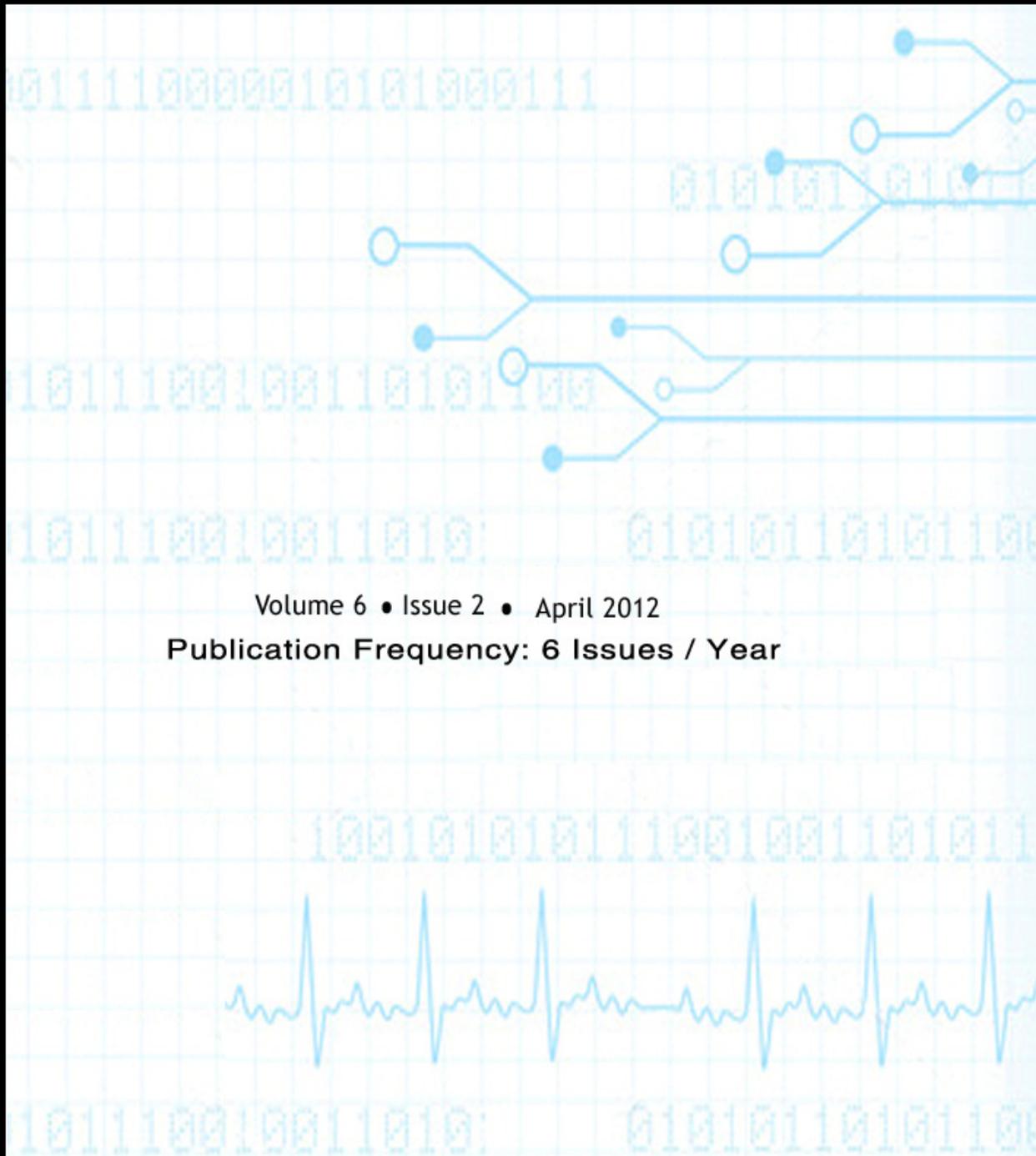


Editor-in-Chief
Dr Saif alZahir

SIGNAL PROCESSING (SPIJ)

AN INTERNATIONAL JOURNAL

ISSN : 1985-2339



Volume 6 • Issue 2 • April 2012

Publication Frequency: 6 Issues / Year

Copyrights © 2012 Computer Science Journals. All rights reserved.

CSC PUBLISHERS
<http://www.cscjournals.org>

SIGNAL PROCESSING: AN INTERNATIONAL JOURNAL (SPIJ)

VOLUME 6, ISSUE 2, 2012

**EDITED BY
DR. NABEEL TAHIR**

ISSN (Online): 1985-2339

International Journal of Computer Science and Security is published both in traditional paper form and in Internet. This journal is published at the website <http://www.cscjournals.org>, maintained by Computer Science Journals (CSC Journals), Malaysia.

