

# HOWLING DETECTION IN HEARING AIDS USING DISCRETE ENERGY SEPARATION ALGORITHM-2 AND GENERALIZED TEAGER-KAISER OPERATOR

Soudeh A. Khoubrouy, Issa M.S. Panahi, *IEEE, Senior Member*, Nasser Kehtarnavaz, *IEEE, Fellow*

Department of Electrical Engineering, University of Texas at Dallas

## ABSTRACT

Howling is one of the most annoying consequences of the acoustical coupling in hearing aids. In this paper, a method of howling detection is proposed applying Discrete Energy Separation Algorithm-2 (DESA-2) and Generalized Teager-Kaiser Operator (GTKO). Several GTKOs with various resolution parameters monitor the input signal to recognize the howling occurrence while DESA-2 is used to estimate the howling frequency. Performance of the proposed method is compared with two known howling detection approaches, i.e. Peak-to-Harmonic Power Ratio (PHPR) and Adaptive Notch Filter (ANF). Simulation results show the advantage of the proposed method in terms of having lower false detection and shorter detection time.

**Index Terms**— Hearing aid, howling detection, Teager-Kaiser operator, discrete energy separation

## 1. INTRODUCTION

The small size of many hearing aid devices allows a signal leakage, called acoustic feedback, between the loudspeaker and the microphone(s) [1]. One of the serious effects caused by acoustic feedback is howling which should be detected and cancelled.

Existing howling detection methods could be classified into two major categories, i.e. frame-based and sample-based methods. The former group processes the Short Time Fourier Transform (STFT) of the input signal and typically checks some frequency domain properties of the howling [2]. Peak to Harmonic Power Ratio (PHPR) is one of the known methods in this category [2], [3]. The latter group processes the input signal sample by sample in time domain, e.g. Adaptive Notch Filter (ANF) method [4].

Due to some similarities howling frequency component cannot be easily distinguished from speech formants (spectral peaks) [2]. Thus, the accuracy of the howling detection is degraded because of detecting formants as howling frequencies (false detection).

This paper presents an approach for howling detection taking advantage of two methods derived from Teager-Kaiser Operator (TKO), i.e. Discrete Energy Separation Algorithm-2 (DESA-2) [5] and Generalized TKO (GTKO) [6]. TKO was first proposed by Teager [7] and investigated by Kaiser [8]. DESA-2 developed by Maragos et al. using nonlinear combination of TKOs with two different inputs and GTKO is a variation of TKO in which the resolution parameter is not necessarily one [6].

In this paper DESA-2 is applied to estimate the howling frequency. This algorithm is normally used to estimate the amplitude and frequency of an AM-FM signal [5]. Since a speech signal can be represented by AM-FM models around its formants, DESA-2 is also applied to estimate the frequencies and amplitudes of the formants or resonances in the speech signal. The howling

phenomenon is also an oscillation or in other words, a resonance. Hence, DESA-2 is selected here to estimate the howling frequency. This algorithm is first tested on the whole band of the speech signal. Since the whole band contains different formants, it cannot be represented by an AM-FM model properly, and the DESA-2 does not show stable frequency unless the howling happens. This algorithm is then applied on the individual bands of the speech signal after passing through a Cosine Modified filter bank [9]. Compared to the whole-band signal, the output of each band can be more likely modeled as an AM-FM signal. Therefore, applying DESA-2 on the output of each band gives more accurate estimate of the howling frequency. Furthermore, to increase the accuracy of the proposed method, before applying DESA-2, the howling occurrence is checked by using GTKO. The GTKO has been introduced to resolve two closely-spaced tones [6] but it is modified and applied for a new application in this paper. Several GTKOs with different resolution parameters are used to detect the howling occurrence in the frequency range of the speech signal. As it will be seen later, each particular resolution parameter improves the frequency sensitivity of GTKO for a certain frequency range. Therefore, to achieve better frequency sensitivity and consequently higher howling detection capability, the input signal is passed through a filter bank and a GTKO with a specific resolution parameter is applied to the output of each band. Once a howling is detected in a band, the output of that particular band is sent to DESA-2 for identifying the howling frequency.

