International Journal of Electronic Engineering Research.
ISSN 0975-6450 Volume 4, Number 3 (2012) pp. 205-221

http://www.ripublication.com/ijeer.htm

Speech Signal Processing: Non-Linear Energy Operator Centric Review

***Subhradeep Pal**

**Department of AEIE, Sir J.C. Bose School of Engineering, 1, Khan Road, Mankundu, Hooghly-712139, West Bengal, India*
**Corresponding Author E-mail: subhradeeppal@gmail.com*

Abstract

In modern days the need of speech signal processing using energy operators has drawn a lot of attention for researchers. The use of Teager Energy Operator and other differential energy operators in signal processing enables us a lot of new techniques to estimate and process the speech signal in noisy environment. In this paper we shall discuss the speech signal processing using the nonlinear energy operators. The simulation results are incorporated in support of theoretical analysis.

Index Terms: TEO, Dyn operator, AM-FM signal, Instantaneous Frequency, Formants, ESA

Introduction

The Information we want to express is contained in our spoken words and the spoken words are conveyed by the speech signal. Before we can proceed to any further details of the speech signal we have to understand the characteristics and describe a suitable model for the speech signal productions. The term speech signal processing has a wide sense of meaning. Now the speech signal processing refers to the manipulation, acquisition, storage and transfer of the human utterances or voice signals by a computer. The signal processing is mainly done by keeping three major goals in mind. Firstly the original human voice signals are captured and process it to convert into the computer readable or recognizable format. This step is known as the *speech recognition*. Although the *speech processing* is the reverse process of the speech recognition but its applicability is totally dependent on the computer's usability and versatility. The *speech compression* is the technique where the original signal is compressed to hear, to store or to transmit the maximum amount of data transmission possible for a given set of time and space constraints. In this paper we start from the human vocal system and discuss the amplitude and frequency modulation of speech signal and also the nonlinear energy operators and their applications in speech signal processing. In this paper we shall compare the TEO and Dyn operators in the context of AM-FM signals. Simulation results are incorporated in the result part of the each section in support of the theory.

Speech Production Process

The Human Vocal System

The vocal tract system is main important thing in the considerations of the speech signal production. It is the shape of the vocal tract which determines the physical distinguishing factor of the speech. Generally the vocal tract system is considered as the speech production organ. As depicted in the Fig 1, vocal tract includes laryngeal pharynx, oral pharynx, oral cavity, nasal pharynx and nasal cavity. Typically in normal adult male it is approximately 17 cm long [1].

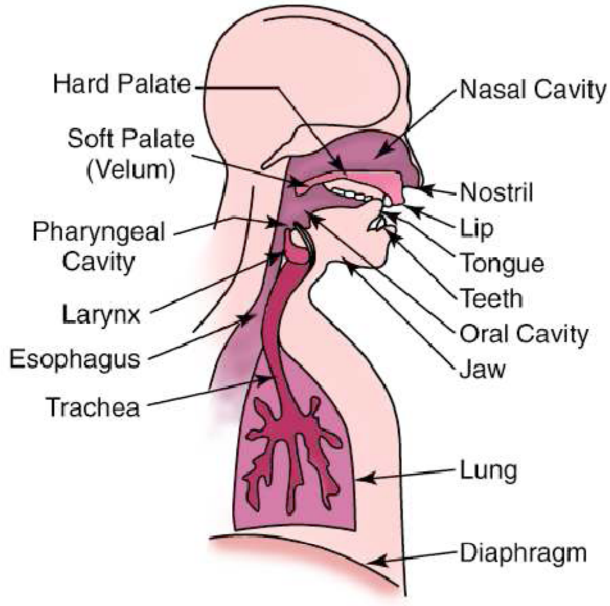


Figure 1: The human vocal system.

As the acoustic wave passes through the vocal tract its frequency content or spectrum is altered by the resonances in the vocal tract known as *formants*. Hence we can easily estimate the structure of the vocal tract from the spectral shape of the vocal signal. Voice verification systems use the human feature derived from the vocal tract. The vocal mechanism is completely driven by the excitation source containing the information of the speaker. In broader sense this excitation can be classified in to six categories. They are phonation, whispering, vibration, compression, frication and combination of all of these.

In the phonation the excitation is occurred when the air flow is modulated by the vocal cords. Whispered excitation is produced by airflow rushing through a small triangular opening between the arytenoids cartilages at the rear of the nearly closed vocal folds. The resultant signal is a wideband noise. Frication excitation is produced by constrictions in the vocal tract. Sounds generated by frication are called *fricatives* or *sibilants*. Compression excitation results from releasing a completely closed and pressurized vocal tract. Vibration excitation is caused by air being forced through a closure other than the vocal folds, especially at the tongue. Speech produced by phonated excitation is called *voiced,* speech produced by phonated excitation plus frication is called *mixed voiced,* and speech produced by other types of excitation is called *unvoiced.*