SPIJ Journal is a part of CSC Publishers

Computer Science Journals

<http://www.cscjournals.org>

SIGNAL PROCESSING: AN INTERNATIONAL JOURNAL (SPIJ)

Book: Volume 6, Issue 2, April 2012

Publishing Date: 16-04-2012

ISSN (Online): 1985-2339

This work is subjected to copyright. All rights are reserved whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, re-use of illustrations, recitation, broadcasting, reproduction on microfilms or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provision of the copyright law 1965, in its current version, and permission of use must always be obtained from CSC Publishers.

SPIJ Journal is a part of CSC Publishers

<http://www.cscjournals.org>

© SPIJ Journal

Published in Malaysia

Typesetting: Camera-ready by author, data conversion by CSC Publishing Services – CSC Journals, Malaysia

CSC Publishers, 2012

EDITORIAL PREFACE

This is second issue of volume six of the Signal Processing: An International Journal (SPIJ). SPIJ is an International refereed journal for publication of current research in signal processing technologies. SPIJ publishes research papers dealing primarily with the technological aspects of signal processing (analogue and digital) in new and emerging technologies. Publications of SPIJ are beneficial for researchers, academics, scholars, advanced students, practitioners, and those seeking an update on current experience, state of the art research theories and future prospects in relation to computer science in general but specific to computer security studies. Some important topics covers by SPIJ are Signal Filtering, Signal Processing Systems, Signal Processing Technology and Signal Theory etc.

The initial efforts helped to shape the editorial policy and to sharpen the focus of the journal. Starting with volume 6, 2012, SPIJ appears in more focused issues. Besides normal publications, SPIJ intend to organized special issues on more focused topics. Each special issue will have a designated editor (editors) – either member of the editorial board or another recognized specialist in the respective field.

This journal publishes new dissertations and state of the art research to target its readership that not only includes researchers, industrialists and scientist but also advanced students and practitioners. The aim of SPIJ is to publish research which is not only technically proficient, but contains innovation or information for our international readers. In order to position SPIJ as one of the top International journal in signal processing, a group of highly valuable and senior International scholars are serving its Editorial Board who ensures that each issue must publish qualitative research articles from International research communities relevant to signal processing fields.

SPIJ editors understand that how much it is important for authors and researchers to have their work published with a minimum delay after submission of their papers. They also strongly believe that the direct communication between the editors and authors are important for the welfare, quality and wellbeing of the Journal and its readers. Therefore, all activities from paper submission to paper publication are controlled through electronic systems that include electronic submission, editorial panel and review system that ensures rapid decision with least delays in the publication processes.

To build its international reputation, we are disseminating the publication information through Google Books, Google Scholar, Directory of Open Access Journals (DOAJ), Open J Gate, ScientificCommons, Docstoc and many more. Our International Editors are working on establishing ISI listing and a good impact factor for SPIJ. We would like to remind you that the success of our journal depends directly on the number of quality articles submitted for review. Accordingly, we would like to request your participation by submitting quality manuscripts for review and encouraging your colleagues to submit quality manuscripts for review. One of the great benefits we can provide to our prospective authors is the mentoring nature of our review process. SPIJ provides authors with high quality, helpful reviews that are shaped to assist authors in improving their manuscripts.

Editorial Board Members

Signal Processing: An International Journal (SPIJ)

EDITORIAL BOARD

EDITOR-in-CHIEF (EiC)

Dr Saif alZahir

University of N. British Columbia (Canada)

ASSOCIATE EDITORS (AEiCs)

Professor. Wilmar Hernandez

Universidad Politecnica de Madrid
Spain

Dr Tao WANG

Universite Catholique de Louvain
Belgium

Dr Francis F. Li

The University of Salford
United Kingdom

EDITORIAL BOARD MEMBERS (EBMs)

Dr Jan Jurjens

University Dortmund
Germany

Dr Jyoti Singhai

Maulana Azad National institute of Technology
India

Assistant Professor Weimin Huang

Memorial University
Canada

Dr Lihong Zhang

Memorial University
Canada

Dr Bing-Zhao Li

Beijing Institute of Technology
China

Dr Deyun Wei

Harbin Institute of Technology
China

TABLE OF CONTENTS

Volume 6, Issue 2, April 2012

Pages

- 22 - 43 Teager Energy Operation on Wavelet Packet Coefficients for Enhancing Noisy Speech Using a Hard Thresholding Function
Tahsina Farah Sanam, Celia Shahnaz
- 44 - 55 DSP Based Speech Operated Home Appliances Using Zero Crossing Features
Deepali Y Loni
- 56 - 64 Peak-to-Average Power Ratio Reduction in NC-OFDM based Cognitive Radio.
Mohammad Zavid Parvez, Md. Abdullah Al Baki, Mohammad Hossain
- 65 - 77 A Simple Design to Mitigate Problems of Conventional Digital Phase Locked Loop
Mohamed Saber Saber Elsayes, Yutaka Jitsumatsu, Mohamed tahir Abasi Khan
- 78 - 85 Adaptive Variable Step Size in LMS Algorithm Using Evolutionary Programming: VSSLMSEV
Ajjaiah H.B.M, Prabhakar V. Hunagund, Manoj Kumar Singh, P.V.rao