Performance of the proposed algorithm is compared with two well-known howling detection approaches, i.e. a frame-based approach called PHPR and a sample-based approach called ANF. The PHPR approach checks one of the main properties of the howling, which is not exhibiting significant power at the harmonics and sub-harmonics of the howling frequency. Among single-feature howling detection methods reported in [3], the PHPR approach has shown the best performance. Therefore, it is selected as a reference frame-based approach to compare the proposed method with. The ANF method tracks the spectral peaks of the signal and has been introduced as a howling detection method by Pandey et al. [4].

The paper is organized as follows. Section 2 introduces the PHPR and ANF approaches. TKO, DESA-2, and GTKO are briefly described in section 3. The proposed method is explained in Section 4. Section 5 presents the experimental results and conclusion is provided in Section 6.

## 2. HOWLING DETECTION APPROACHES

### 2.1. Frame-based algorithms

These algorithms process the input signal frame by frame and typically check some frequency-domain characteristics of the

howling, e.g. possessing large frequency magnitudes and not exhibiting significant power at the harmonics and sub harmonics of the howling frequency component.

The  $l$ th frame ( $l > 0$ ) of the input signal with length  $M$  and overlap of  $P$  samples is represented as

$$y(l) = [y((l-1)M - (l-1)P + 1) \dots y(lM - (l-1)P)]^T. \quad (1)$$

STFT corresponding to each frame is obtained by [2]

$$Y(\omega_k, l) = \sum_{n=(l-1)M-(l-1)P+1}^{lM-(l-1)P} w(n)y(n)e^{-j\omega_k n} \quad (2)$$

where  $\omega_k \triangleq 2\pi k/M$  and  $w(n)$  is the corresponding sample of a window usually used, i.e. Rectangular, Hamming (used in this paper), Hanning, or other types of windows. Once STFT is computed for a frame,  $N$  peaks of the spectrum are selected as the ‘‘howling component candidate’’ and collected in the set  $D_{\tilde{\omega}}(l) = \{\tilde{\omega}_i\}_{i=1}^N$ , where  $N$  is normally in the range of 1-10 [2]. In the PHPR approach,  $\tilde{\omega}_i$  is recognized as a howling component if the powers of its harmonics are negligible. In other words, the power of  $\tilde{\omega}_i$  is compared with the power of its  $m$ th harmonic by [2], [3]:

$$PHPR(\tilde{\omega}_i, l, m) = 10 \log_{10}(|Y(\tilde{\omega}_i, l)|^2 / |Y(m\tilde{\omega}_i, l)|^2). \quad (3)$$

Consequently,  $\tilde{\omega}_i$  is recognized as a howling component if the following condition is satisfied.

$$\cap_{m \in \{0.5, 1.5, 2.3, 4\}} [PHPR(\tilde{\omega}_i, l, m) \geq T_{PHPR}] = 1 \quad (4)$$

$\cap$  is the intersection operator, and  $T_{PHPR}$  is a threshold ( $T_{PHPR} = 33$  dB is suggested in [2]).

## 2.2. Sample-based algorithms

A few howling detection methods process the input signal in time domain and sample by sample, where the ANF is one of them [4]. Spectral peaks of the speech signal can be tracked and identified by the ANF approach. In this method, the input signal is filtered by a second-order ANF with the following transfer function.

$$H_a(z) = \frac{1 - a(n)z^{-1} + z^{-2}}{1 - \rho a(n)z^{-1} + \rho^2 z^{-2}} \quad (6)$$

where  $\rho$  and  $a(n)$  are constant and variable parameters, respectively. The ANF adjusts the center frequency of the notch filter by adjusting parameter  $a(n)$  in a way that at time  $n$  the power of the output of the filter is reduced [4]. The variability of  $a(n)$  is small if the system is tracking a strong frequency component. Hence, Pandey et al. [4] use this property to detect the howling frequency component. Update equations and rest of parameters are described in [4]. In addition to this property and using ANF, the authors have checked the energy growth in the input signal and improved the accuracy of ANF method in later versions [10]. For fair assessment here, simple DESA-2 (without GTKO) is compared with original ANF and PHPR approaches.