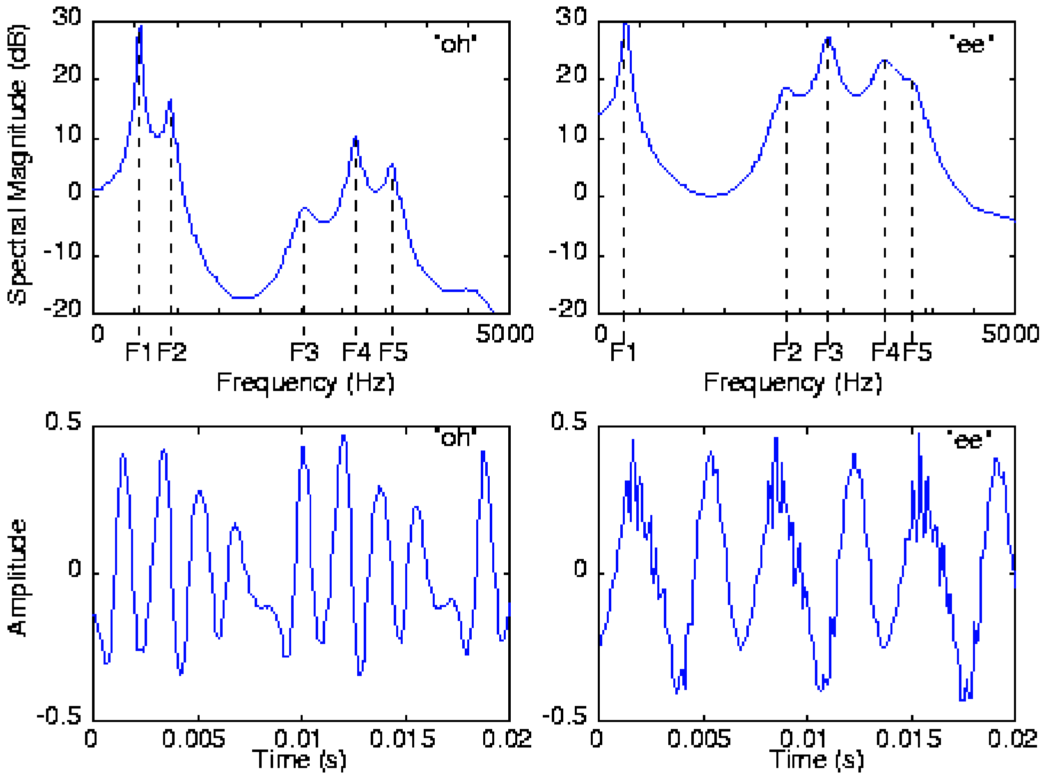


Figure 2: The speech signal spectrum for the sounds 'oh' and 'ee' along with their ideal response of the vocal tract. Picture adapted from Wikipedia.

Linear Prediction Modeling of Speech Signals

The all pole LP models a speech signal s[n] by the linear combination of its past values and present inputs. It can be represented by [1]

$$s[n] = G.u[n] - \sum_{k=1}^{p} a_k s[n-k] \tag{5}$$

Where, s[n] is the present output, p is the predictor co-efficient, a_k is the predictor co-efficients, s [n-k] is the past output, G is the gain scaling factor and u [n] is the present input. The simple LP model based on the past output samples only can be expressed as

$$\hat{s}[n] = -\sum_{k=1}^{p} a_k s[n-k] \tag{6}$$

The prediction error can be defined as the difference between the actual values and predicted value as

$$e[n] = s[n] - \hat{s}[n] = s[n] + \sum_{k=1}^{p} a_k s[n-k] \tag{7}$$

And from the expression (6) it is clear that the prediction error is similar to the scaled input signal. The mean squared error (MSE) can be given by

$$E = \sum_n (e[n])^2 = \sum_n \left[s[n] + \sum_{k=1}^{p} a_k s[n-k] \right]^2 \tag{8}$$

More detailed analysis is beyond of this paper and this portion is omitted. In the linear prediction model of speech signals there are two process of analysis are present. Both of the process has some advantages and disadvantages when compared to each other.

Auto correlation method

In this method it is assumed that the signal is identically zero outside the analysis interval. Then it tries to minimize the prediction error whenever it is non zero. Due to the use of the windowed speech signal the error is most likely to be large at the start and end of the speech interval. The use of window like Hamming window utilizes the tapering of the speech segment to be analyzed to minimize the error. The most advantage of this method is that stability of the resulting model is ensured.

Covariance method

In contrast to the auto correlation method the covariance method of the LP, here the interval over which the mean-squared error (MSE) is minimized and speech is not taken to be zero outside this interval. Stability of this model cannot be guaranteed but usually for large analysis interval the predictor co-efficients are stable. Here the error auto correlation and spectrum are calculated in order to measure the whiteness of the speech signal.

AM-FM Speech Signal Modeling

We know that, $x(t) = a(t)\cos(\omega_c t + \omega_m \int_0^t m(\tau)d\tau + \theta)$ is a real valued signal which has both an AM and FM structure and thus we call this type of signal as AM-FM signal. It is interesting to note that two different information signals can be simultaneously transmitted in the amplitude a (t) and the frequency ω_i (t), where we can write the instantaneous frequency [4, 5] of the signal as,

$$\omega_i(t) = \frac{d}{dt}\phi(t) = \omega_c + \omega_m m(t) \tag{1}$$

This type of AM-FM signals are widely and extensively used in the communication systems. By the term "speech resonances", we refer to the oscillatory systems formed by the local cavities of the vocal tract which emphasizes and de-emphasizes certain frequencies during the speech production as described earlier. In the LP modeling all the speech resonances are characterized by the poles of the transfer function of the linear filters that models the vocal tract [1, 6]. Now each pair of the complex conjugate poles refers to the second order resonator with exponentially damped cosine impulse response and so we have,

$$R(t) = Ae^{-\sigma t}\cos(\omega_c t + \theta) \tag{2}$$

The formant frequency is ω_c where the factor σ controls the formant bandwidth (BW). This approach is based on the assumption that there is local stationarity present in the speech signal.