Teager Energy Operation on Wavelet Packet Coefficients for Enhancing Noisy Speech Using a Hard Thresholding Function

Tahsina Farah Sanam

*Institute of Appropriate Technology
Bangladesh University of Engineering and Technology
Dhaka, Bangladesh*

tahsina@iat.buet.ac.bd

Celia Shahnaz

*Department of Electrical and Electronic Engineering
Bangladesh University of Engineering and Technology
Dhaka, Bangladesh*

celia@eee.buet.ac.bd

Abstract

In this paper a new thresholding based speech enhancement approach is presented, where the threshold is statistically determined by employing the Teager energy operation on the Wavelet Packet (WP) coefficients of noisy speech. The threshold thus obtained is applied on the WP coefficients of the noisy speech by using a hard thresholding function in order to obtain an enhanced speech. Detailed simulations are carried out in the presence of white, car, pink, and babble noises to evaluate the performance of the proposed method. Standard objective measures, spectrogram representations and subjective listening tests show that the proposed method outperforms the existing state-of-the-art thresholding based speech enhancement approaches for noisy speech from high to low levels of SNR.

Keywords: Teager Energy Operator, Wavelet Packet Transform, Statistical Modeling, Thresholding Function

1. INTRODUCTION

Estimating a signal that is corrupted by additive noise has been of interest to many researchers for practical as well as theoretical reasons. The problem is to recover the original signal from the noisy data. We want the recovered signal to be as close as possible to the original signal, retaining most of its important properties. There has been an increasing interest in noisy speech enhancement in a broad range of speech communication applications, such as mobile telephony, speech coding and recognition, and hearing aid devices [1]-[5]. Since the presence of noise seriously degrades the performance of the systems in such applications, the efficacy of the systems operating in a noisy environment is highly dependent on the speech enhancement techniques employed therein.

Various speech enhancement methods have been reported in the literature describing the know how to solve the problem of noise reduction in the speech enhancement methods. Speech enhancement methods can be generally divided into several categories based on their domains of operation, namely time domain, frequency domain and time-frequency domain. Time domain methods includes the subspace approach [6]-[10], frequency domain methods includes speech enhancement methods based on discrete cosine transform [11], the spectral subtraction [12]-[16], minimum mean square error (MMSE) estimator [17]-[21], Wiener filtering [22]-[25] and time frequency-domain methods involve the employment of the family of wavelet [26]-[34]. All the methods have their own advantages and drawbacks. In the subspace method [6]-[10], a mechanism to obtain a tradeoff between speech distortion and residual noise is proposed with the cost of a heavy computational load. Frequency domain methods, on the other hand, usually need less computation. In particular, although spectral subtraction method [12]-[16], is simple and provides a tradeoff between speech distortion and residual noise to some extent, it suffers from an artifact known as "musical noise" having an unnatural structure that is perceptually annoying, composed of tones at random frequencies and has an increased variance. In the MMSE estimator [17]-[21], the frequency spectrum of the noisy speech is modified to reduce the noise from noisy speech in the frequency domain. A relatively large variance of spectral coefficients is the problem of such an estimator. While adapting filter gains of the MMSE estimator, spectral outliers may emerge, that is especially difficult to avoid under noisy conditions.

One of the major problems of Wiener filter based methods [22]-[25] is the requirement of obtaining clean speech statistics necessary for their implementation. Both the MMSE and the Wiener estimators have a moderate computation load, but they offer no mechanism to control tradeoff between speech distortion and residual noise. Among the methods using time-frequency analyses, an approach of reducing different types of noise that corrupt the clean speech is the use of nonlinear techniques based on Discrete Wavelet Transform (DWT) [26]-[34], which is a superior alternative to the analyses based on Short Time Fourier Transform (STFT). The main challenge in such denoising approaches based on the thresholding of the wavelet coefficients of the noisy speech is the estimation of a threshold value that marks a difference between the wavelet coefficients of noise and that of clean speech. Then, by using the threshold, the designing of a thresholding scheme to minimize the effect of wavelet coefficients corresponding to the noise is another difficult task considering the fact that the conventional DWT based denoising approaches exhibit a satisfactory performance only at a relatively high signal-to-noise ratio (SNR).