## 3. TKO AND RELATED OPERATORS

### 3.1. TKO

TKO has been defined for both continuous and discrete-time signals, where the discrete-time representation is used in this paper. For the input signal,  $y(n)$ , TKO is defined as [5]:

$$\psi(n) = y^2(n) - y(n+1)y(n-1). \quad (6)$$

For a pure sinusoidal signal,  $y(n) = A \cos(\omega_c n + \theta)$ , the operator output is  $A^2 \sin^2(\omega_c)$  which is proportional to the energy of a simple oscillation [8]. With a good approximation, this result is also valid for AM-FM signals [5].

### 3.2. DESA-2

Although TKO provides the amplitude and frequency of the sinusoidal or AM-FM signal, it does not determine their values separately. For a pure sinusoidal signal, the frequency can be exactly estimated by DESA-2 [5]:

$$\omega_c = \arcsin \left( \sqrt{\frac{\Psi[x(n+1) - x(n-1)]}{4\Psi[x(n)]}} \right) \quad (7)$$

The algorithm can also provide a good estimate while the input signal is an AM-FM signal [5]:

$$\begin{aligned} y(n) &= A(n) \cos(\Phi(n)) \\ &= A(n) \cos \left( \omega_c n + \omega_m \int_0^n q(m) d_m + \theta \right) \end{aligned} \quad (8)$$

where  $A(n)$  is the time-varying amplitude,  $q(m)$  is the frequency modulating signal, and  $\omega_m$  is the maximum frequency deviation from  $\omega_c$ . Maragos et al. [5] have shown that a speech signal band-limited around one of its formants can be modeled by an AM-FM signal. Therefore,  $f_c = \frac{\omega_c f_s}{2\pi}$  in (8) is the formant frequency and  $f_s$  is the sampling frequency. For AM-FM signals, the instantaneous frequency defined as  $f(n) = f_c + q(m)$  can be estimated by DESA-2 using (7) [5]. It should be mentioned that the total speech signal,  $s(n)$ , is modeled as a sum of such AM-FM signals:

$$s(n) = \sum_{j=1}^K A_j(n) \cos(\Phi_j(n)). \quad (9)$$

$j$  refers to the  $j$ th resonance and  $K$  is the number of formants [5].

### 3.3. GTKO

The GTKO has been originally introduced by Lin et al. [6] to resolve two closely-spaced tones and is defined as:

$$\psi_k(n) = y^2(n) - y(n+k)y(n-k) \quad (10)$$

where  $k$  is the resolution parameter which is an integer value not necessarily equal to one. Equation (10) is known as a general form of TKO because of having this unrestricted parameter.

For the sinusoidal signal, the value obtained by GTKO can be seen in (11) which is roughly valid for an AM-FM signal [11].

$$\psi_k(n) = (A \sin(\omega_c k))^2. \quad (11)$$

In this equation,  $k = \pi/2\omega_c$  maximizes the value of  $\Psi_k(n)$ . On the other hand, each frequency has its own optimum resolution parameter which results in the largest value in (11). The optimum resolution parameter may take a non-integer value which is not practical for discrete time signals. Therefore, the nearest integer resolution parameter to each optimum value is used in this paper.

$$k = [\pi/2\omega_c] \quad (12)$$

where the function  $[x]$  rounds up the value  $x$  to the smallest integer that is not smaller than  $x$ . By rounding up, the resolution

parameter takes integer values greater than 0. According to (12), a resolution parameter of 1 is a proper choice for the frequencies higher than 4 kHz. As the frequency decreases, the corresponding optimum resolution parameter increases. The frequency dependency of resolution parameter is used in applying GTKO for detecting energy growth and howling in this paper.