Teager had shown earlier that speech resonances can change rapidly both in frequency and amplitude within a single pitch period due to rapidly time-varying and separated speech air flow in the vocal tract. Now we know that the air passes through the vocal tract cavities and the effective cross-sectional areas of the air flow can vary rapidly causing modulations of the air pressure or velocity filed [6]. Thus we can model the modulation of the each speech resonance as

$$R(t) = a(t)\cos[\phi(t)]$$

$$= e^{-at}A(t)\cos[\omega_c t + \omega_m \int_0^t m(\tau)d\tau + \theta] \tag{3}$$

And the total speech signal can be modeled as

$$s(t) = \sum_{k=1}^{K} a_k(t)\cos[\phi_k(t)] \tag{4}$$

Where, the k^{th} subscript refers to the k^{th} resonances and K is number of speech formants [2].

Nonlinear Energy Operators

The nonlinear energy operators can be capable of estimating the speech signal energy as these operators are basically energy tracking operators. The nonlinear differential energy operators like Teager-Kaiser Energy Operator (TEO) can detect formant AM-FM modulations by estimating the product of their time varying amplitude and frequency. The Teager Energy Operator is considered to be a very high resolution energy operator [7]. As described earlier the speech resonances can be modeled as AM-FM signals so demodulate the speech signals we have to study the AM-FM signals (Fig 1) in terms of nonlinear energy operators or in other words if we can demodulate the AM-FM signals we can easily extend our results for the so called speech signals. Maragos, Kaiser and Quatieri had developed the various energy separation algorithms (ESA) in order to separate the amplitude and frequency from the original signal. Brief discussions on those algorithms are discussed here. Instead of going through the detailed mathematical analysis of the algorithms we show the simulation results of those algorithms.

Definition of TEO

The nonlinear Teager operator can be defined as

$$\psi(x(t)) = \left[\dot{x}(t)\right]^2 - \left[x(t)\ddot{x}(t)\right] \tag{5}$$

And the discrete version of the operator can be defined as

$$\psi[x[n]] = x^2[n] - x[n-1]x[n+1] \tag{6}$$

The TEO can also be defined from the system equation of an undriven linear undamped oscillator [11, 12] as well as from the Lie Bracket operator [13] as

$$\left[x, \dot{x}\right] = \left(\dot{x}\right)^2 - x\ddot{x} = \psi(x) \tag{7}$$

Definition of Dyn Operator

The definition of Teager Energy Operator was first given by J. F. Kaiser in the year 1991 [12]. It was a nonlinear approach to the speech signals. Similarly J. Rouat defined another nonlinear energy operator called *Dyn Operator* in the same year. The Dyn operator is also an energy tracking operator and it can be defined as

$$Dyn[x(t)] = x(t)\,\dot{x}(t) \tag{8}$$

This operator is not so conventional as compared to the TEO and has some less importance in the field of speech signal processing. Both the energy operator defined here has some special advantages and disadvantages which will be discussed later. Other higher order differential energy operators like energy velocity and energy acceleration operators [13] are used also in the field of the speech signal processing which are omitted here.

Dyn Operator vs. Teager Energy Operator

The Dyn operator has the advantage of less sensitive to noise and quantization errors as compared to the TEO. But the most important disadvantage of it is that we have to post process the output of the Dyn operator in order to extract the frequency and amplitude information. The TEO is free from this problem and it's our perception that being free from post processing of the signals the TEO offers less circuit complexity as compared to the Dyn operator.

Energy Separation Algorithms

The energy separation algorithms are the various processes to separate the energy, amplitude and frequency components using energy operators from a signal. The ESA can be broadly classified into two major categories as:

1. Continuous Energy Separation Algorithm
2. Discrete Time Energy Separation Algorithms

The discrete time energy separation algorithms are again of three major types

1. DESA 1
2. DESA 2
3. DESA 1a

We shall discuss all of these algorithms in brief starting from the CESA and then followed by the DESAs.

Continuous Energy Separation Algorithm

For constant amplitude/ frequency signals

The continuous energy separation algorithm (CESA) for a fixed amplitude or frequency cosine signals are discussed here. We assume that the signal can be expressed as, $x(t) = A\cos(\omega_c t + \theta)$ where its amplitude A and frequency ω_c are constant. Then we have its derivative as

$$\dot{x}(t) = -A\omega_c \sin(\omega_c t + \theta)$$

We have the TEO output of the original signal and its time derivative signal as

$$\left.\begin{aligned} \psi\left(\dot{x}(t)\right) &= A^2\omega_c^4 \\ \psi\left(x(t)\right) &= A^2\omega_c^2 \end{aligned}\right\} \tag{9}$$

Then from the relation (9) the amplitude and frequency can be estimated as

$$\left.\begin{aligned} |A| &= \frac{\psi[x(t)]}{\sqrt{\psi[\dot{x}(t)]}} \\ \omega_c &= \sqrt{\frac{\psi[\dot{x}(t)]}{\psi[x(t)]}} \end{aligned}\right\} \tag{10}$$

For AM-FM signals

We consider the real valued AM-FM signal as defined in Section II for CESA analysis of AM-FM signals. We have the TEO output for the real valued AM-FM signals

$$\psi[a(t)\cos(\phi(t))] = (a\dot{\phi})^2 + \frac{a^2\ddot{\phi}\sin(2\phi)}{2} + \cos^2(\phi)\psi(a) \tag{11}$$