For zero-mean, normally distributed white noise, Donoho and Johnstone proposed Universal threshold based method [30] for denoising the corrupted speech. For noisy speech, applying a unique threshold for all the wavelet or WP coefficients irrespective of the speech and silence segments may suppress noise to some extent, but it may also remove unvoiced speech segments thus degrading the quality of the enhanced speech. Statistical modeling is another approach of thresholding [28], where the threshold of WP coefficients is determined using the similarity distances between the probability distributions of the signals. Since speech is not always present in the signal, the thresholding must be adapted over time so that it is larger during portions without speech and smaller for those with speech. This will eliminate as much of the noise as possible while still maintaining speech intelligibility. However, the method in [28] requires an estimate of noise variance to distinguish speech frames from that of the noise ones with a view to set different thresholds for them. In order to decide a time adaptive threshold considering speech or silence frame, some estimate of the signal energy over time is necessary. A popular technique to estimate the required energy of the signal is Teager Energy Operator [35]. In [29], Teager energy operator (TEO) proposed by Kaiser [35] is employed to compute a time-adaptive threshold (TAT) value to threshold the WP coefficients of the noisy speech. But, TAT method suffers from an over thresholding problem if the speech signal is just contaminated by slight noises as this method uses an absolute offset parameter to distinguish speech frames from that of the noise ones.

In this paper, we develop a new thresholding method in the wavelet packet domain, where the threshold is adapted with respect to speech and silent segments. Since, TEO is a popular way to estimate the speech signal energy, instead of direct employment of the TEO on the noisy speech, we apply the TEO on the WP coefficients of the noisy speech. Unlike the approach of threshold determination directly from the WP coefficients of the noisy speech, we determine an appropriate threshold by performing the statistical modeling of the TE operated WP coefficients of noisy speech and employed a hard thresholding function for obtaining an enhanced speech.

2. BRIEF BACKGROUND

2.1. Wavelet Packet Transform

A method based on the Wavelet Packet Transform is a generalization of the Wavelet Transform based decomposition process that offers a richer range of probabilities for the analysis of signals, namely speech. In wavelet analysis, a speech signal is split into sets of approximation and detail coefficients. The set of approximation coefficients is then itself split into a second-level approximation and detail coefficients, and the process is repeated. Mallat algorithm is one of the efficient ways to construct the discrete wavelet transform (DWT) by iterating a two-channel perfect reconstruction filter bank over the low pass scaling function branch. However, this algorithm results in a logarithmic frequency resolution, which does not work well for all the signals. In order to overcome the drawback as mentioned above, it is desirable to iterate the high pass wavelet branch of the Mallat algorithm tree as well as the low pass scaling function branch. Such a wavelet decomposition produced by these arbitrary subband trees is known as wavelet packet (WP) decomposition.

In wavelet analysis, only scale space is decomposed, but wavelet space is not decomposed. By the restriction of Heisenberg's uncertainty principle, the spatial resolution and spectral resolution of high frequency band become poor thus limiting the application of wavelet transform. In particular, there are some problems with the basic wavelet thresholding method, when it is applied to the noisy speech for

the purpose of enhancement. An important shortcoming is the shrinkage of the unvoiced segments of speech which contain many noise-like speech components leading to a degraded speech quality. On the other hand, in wavelet packet analysis, the wavelet space is also decomposed thus making the higher frequency band decomposition possible. Since, both the approximation and the detail coefficients are decomposed into two parts at each level of decomposition, a complete binary tree with superior frequency localization can be achieved. This particular feature of the WP transform is indeed useful for enhancing speech in the presence of noise.

2.2. Teager Energy Operator

The Teager Energy Operator (TEO) is a powerful nonlinear operator proposed by Kaiser [36], capable to extract the signal energy based on mechanical and physical considerations. The continuous form of the TEO is given as,

$$\Psi_c[y(t)] = \left(\frac{d}{dt}y(t)\right)^2 - y(t)\frac{d^2}{dt^2}y(t), \tag{1}$$

where, $\Psi_c[\cdot]$ and $y(t)$ represent the continuous TEO and a continuous signal, respectively. For a given bandlimited discrete signal $y[n]$, the discrete-time TEO can be approximated by,

$$\Psi_d(y[n]) = y[n]^2 - y[n+1]y[n-1], \tag{2}$$

The discrete time TEO in (2) is nearly instantaneous since only three samples are required for the energy computation at each time instant. Due to this excellent time resolution, the output of a TEO provides us with the ability to capture the energy fluctuations and hence gives an estimate of the energy required to generate the signal. Note that, in case of speech signal, directly using the TEO on original speech may result in much undesired artifact and enhanced noises as TEO is a fixed-sized local operator.