#### 4. PROPOSED HOWLING DETECTION METHOD

This paper uses DESA-2 to recognize the howling frequency. This algorithm has been used in howling detection. Here, its capability is checked when it is applied on the whole band and sub bands of the speech signal (Sections 4.1 and 4.2). Thereafter, to reduce the chance of detecting formants as howling frequencies the GTKO blocks are attached to DESA-2 (Section 4.3).

##### 4.1. Whole-band situation

The whole band of the speech signal contains various formants and cannot be represented with an AM-FM model. It means that, DESA-2 not often converges to a frequency. Therefore, it rarely detects a formant and thus, false detection probability is low. Compared to the formants, DESA-2 has less difficulty in recognizing the howling frequencies as they are typically stronger than the formants.

##### 4.2. Sub-band situation

By dividing the speech signal into the bands with 300 Hz ~ 500 Hz bandwidth, two formants barely sit in one band. Therefore, the outputs of the filter bank can be more likely represented as AM-FM models. This is typically true for low and middle-frequency bands and not for high-frequency bands (e.g. frequencies higher than 4 kHz). Since the formants rarely happen in high frequencies, the outputs of these bands are not AM-FM model. However, once a howling occurs in any band, no matter in low, middle, or high-frequency, it plays the role of carrier frequency and the output is an AM-FM signal. By applying DESA-2 on each band the chance of detecting a formant goes up. However, the probability of more accurately detecting the howling frequency increases, too.

In order to be able to apply DESA-2 on each band, the filter bank should provide real data in each band. As a result, the Cosine Modulated filter bank [9] is selected in this paper.

##### 4.3. GTKO and DESA-2

In order to reduce the false detection probability, the GTKO blocks are placed as interfaces between the filter bank and DESA-2. The GTKO is applied on the output of each band to detect instantaneous energy of that band which increases while howling happens. The resolution parameter of each GTKO block is selected using (12), while the central frequency of each band is chosen as  $\omega_c$  in this equation. Therefore, each GTKO block is tuned to maximally detect the energy of its corresponding band. This means that, several GTKO blocks, indeed with different resolution parameters, are applied simultaneously to monitor the whole band of the signal.

Once the howling is detected in a band, the output of that band is passed to DESA-2 for frequency recognition. Therefore, DESA-2 blocks are substituted by GTKO blocks which are computationally simpler and only one DESA-2 block is used in the final stage. To reduce the computational complexity even more, instead of monitoring the convergence of (7), one can only track

$\frac{\psi[x(n+1)-x(n-1)]}{\psi[x(n)]}$ . Once a convergence is detected, the frequency can be computed by (7).

## 5. EXPERIMENTAL RESULTS AND DISCUSSION

The howling detection methods presented in Section 4 are implemented here. The input signals are clean signals of Noizeus database [12], i.e. 30 IEEE standard speech files with  $f_s=16$  kHz.

The hearing aid device starts working in the steady situation without any howling occurrence. The gain of it then increases up to a level that generates howling. The filter bank has 16 sub bands and its prototype filter is FIR filter of order 32 with normalized cutoff frequency of 1/64 designed based on the MATLAB built-in function (fir1). The howling detection methods are evaluated by receiver operating characteristic (ROC) curves [3]. An appropriate howling detection method should have high detection probability ( $P_D$ ) or low missed detection probability. It should also have low false detection probability ( $P_{FA}$ ). By plotting  $P_D$  versus  $P_{FA}$  in a ROC curve, the overall performance of the algorithm is evaluated. Also, the appropriate threshold or parameter used in the howling detection algorithm can be selected based on the ROC curve.