Here the desired term is $D(t) = (a\dot{\phi})^2$ while the error term is $E(t) = \frac{a^2\ddot{\phi}\sin(2\phi)}{2} + \cos^2(\phi)\psi(a)$ in the expression (11). Thus the expression for the AM-FM signals can be rewritten as

$$\psi[x(t)] = D(t) + E(t) \approx D(t) \tag{12}$$

Since we concentrate on the narrowband AM-FM signals so we have error term having maximum value much less than that of the desired value D (t) or in other words we have,

$$\psi[a\cos(\phi)] = (a\dot{\phi})^2$$

When the following condition is satisfied,

$$\psi(a)_{\max} + 0.5(a^2\ddot{\phi})_{\max} << (a\dot{\phi})^2_{\max} \tag{13}$$

To separate the amplitude/ frequency information from the signal we need to apply the TEO or ψ to the derivative of AM-FM signal i.e.

$$\dot{x}(t) = \dot{a}(t)\cos\phi - a(t)\dot{\phi}(t)\sin\phi = y_1 - y_2 \tag{14}$$

From (11) we get that,

$$\psi[a(t)\dot{\phi}(t)\sin\phi] = a^2\dot{\phi}^4 - 0.5a^2\dot{\phi}^2\ddot{\phi}\sin(2\phi)$$

$$+[a^2\psi[\dot{\phi}] + \dot{\phi}^2\psi[a]]\sin^2(\phi) \tag{15}$$

Now we see that the desired term for energy separation is $a^2\dot{\phi}^4$ which can be combined with D(t) to have simple equations for the frequency and amplitude signals. But the analysis does not consider the effect of the other cross energy terms which occurs during the analysis. This detailed analysis can be found in [6, 15].

Discrete Time Energy Separation Algorithm-1 (DESA-1)
This algorithm was proposed by Maragos, Kaiser and Quatieri in 1992 in order to estimate the amplitude envelope and the instantaneous frequency of discrete time AM-FM signals using discrete time energy operators. In this case we shall discuss only the constant amplitude/frequency signals for the discussion.

We have the discrete version of TEO defined as

$$\psi[x[n]] = \frac{x^2[n] - x[n-1]x[n+1]}{T^2} \tag{16}$$

A constant amplitude/frequency signal having the structure of

$x[n] = a\cos(\Omega_c n + \theta)$, gives the TEO output of

$$\psi[x[n]] = A^2 \omega_c^2 \left(\frac{\sin \Omega_c}{\Omega_c} \right)^2 \tag{17}$$

Now rewriting the derivative with help of the two-sample backward difference and we have [6]

$$\begin{aligned} \psi[y[n]] &= \psi[x[n] - x[n+1]] \\ &= 4A^2 \sin^2\left(\frac{\Omega_c}{2}\right) \sin^2(\Omega_c) \end{aligned} \tag{18}$$

Then we have the result of

$$\frac{\psi[y[n]]}{2\psi[x[n]]} = 1 - \cos \Omega_c \tag{19}$$

This result will lead us to the following results as

$$\left. \begin{aligned} \Omega_c &= \cos^{-1}\left(1 - \frac{\psi[x[n] - x[n-1]]}{2\psi[x[n]]} \right) \\ |A| &= \sqrt{\frac{\psi[x[n]]}{1 - \left(1 - \frac{\psi[x[n] - x[n-1]]}{2\psi[x[n]]} \right)^2}} \end{aligned} \right\} \tag{20}$$

The detailed analysis for the AM-FM signals can be found in [6, 15].

Discrete Time Energy Separation Algorithm-2 (DESA-2)
Instead of using two sample derivatives in DESA-1 if we use three sample derivative or three sample symmetric differences i.e. if we use $y[n] = x[n+1] - x[n-1]$, then we

have the DESA-2 algorithm. The new expression for the amplitude and frequencies are given as:

$$\left.\begin{aligned} \Omega_c &= 0.5\cos^{-1}\left(1-\frac{\psi[x[n+1]-x[n-1]]}{2\psi[x[n]]}\right) \\ |A| &= \frac{2\psi[x[n]]}{\sqrt{\psi[x[n+1]-x[n-1]]}} \end{aligned}\right\} \tag{21}$$

Discrete Time Energy Separation Algorithm -1a (DESA-1a)

The DESA-1a is the special form of DESA-1 where 1 refers to two sample difference and ‘a’ denotes that we have used asymmetric derivative in this case. The detailed analysis of DESA-1a is discussed in [6, 14, 15] and hence we omit the analysis portion of this algorithm.

Alternative Energy Separation Algorithms

We have previously discussed the CESA and three DESAs in brief. In addition to those ESAs there are two more alternative ESAs are present and we shall give a very brief introduction to them. With the help of these alternative ESAs we are also able to find out the amplitude and frequency component of the AM-FM signals.