In context of the speech enhancement by thresholding via WP analysis, the threshold must be adapted over time, since speech is not always present in the signal. It is expected that the threshold should be larger during periods without speech and smaller for those with speech. In order to obtain an idea of speech/nonspeech activity for deciding the corresponding threshold value, it is required to estimate the signal energy over time. Since TEO is a popular way to estimate the speech signal energy, instead of direct employment of the TEO on the original speech, it is reasonable to apply the TEO on the WP coefficients. In comparison to the approach of threshold determination from the WP coefficients of noisy speech, the approach intended to determine threshold from the TE operated WP coefficients has the potential to eliminate as much of the noise as possible, while still maintaining speech intelligibility in enhanced speech.

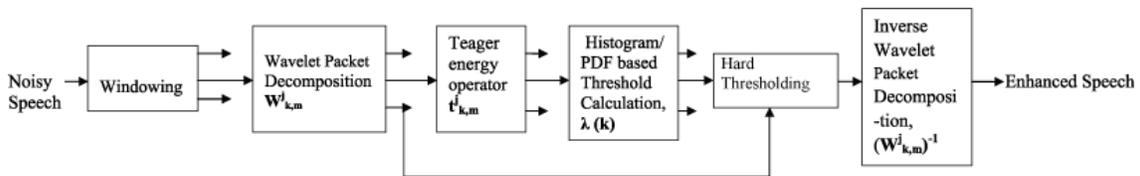


FIGURE 1: Block Diagram of the Proposed Method

3. PROPOSED METHOD

The block diagram of our proposed system is shown in Fig. 1. It is seen from Fig. 1 that WP transform is first applied to each input speech frame. Then, the WP coefficients are subject to Teager Energy approximation with a view to determine a threshold value for performing thresholding operation in the WP domain. On thresholding, an enhanced speech frame is obtained via inverse wavelet packet (IWP) transform.

3.1. Wavelet Packet Analysis

For a j level WP transform, the noisy speech signal $y[n]$ with frame length N is decomposed into 2^j subbands. The m -th WP coefficient of the k -th subband is expressed as,

$$W^j_{k,m} = WP[y[n], j], \quad n = 1, \dots, N. \tag{3}$$

where, $m = 1, \dots, N/2^j$ and $k = 1, \dots, 2^j$. Since the frequency resolution provided by the full WP transform is not sufficient to separate speeches from low-frequency noises, WP coefficients of the noisy speech have some noise handling limitations. To this end, we apply the discrete time TEO on the obtained WP coefficients, $W_{k,m}^j$.

3.2. Teager Energy (TE) Approximation

The application of the discrete-time TEO on the $W_{k,m}^j$ results in a set of TEO coefficients $t_{k,m}^j$. The m -th TEO coefficient corresponding to the k -th subband of the WP is given by,

$$t_{k,m}^j = \Psi_d[W_{k,m}^j], \quad k = 1, \dots, 2^j \quad (4)$$

i.e; using (2)

$$t_{k,m}^j = [W_{k,m}^j]^2 - [W_{k,(m-1)}^j][W_{k,(m+1)}^j], \quad k = 1, \dots, 2^j \quad (5)$$

In comparison to the operation of WP transform on the noisy speech, the TEO operation on the WP coefficients of the noisy speech is able to enhance the discriminability of speech coefficients among those of noise. This energy tracking operator can be understood when considering sinusoidal oscillation that occur with a simple harmonic oscillator. If TEO is applied to a discrete time signal, it causes an effect as if a single sinusoid of amplitude A and frequency ω is passed through three adjacent points with index $n - 1, n, n + 1$ thus yielding an output sequence that is a varying signal proportional to $A^2 \sin^2 \omega$, where the frequency ω is normalized with respect to the sampling frequency. This signal in essence is a measure of the 'energy' in that signal as a function of time. It is thus important that the original discrete signal consist primarily of a single component. Since a wavelet coefficient stream in a single subband is primarily a single component signal, it is valid to apply TEO on the WP coefficients of the noisy speech.

3.3. Statistical Modeling of TEO Operated WP Coefficients

The outcome of a speech denoising method based on the thresholding in a transform domain depends mainly on two factors, namely the threshold value and the thresholding functions. The use of a unique threshold for all the WP subbands is not reasonable. As a crucial parameter, the threshold value in each subband is required to be adjusted very precisely so that it can prevent distortion in the enhanced speech as well as decrease annoying residual noise. In order to remove the noisy coefficients with low distortion in the enhanced speech signal, the value of threshold has to be different in the speech and silent frames. The value of the threshold in the silent frames is smaller than it in the speech frames. Also, the use of conventional thresholding functions, for example, Hard and Soft thresholding functions often results in time frequency discontinuities. In order to handle such problems, we propose a new thresholding function employing a threshold value determined for each subband of the WP by statistically modeling the TE operated WP coefficients $t_{k,m}^j$ with a probability distribution rather than choosing a threshold value directly from the $t_{k,m}^j$. This idea is also exploited to determine a threshold value for each subband of an silent frame which is different from that of each subband of a speech frame.