Fig. 1 compares the ROC curves of DESA-2 (applied on whole band of the signal), PHPR, and ANF. One advantage of DESA-2 over the other two methods is possessing less number of parameters. Once the detected frequencies converge for  $N_s$  samples, the algorithm declares a howling detection. Four different values are considered for  $N_s$ , i.e. 1, 5, 10, and 100. By convergence it means that two successive detected frequencies are within  $\nu$ -Hz vicinity of each other. Each ROC curve of DESA-2 is plotted by assuming the value of 1 Hz for  $\nu$  and increasing it upward. The PHPR is a frame-based algorithm. Therefore, instead of  $N_s$  the number of successive frames are considered,  $N_f=1, 5, \text{ and } 10$ . Moreover, it detects the howling frequency based on the computed STFT. For a 20-ms frame (considered in the simulation) the number of Fast Fourier Transform (FFT) points are typically about 512 (The next power of two from length of the frame). Therefore, each point represents about 16 Hz (half of the sampling frequency/512) of the 8 kHz bandwidth of the signal. In other words, this method cannot detect the frequency very accurately. The curves in the second panel of Fig. 1 are corresponding to the PHPR approach with  $\nu=20$  Hz. This algorithm could not detect the howling with higher accuracy. For each curve in this panel, the threshold  $T_{PHPR}$  increases from 0 to infinity. Therefore, this method has one more parameter compared to DESA-2. The performance of the algorithm is highly dependent on this threshold value. One threshold may be appropriate for one situation but be inappropriately small or large for another situation. The third panel corresponds to the ANF method. Similar to DESA-2 this algorithm can detect howling for  $\nu \geq 1$  Hz, although  $P_D$  is too low for small  $\nu$ . The curves plotted in this panel correspond to  $\nu=20$  Hz. This method has several parameters [4] where two of them ( $\delta_q$  and  $T_a$ ) whose changes have more effect on  $P_D$  and  $P_{FA}$  are chosen as variables. The other parameters are chosen in such a way to maximize the performance ( $\rho = 0.95$ ,  $\alpha_a = 0.1$ ,  $\lambda_x = 0.9$ , and  $\lambda_m = 0.999$ ). Each curve is plotted by increasing  $T_a$  from 1 to infinity. According to this figure, DESA-2 outperforms the PHPR and ANF in terms of having lower  $P_{FA}$ .

The panels in Fig. 2 plot the required time (in ms) to detect the howling in each method. Depending on the value of  $N_s$  in DESA-2 with  $\nu=20$  Hz, detection time varies from 50 ms to 350 ms. For the PHPR and ANF it is within the same range.

Fig. 3 shows ROC curves and detection time for DESA-2 applied on the output of each sub-band.  $P_D$  increases in this case, e.g. for  $\nu=20$  Hz and  $N_s = 1$  it has increased from 0.45 to 0.8. Also, convergence happens faster and the detection time is much shorter.  $P_{FA}$  degrades when DESA-2 is applied on the sub-bands which will be addressed and solved in the GTKO+DESA-2 method.

Fig. 4 corresponds to the GTKO+DESA-2 algorithm. Using GTKO reduces both  $P_{FA}$  and  $P_D$ . However, the amount of  $P_{FA}$  improvement is more significant than the amount of  $P_D$  degradation, e.g. for  $\nu=20$  Hz and  $N_s = 1$ ,  $P_{FA}$  is ten times better than the value in Fig. 3 but  $P_D$  has minor degradation (Note the change in the horizontal-axis scale.)