Alternative CESA

For a constant cosine signal,

$$x(t) = A\cos(\omega_c t+\theta)$$

using the alternative CESA we can easily write for the amplitude and frequency as

$$\left.\begin{aligned} \omega_c^2 &= -\frac{\ddot{x}(t)}{x(t)} \\ A^2 &= x^2(t) - \frac{x(t)\,\dot{x}^2(t)}{\ddot{x}(t)} \end{aligned}\right\} \tag{22}$$

The amplitude estimator can be represented with the energy operator terms as

$$A^2 = \frac{\psi[x(t)]}{\omega_c^2} \tag{23}$$

This alternative CESA is not a low bandwidth function due to presence of the derivative terms of the original signal. But the previous CESA is a low bandwidth function and hence it is expected to give out stable and less noisy outputs when compared to this alternative CESA [6].

Alternative DESA

In the alternative DESA we combine various time shifted version of the signals and their output form the **ψ (.)** in order to obtain a system of equations whose solutions will lead us to the amplitude and frequency separation from the original signal.

For a discrete cosine signal we have $x[n] = a\cos(\Omega_c n + \theta)$ and thus we can write that,

$$x[n+1] + x[n-1] = 2A\cos\Omega_c \cos(\Omega_c n + \theta)$$
$$= 2x[n]\cos\Omega_c$$

Thus the frequency and amplitude can be separated as

$$\left.\begin{aligned} \Omega_c &= \cos^{-1}\left(\frac{x[n+1]+x[n-1]]}{2x[n]}\right) \\ |A| &= \sqrt{\frac{\psi[x[n]]}{1-\left(\frac{(x[n+1]+x[n-1])^2}{4x^2[n]}\right)}} \end{aligned}\right\} \tag{24}$$

This algorithm has the obvious advantage of having no signal difference and instead of that it averages the signal and hence the alternative DESA may more robust than the previous DESAs. However the algorithm does not generalize to the time varying cases [6].

Computational Complexities in ESAS

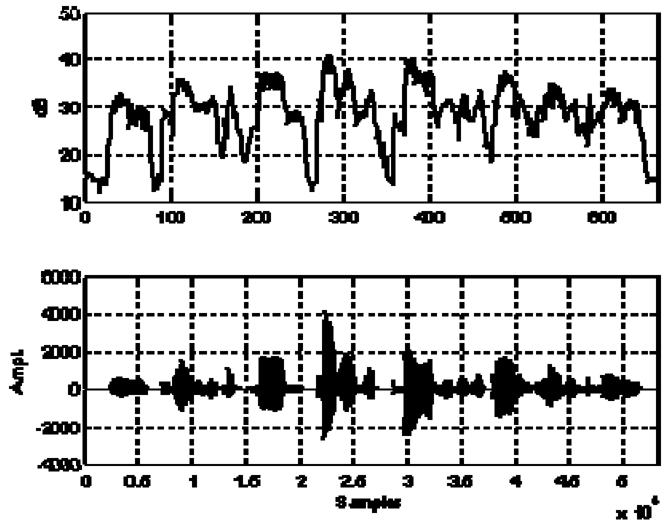
The computational complexity is one of the important parameters during the computational delays of the ESAs. The DESAs are very simple algorithms to compute and all the complexity are linear type. A table listing the complexities of the describe DESAs in terms of number of operations in per sample is listed in Table-1. Among the DESAs the DESA-2 algorithm is the fastest while the DESA-1 is the slowest but it worth to note that the difference is very small in practical cases.

Table 1: Computational Complexities of ESAs

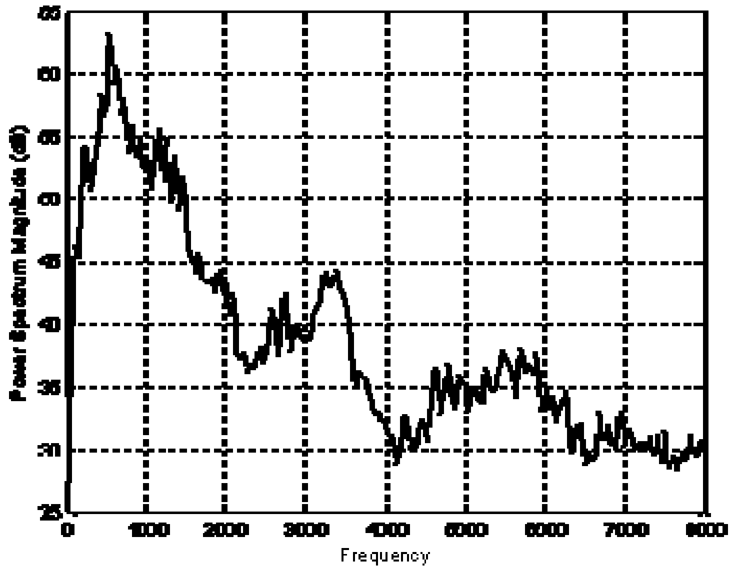
Operations	Energy Operator	DESA-1	DESA-1a	DESA-2
Addition	1	6	5	4
Multiplication	2	8	8	8
Square root	0	1	1	1
Cos^{-1}	0	1	1	1
w	3	5	4	5