The main issue in wavelet thresholding is estimating an appropriate threshold value λ . In the range of $-\lambda$ and λ , the noisy speech wavelet coefficients are similar to the noise wavelet coefficients and outside of this range the wavelet coefficients of the noisy speech are similar to that of the clean speech [38]. So it is expected that in the range of $-\lambda$ and λ , [30],[26],[37], the probability distribution of the noisy speech coefficients would to be nearly similar to that of the noise coefficients. Furthermore, the probability distribution of the noisy speech coefficients is expected to be similar to that of the clean speech coefficients outside of this range. In a certain range, the probability distribution of the $t_{k,m}^j$ of the noisy speech is expected to be nearly similar to those of the noise. Also, outside that range, the probability distribution of the $t_{k,m}^j$ of the noisy speech is expected to be similar to those of the clean speech. Thus by considering the probability distributions of the $t_{k,m}^j$ of the noisy speech, noise and clean speech, a more accurate threshold value can be obtained using a suitable scheme of pattern matching or similarity measure between the probability distributions. Since speech is a time-varying signal, it is difficult to realize the actual probability distribution function (pdf) of

speech or its $t_{k,m}^j$. As an alternative to formulate a pdf of the $t_{k,m}^j$ of speech, we can easily formulate the histogram of the $t_{k,m}^j$ and approximate the histogram by a reasonably close probability distribution function, namely Gaussian distribution [28]. In frequency domain, the use of a Gaussian statistical model is motivated by the central limit theorem since each Fourier expansion coefficient can be seen as a weighted sum of random variables resulting from the observed samples [39]. Other distributions were also proposed for the real and imaginary parts of the STFT coefficients [40], [24], the STSA coefficients [40], and the complex STFT coefficients [42], [17]. While it has been proposed that the Fourier expansion coefficients of speech signals may not be Gaussian-distributed, those assumptions are usually motivated by long-term averages of the speech signal which may not be applicable to specific short-time utterances. Moreover, many estimators using a Gaussian distribution do not have an analytical counterpart when using other distributions [41]. Therefore, many researchers consider only Gaussian distributed complex STFT coefficients in their works [43]. However, in our work, we have implemented our algorithm for 30 sentences of the NOIZEUS database and 4 different noise signals (white, car, pink, and babler noises) and it is verified that, in each subband, the pdfs of the $t_{k,m}^j$ of the noise, clean speech and noisy speech can be sufficiently well described by the Gaussian distribution. Fig. 2, Fig. 3, and Fig. 4 shows instances of these results for clean speech, noisy speech, and noise, respectively.

3.4. Adaptive Threshold Calculation

Analysis on the speech signal shows that, the value of entropy has the ability to detect speech/silence frames [38], [45]. Also, the entropy of each subband of the $t_{k,m}^j$ is found different from each other. So, an entropy measure may be chosen to select a suitable threshold value adaptive to each subband as well as adaptive to the speech/silence frames. Some popular similarity measures that are related to the entropy functions are the Variational distance, the Bhattacharyya distance, the Harmonic mean, the Kullback Leibler(K-L) divergence, and the Symmetric K-L divergence [47]. All these measures are used to estimate the similarity between two pdfs. As all of these distance measures have nonnegative values, zero flag is a very suitable distinctive for recognizing the similarity between two pdfs. Note that, if two pdfs are exactly the same, only two measures (the K-L divergence and the symmetric K-L divergence) will be equal to zero. So, in order to determine an adaptive threshold value based on the idea of entropy quantified by an appropriate similarity measure, we proceed as follows,

1. The average of the $t_{k,m}^j$ of different segments is calculated.
2. The histograms of the averaged $t_{k,m}^j$ in each sub-band is obtained. The number of bins in the histogram has been set equal to the square root of the number of samples divided by two.
3. Since $t_{k,m}^j$ of clean speech, noisy speech and noise are positive quantity, there histograms in each sub-band can be approximated by the positive part of a pdf following the Gaussian distribution as shown in Fig. 2, 3 and 4.