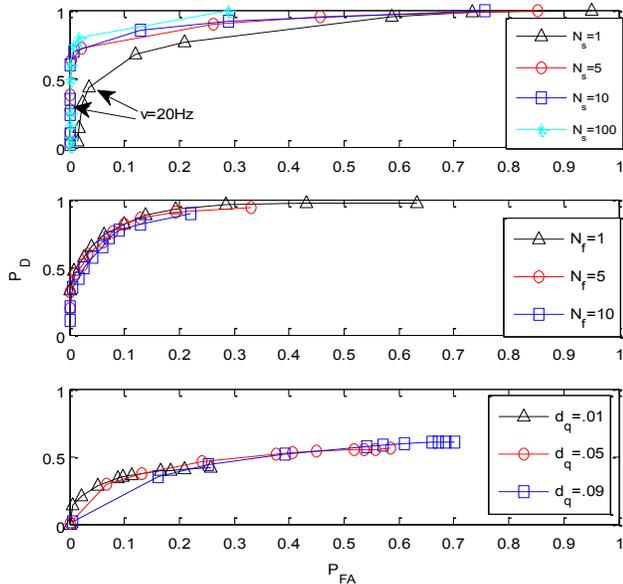


Fig. 1. ROC curves for DESA-2 (the first panel), PHPR (the second panel), and ANF (the third panel).

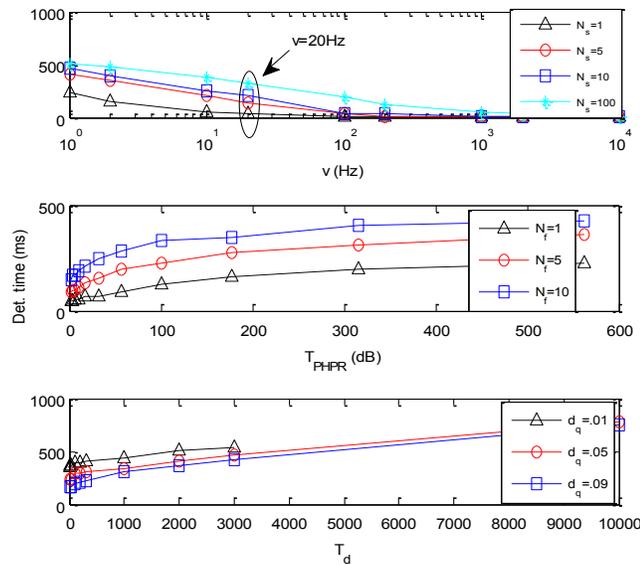


Fig. 2. Detection time for DESA-2 (the first panel), PHPR (the second panel), and ANF (the third panel).

## 6. CONCLUSION

A howling detection method was proposed using the GTKO as the howling detector and DESA-2 as the frequency recognizer. The paper first evaluated the capability of DESA-2 in recognizing the howling frequency. This algorithm had less number of parameters which made it simpler. Scanning the parameters and finding proper values for each environment or each input signal was not a concern in this method. The algorithm also showed low false detection probability. By applying DESA-2 on the sub-band outputs, detection probability and detection time were improved. Attaching the GTKO blocks also helped decreasing false detection probability.

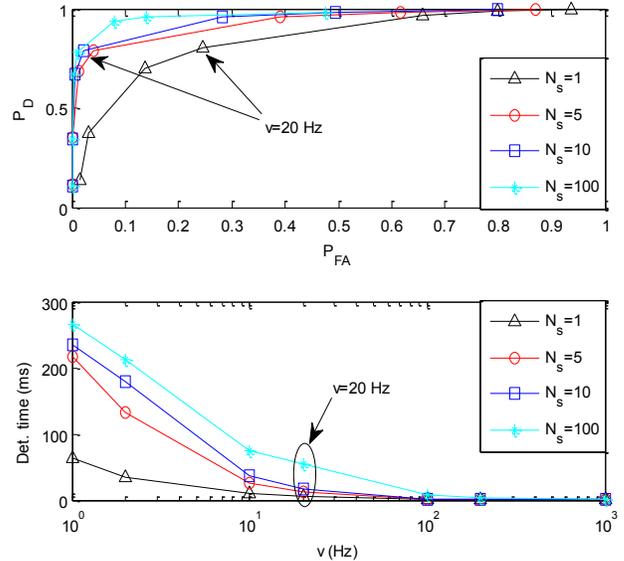


Fig. 3. ROC curves (the first panel) and detection time (the second panel) for DESA-2 applied on the sub bands.