Simulation Results

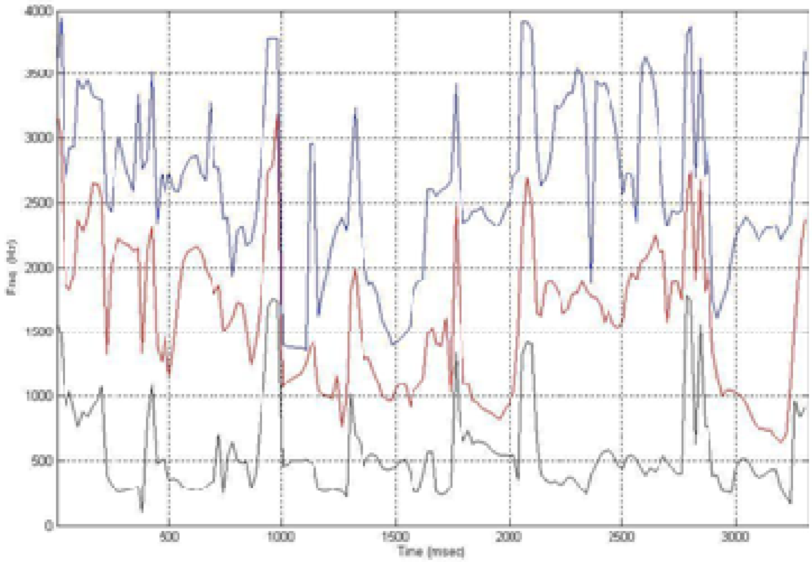
For the simulation purpose MATLAB and SIMULINK are extensively used. We start our simulation from the speech signal and its determination of formants along with FFT is shown here. The Fig. 3 shows a typical speech signal along with its energy plotting and power spectrum density and determined formants.



(a)



(b)



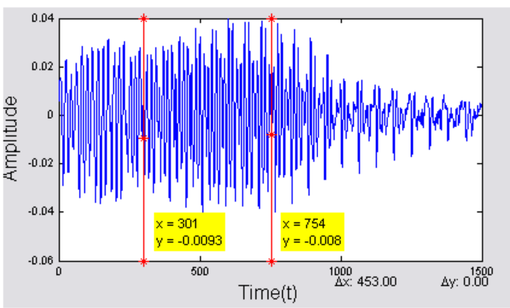
(c)

Figure 3: A typical speech signal showing its (a) energy plot (b) power spectral density (c) determined formants.

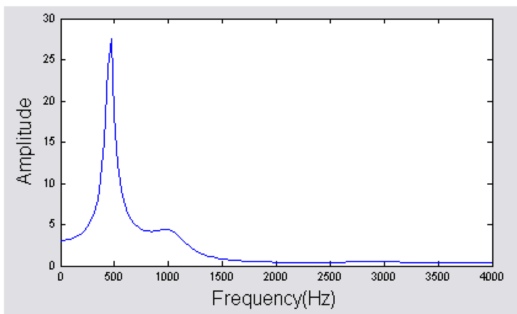
Another analysis of speech signal with other uttered sentence is shown in Fig 4. A close comparison of the two portions of the shown speech will reveal its different formant frequencies which are listed in Table-2.

Table 2: Various Formant Frequencies of Speech Signal in Fig 4

Formant Frequencies (Hz)	Time duration of speech signal	
	301ms- 754 ms	459ms-840ms
F_1	458.8917	504.881
F_2	1042.7825	1088.9779
F_3	2611.0224	2844.3479



(a)



(b)

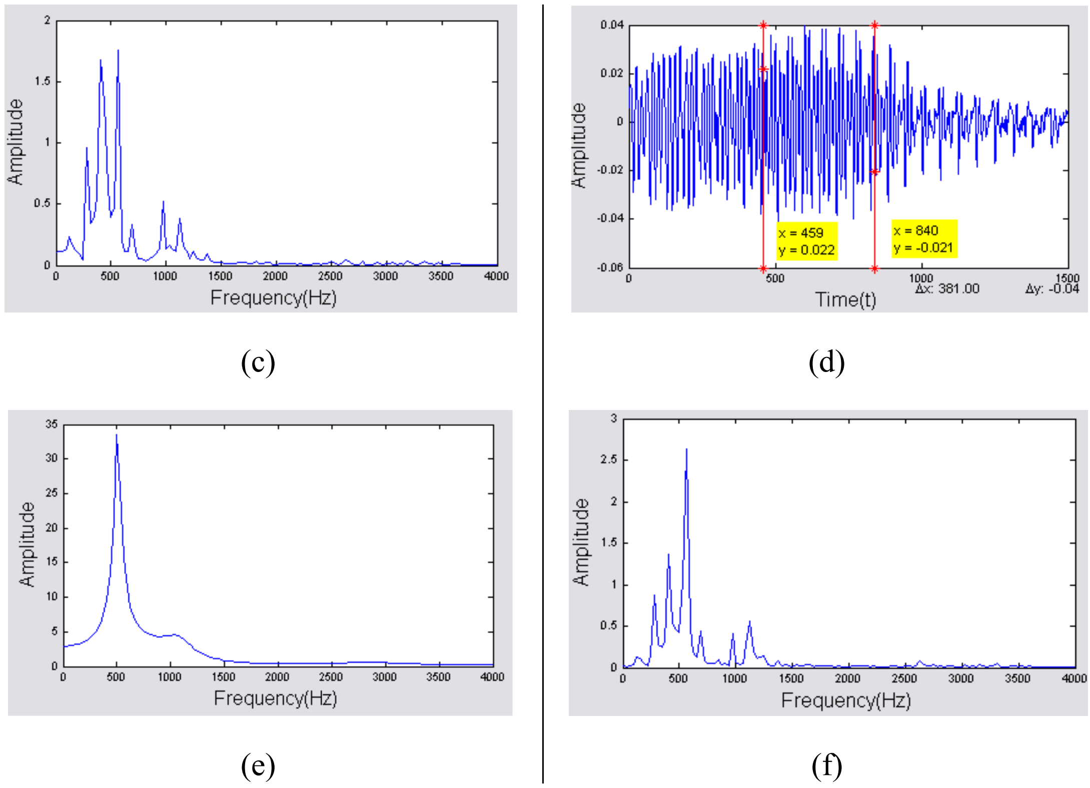


Figure 4: Comparative study of the various portions of a single speech signal. (a) and (d) are the original signals with red lines showing time duration selected for analysis; (b), (e) are analyzed formant frequency plotting of the selected portion of the speech signal; (c), (f) are the FFT plot of the selected portion of the speech signals.