The K-L divergences is always nonnegative and zero if and only if the approximate Gaussian distribution functions of the $t_{k,m}^j$ of noisy speech and that of the noise or the approximate Gaussian distribution functions of the $t_{k,m}^j$ of the noisy speech and that of the clean speech are exactly the same. In order to have a symmetric distance between the any two approximate Gaussian distribution functions as mentioned above, the Symmetric K-L divergence has been adopted in this paper. The Symmetric K-L divergence is defined as,

$$SKL(p, q) = \frac{KL(p, q) + KL(q, p)}{2}, \tag{6}$$

where, p and q are the two approximate Gaussian pdfs calculated from the corresponding histograms each having M number of bins and $KL(\cdot)$ is the K-L divergence given by,

$$KL(p, q) = \sum_{i=1}^M p_i(t_{k,m}^j) \ln \frac{p_i(t_{k,m}^j)}{q_i(t_{k,m}^j)} \tag{7}$$

In (7), $p_i(t_{k,m}^j)$ represents the approximate Gaussian pdf of the $t_{k,m}^j$ of the noisy speech estimated by,

$$\hat{p}_i(t_{k,m}^j) = \frac{\text{Number of coefficients in the } i\text{th bin of histogram}}{\text{Total number of coefficients in each subband}} \tag{8}$$

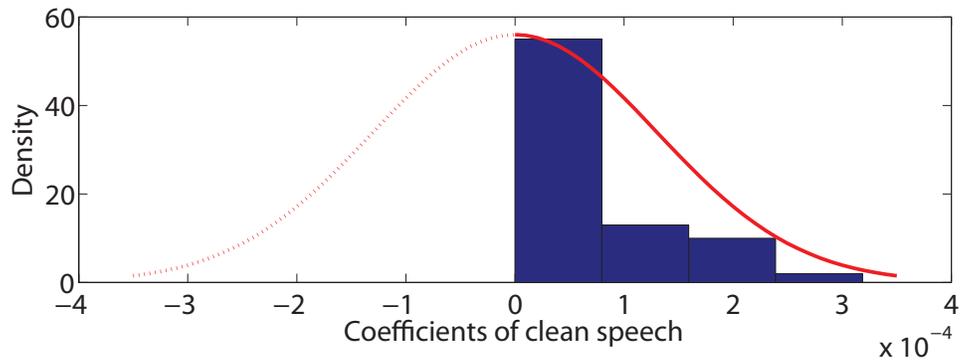


FIGURE 2: Probability distribution of TEO coefficients of clean speech

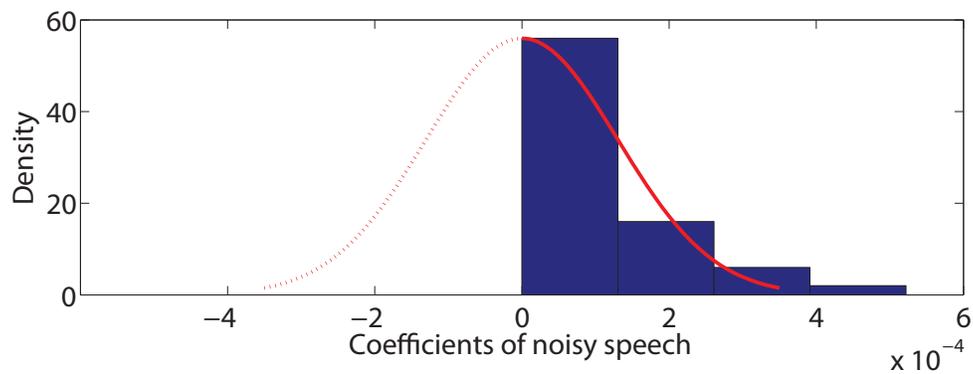


FIGURE 3: Probability distribution of TEO coefficients of noisy speech

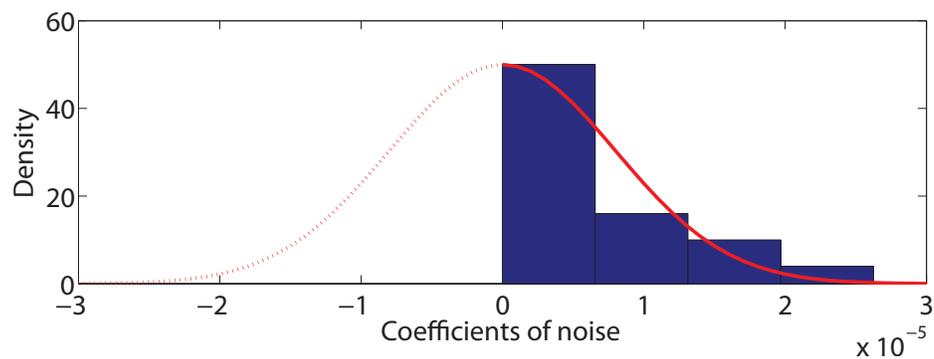


FIGURE 4: Probability distribution of TEO coefficients of noise

Similarly, the approximate Gaussian pdf of the $t_{k,m}^n$ of the noise and that of the $t_{k,m}^c$ of the clean speech can be estimated from (8) and denoted by $\hat{q}_i(t_{k,m}^n)$ and $\hat{q}_i(t_{k,m}^c)$, respectively. Below a certain value of threshold λ , the $t_{k,m}^n$ of the noisy speech, the Symmetric K-L divergence between $\hat{q}_i(t_{k,m}^n)$ and $\hat{q}_i(t_{k,m}^c)$ is approximately zero, i.e,

