode composition, in each band, using the TKE operator, the IF and amplitude of each of them are calculated. The envelopes of the derived IMFs are then extracted by low-pass filtering (LP).The cutoff frequency of LP is typically 400 Hz. Finally, envelope matching is needed to map the decomposed signal to the dynamic range of the human ear. For this purpose, a nonlinear logarithmic function (NLM, as shown in figure 1) is used [12]. At the same time, in another path the frequency is derived from each band. After low pass filtering (LP), the frequency depth of processed signal in each band is limited at about 500Hz. At last, synthesized speech signal could obtain by summarizing each sub-band's stimuli.

## 2. Results

The cross-correlation between spectrums of synthesized and original signals was calculated, showing the power of this method and its capability in representation of a high percentage of the original signal for the implant user. Correlation coefficients were obtained in different environmental conditions (quiet, 5dB, 10dB, 15dB). In this computer simulation, Noisy92 sentences as database were processed by CIS, FAME, HHTS and proposed algorithm. Table 1 shows correlation coefficient between spectrum of reconstructed signal and original one which is deteriorating for each speech coding algorithm with noise increasing. This indicated that regardless of the type of algorithm, the efficiency of encoding can be reduced by increasing of noise. The performance of proposed algorithm is better than other three algorithms in four different environmental conditions.

Table1-Mean values of absolute correlation coefficients $\bar{r}$

| **Listening condition** | **CIS** | **FAME** | **HHT** | **Proposed algorithm** |
|---|---|---|---|---|
| Quiet | 0.0892 | 0.3174 | 0.4347 | 0.5320 |
| 5 dB | 0.0041 | 0.3060 | 0.4122 | 0.4621 |
| 10 dB | 0.0038 | 0.2974 | 0.4009 | 0.4103 |
| 15 dB | 0.0036 | 0.2795 | 0.3793 | 0.4021 |

## 4. Conclusion

In this paper, we presented a new algorithm based on the EMD technique to decompose the input signal into different frequency bands and Teager-Kaiser Energy Operator (TEO) applied to modes to extract amplitude and IF of each IMF. The TEO is less computationally complex and has better time resolution than other demodulation methods such as the Hilbert transform. Reconstruction of the decomposed signal showed that our technique can produce the processing with higher correlation than other methods.

## References

[1] K. Nie, G. Stickney, and F.-G. Zeng, "Encoding frequency modulation to improve cochlear implant performance in noise," *Biomedical Engineering, IEEE Transactions on,* vol. 52, pp. 64-73, 2005.

[2] P. C. Loizou, "Mimicking the human ear," *Signal Processing Magazine, IEEE,* vol. 15, pp. 101-130, 1998.

[3] A. J. Spahr and M. F. Dorman, "Performance of subjects fit with the Advanced Bionics CII and Nucleus 3G cochlear implant devices," *Archives of Otolaryngology–Head & Neck Surgery,* vol. 130, pp. 624-628, 2004.

[4] M. K. Qin and A. J. Oxenham, "Effects of simulated cochlear-implant processing on speech reception in fluctuating maskers," *The Journal of the Acoustical Society of America,* vol. 114, pp. 446-454, 2003.

[5] B. S. Wilson, D. T. Lawson, M. Zerbi, C. C. Finley, and R. D. Wolford, "New processing strategies in cochlear implantation," *Otology & Neurotology,* vol. 16, pp. 669-675, 1995.

[6] S. Sunny, S. David Peter, And K. Poulose Jacob, "Recognition Of Speech Signals: An Experimental Comparison Of Linear Predictive Coding And Discrete Wavelet Transforms," *International Journal Of Engineering Science & Technology,* Vol. 4, 2012.

[7] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng*, 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, pp. 903-995, 1998.

[8] H. Liu, W. Wang, K. Li, and Z. Zhang, "A novel speech coding algorithm for cochlear implants," in *Biomedical Engineering and Informatics (BMEI), 2012 5th International Conference on*, 2012, pp. 403-406.

[9] L. Angrisani and M. D'Arco, "A measurement method based on a modified version of the chirplet transform for instantaneous frequency estimation," *Instrumentation and Measurement, IEEE Transactions on,* vol. 51, pp. 704-711, 2002.

[10] J. Cexus and A. Boudraa, "Teager-Huang Analysis Applied to Sonar Target Recognition," *International Journal of Signal Processing,* vol. 1, pp. 1-5, 2004.

[11] P. Maragos, J. F. Kaiser, and T. F. Quatieri, "Energy separation in signal modulations with application to speech analysis," *Signal Processing, IEEE Transactions on,* vol. 41, pp. 3024-3051, 1993.

[12] Q.-J. Fu and R. V. Shannon, "Effect of stimulation rate on phoneme recognition by Nucleus-22 cochlear implant listeners," *The Journal of the Acoustical Society of America,* vol. 107, pp. 589-597, 2000.