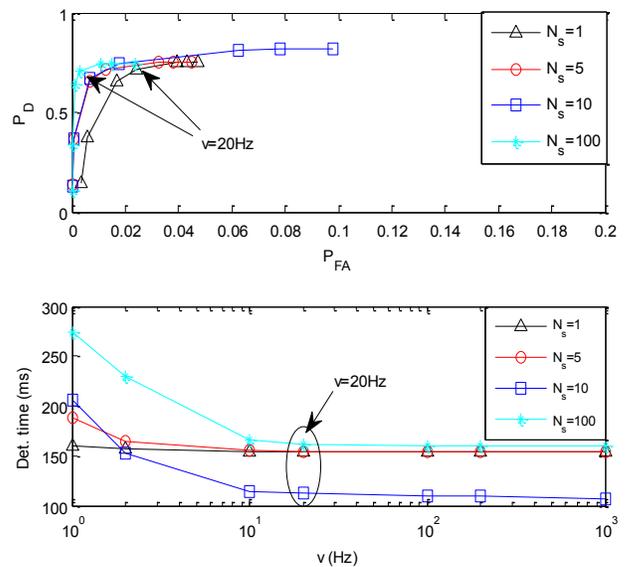


Fig. 4. ROC curves (the first panel) and detection time (the second panel) for DESA-2 applied with GTKO.

## 7. REFERENCES

- [1] S.A. Khoubrouy, I.M.S Panahi, and A.A. Milani, "A new delayless sub-band filtering method for canceling the effect of feedback path in hearing aid system," *IEEE EMBS*, Boston, pp. 7920-7923, 2011.
- [2] T.V. Waterschoot and M. Moonen, "Fifty years of acoustic feedback control: state of the art and future challenges," *Proc. IEEE*, pp. 288-327, 2011.
- [3] T.V. Waterschoot and M. Moonen, "Comparative evaluation of howling detection criteria in notch-filter-based howling suppression," *J. Audio Eng. Soc.*, vol. 58, pp. 923-940, 2010.
- [4] A. Pandey and V.J. Mathews, "Improving adaptive feedback cancellation in digital hearing aids through offending frequency suppression," *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Dallas, pp. 173-176, 2010.
- [5] P. Maragos, J.F. Kaiser, T.F. Quatieri, "Energy separation in signal modulations with application to speech analysis," *IEEE Trans. Signal Process.*, vol. 41, pp. 3024-3051, 1993.
- [6] W. Lin, C. Hamilton, P. Chitrapu, "A generalization on the Teager-Kaiser energy function and application to resolving two closely-spaced tones," *IEEE Int. Conf. Acoust., Speech, Sig. Process.*, vol. 3, pp. 1637-1640, 1995.
- [7] H.M. Teager, "Some observations on oral flow during phonation," *IEEE Trans. Acoust., Speech Sig. Process.*, vol. 28, no. 5, pp. 599-601, Oct. 1980.
- [8] J.F. Kaiser, "On a simple algorithm to calculate the 'energy' of a signal," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Albuquerque, NM, pp. 381-384, 1990.
- [9] S.K. Mitra, "Multirate digital signal processing" in *Digital Signal Processing*, 2nd ed. New York: McGraw-Hill, 2001, pp. 696-734.
- [10] A. Pandey and V.J. Mathews, "Adaptive gain processing with offending frequency suppression for digital hearing aids," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, pp. 1043-1055, 2012.
- [11] J.H. Choi and T. Kim, "Neural action potential detector using multi-resolution TEO," *IET J. Mag.*, vol. 38, pp. 541-543, 2002.
- [12] Y. Hu and P. Loizou, "Subjective evaluation and comparison of speech enhancement algorithms," *Speech Commun.*, vol. 49, pp. 588-601, 2007.