So far we have extensively studied the speech signal. Now we again concentrate on the ESAs in order to show their effectiveness in the case of the speech signals. We start from the CESA. First we have simulated the frequency separating algorithm in the continuous domain and studied it for a wide range of the frequencies. A typical plot of the estimated frequency vs. the original signal frequency is shown in Fig 5. From the analysis we have seen that the algorithm seems to be some what erroneous due approximation of the term pi at high frequency ranges. However the algorithm gives exact frequency information at the low and medium frequency ranges. But a good approximation of the term pi will lead us to a good result at high frequencies too.

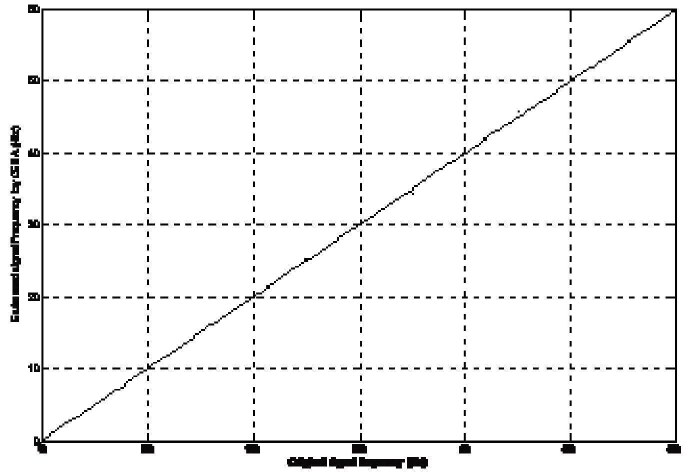


Figure 5: Plot showing the result for the frequency estimation by the CESA.

We now comeback to the simulation part of the DESA-1. The simulation for DESA-1 is done for simple AM signal. In the Fig 6 we have shown the results.

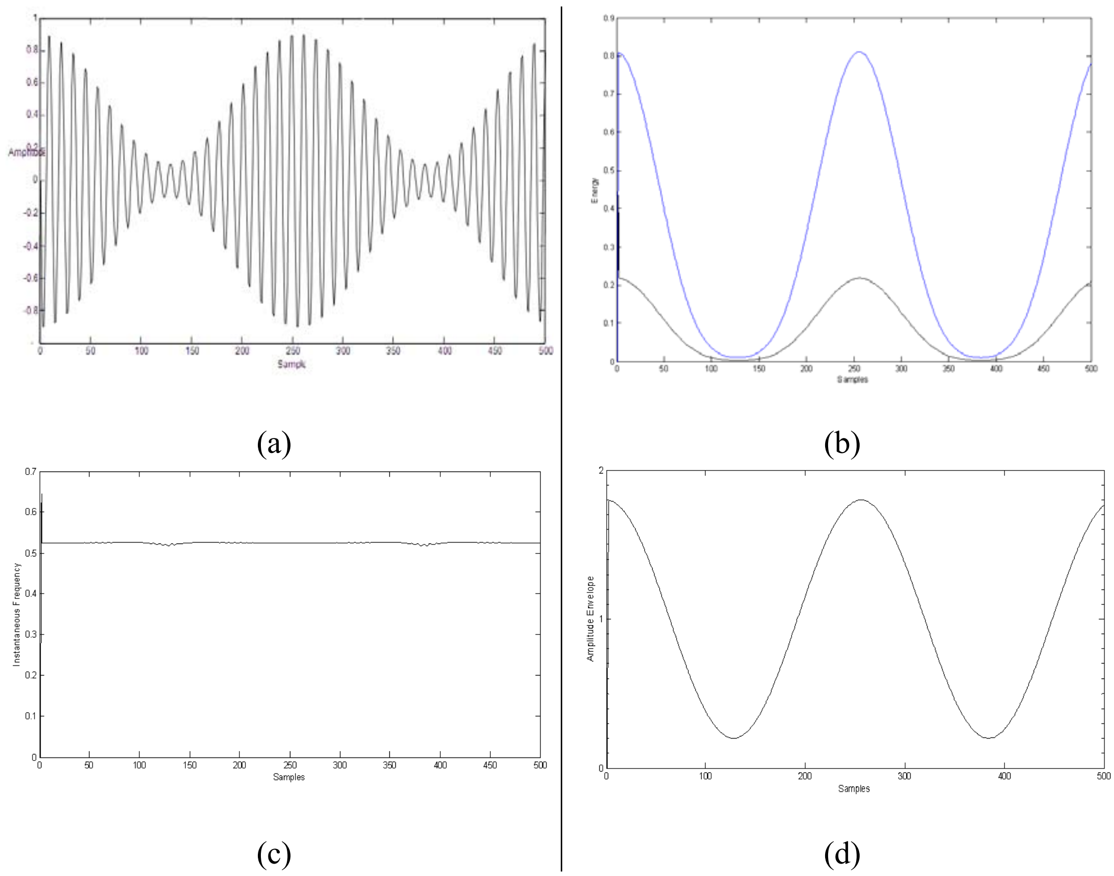


Figure 6: Simulation results for DESA-1. (a) Original signal used for analysis by DESA-1; (b) Estimated energy by the nonlinear energy operator; (c) Estimated instantaneous frequency; (d) Estimated amplitude envelope.

Conclusions

In this paper, modelling and analysis of speech signals is discussed in the context of the Teager energy operator. Simulation results shows a good approximation in the estimation of the amplitude, energy and instantaneous frequency of the speech signal can be done with the help of DESA 1 algorithm. Conventional TEO can be utilized in the extraction of the message signal from the modulated speech signal as discussed in [23].

References

[1] J.P. Campbell, Jr., "Speaker Recognition: A Tutorial", *Proceedings of the IEEE*, vol. 85, no. 5, pp. 1437-1462, September 1997.

[2] B.S. Atal and S.L. Hanauer, "Speech Analysis and synthesis by linear prediction of the speech wave", *Journal of Acoustic Society of America*, vol. 50, pp. 689-740, 1971.

[3] J. Picone, "Signal Modeling Techniques In Speech Recognition", *Proceedings of the IEEE*, June 1993.

[4] B. Boashah, "Estimating and Interpreting the Instantaneous Frequency of a Signal — Part 1: Fundamentals", *Proceedings of the IEEE*, vol. 80, no. 4, pp. 520-538, April 1992.

[5] B. Boashah, "Estimating and Interpreting the Instantaneous Frequency of a Signal — Part 2: Algorithms and Applications", *Proceedings of the IEEE*, vol. 80, no. 4, pp. 540-568, April 1992.

[6] P. Maragos, J. F. Kaiser and T. F. Quatieri, "Energy Separation in Signal Modulations with Applications to Speech Analysis", *IEEE Transactions on Signal Processing*, vol. 41, no. 10, pp. 3024-3051, October 1993.

[7] L.G. Ceballos, J.H.L. Hansen and J.F. Kaiser, "Vocal Fold Pathology Assignment using AM Autocorrelation Analysis of the Teager Energy Operator",

[8] A.S. Spanias, "Speech Coding: A tutorial review", *Proceedings of the IEEE*, vol. 82, no. 10, pp. 1541-1582, October 1994.

[9] H. Hermansky, "Perceptual linear predictive (PLP) analysis of speech", *Journal of Acoustic Society of America*, vol. 87, no. 4, pp. 1738-1752, April 1990.

[10] A. El-Jaroudi and J. Makhoul, "Discrete All-Pole Modeling" , *IEEE Transactions on Signal Processing*, vol. 39, no. 2, pp. 411-423, February 1991.

[11] J.F. Kaiser, "On Teager's Energy Algorithm and Its Generalization to Continuous Signals", *Proceedings of the 4th IEEE Digital Signal Processing Workshop*, Mohonk, NY, September 1990.

[12] J.F. Kaiser, "On a simple algorithm to calculate the 'energy' of a signal", *Proceedings of the IEEE ICASSP-90*, Albuquerque, NM, pp-381-384, April 1990.

[13] P. Maragos and A. Potamianos, "Higher Order Differential Energy Operators", *IEEE Signal Processing Letters*, vol. 2, No. 8, August 1995.

[14] P. Maragos, T.F. Quatieri and J.F. Kaiser, "On Separating Amplitude from Frequency Modulations Using an Energy Operator", *Proceedings of the IEEE* ICASSP-92, vol. 2, pp-1-4.
[15] P. Maragos, J.F. Kaiser, T.F. Quatieri, "On Amplitude and Frequency Demodulation Using Energy Operators", *IEEE Transactions on Signal Processing*, vol. 41, No. 4, pp. 1532-1550, April 1993.
[16] J. Rouat, "Nonlinear Operators for Speech Analysis", _________.
[17] B.N. Biswas, "On the Fundamental Aspects of Demodulation", unpublished work.
[18] A.C. Bovik, P. Maragos, "Conditions for Positivity of an Energy Operator", *IEEE Trans. Signal Proc.*, vol. 42, No. 2, February 1994.
[19] Valentin Emiya, Lúcio F.C. Pessoa, Delphine Vallot, and David Melles, "Generic Tone Detection Using Teager-Kaiser Energy Operators on the StarCore SC140 Core", *Freescale Semiconductor Application Note*, AN2384, Revised: December2004.
[20] Fabien Salzenstein and et al.: "Teager-Kaiser Energy and Higher-Order Operators in White-Light Interference Microscopy for Surface Shape Measurement", *EURASIP Journal on Applied Signal Processing*, Page-2804-2815, 2005.
[21] Zhao Junhui, Kuang Jingming, Kuang Jingming, Huang Shilei: "Noise suppression based on Teager energy operator for improving the robustness of ASR FRONT-END", *International Workshop on Acoustic Echo and Noise Control*, September 2003, Kyoto, Japan.
[22] R. Hamilia et al.: "Teager Energy and Ambiguity Function", *IEEE Trans. Signal Processing*, vol. - 47, no.-1 1999.
[23] S. Pal et. al, "Energy Based Discriminator: Some Issues", *International Journal of Engineering Research and Applications,* vol. 2, issue 3, pp. 1735-1741, May-June 2012.