$$SKL(\hat{q}_i(t_{k,m}^n), \hat{q}_i(t_{k,m}^c)) \approx 0, \quad (9)$$

where the bins lie in the range $[1, \lambda]$ in both $\hat{p}_i(t_{k,m}^j)$ and $\hat{q}_i(t_{k,m}^j)$. Alternatively, above the value λ of the $t_{k,m}^j$ of the noisy speech, the Symmetric K-L divergence between $\hat{p}_i(t_{k,m}^j)$ and $\hat{q}_i(t_{k,m}^j)$ is closely zero, i.e.,

$$SKL(\hat{p}_i(t_{k,m}^j), \hat{q}_i(t_{k,m}^j)) \approx 0, \tag{10}$$

In (10), the bins lie in the range $[\lambda+1, M]$ in both $\hat{p}_i(t_{k,m}^j)$ and $\hat{q}_i(t_{k,m}^j)$. Using (6) and (7) in evaluating (9) and (10), we get,

$$\sum_{i=1}^{\lambda} [\hat{p}_i(t_{k,m}^j) - \hat{q}_i(t_{k,m}^j)] \ln \frac{\hat{p}_i(t_{k,m}^j)}{\hat{q}_i(t_{k,m}^j)} \approx 0 \tag{11}$$

$$\sum_{i=\lambda+1}^M [\hat{p}_i(t_{k,m}^j) - \hat{q}_i(t_{k,m}^j)] \ln \frac{\hat{p}_i(t_{k,m}^j)}{\hat{q}_i(t_{k,m}^j)} \approx 0 \tag{12}$$

From (11), it is apparent that $t_{k,m}^j$ of the noisy speech lying in the range $[1, \lambda]$ can be marked as $t_{k,m}^j$ of noise and needed to be removed. Similarly, (12) attests that the $t_{k,m}^j$ of the noisy speech residing outside $[1, \lambda]$ can be treated as similar to the $t_{k,m}^j$ of the clean speech and considered to be preserved. For obtaining a general formula for the threshold value λ in each subband, we use continuous real mode in (11) and (12), thus obtain,

$$\int_1^{\lambda} \left[\frac{\sqrt{\vartheta}}{\sqrt{2\pi}\sigma_S} \exp\left(-\frac{\vartheta x^2}{2\sigma_S^2}\right) - \frac{1}{2\pi\sigma_N} \exp\left(-\frac{x^2}{2\sigma_N^2}\right) \right] \ln \left((1 - \sqrt{\vartheta}) \exp\left(-\frac{\vartheta x^2}{2\sigma_S^2} + \frac{x^2}{2\sigma_N^2}\right) \right) dx \approx 0 \tag{13}$$

$$\int_{\lambda+1}^{\infty} \left[\frac{\sqrt{\vartheta}}{\sqrt{2\pi}\sigma_S} \exp\left(-\frac{\vartheta x^2}{2\sigma_S^2}\right) - \frac{1}{2\pi\sigma_S} \exp\left(-\frac{x^2}{2\sigma_S^2}\right) \right] \ln \left((\sqrt{\vartheta}) \exp\left(\frac{(1 - \vartheta)x^2}{2\sigma_S^2}\right) \right) dx \approx 0 \tag{14}$$

where,

$$\vartheta = \sigma_S^2 / (\sigma_N^2 + \sigma_S^2) \tag{15}$$

where, $\vartheta = \sigma_S^2 / (\sigma_N^2 + \sigma_S^2)$, σ_N [44], [46] and σ_S are the variances of noise and clean speech in each subband, respectively. For computing λ , we first simplify the equations (13) and (14) to solve. Since the symmetric K-L is a nonnegative distance, in a specified range, its minimum value can be found to be nearly zero. Thus the value of $t_{k,m}^j$ for which the threshold reaches its optimum value can be determined by minimizing (13) or (14). It is well known that an optimum value of a function in a given range can be calculated by setting its derivative, with respect to the variable expected to optimize the function value, to zero. Since (13) is a definite integral, the derivative of the function defined in the left hand side (L.H.S) of (13) representing the Symmetric K-L divergence between $\hat{p}_i(t_{k,m}^j)$ and $\hat{q}_i(t_{k,m}^j)$ is zero. On the other hand, the derivative of the function obtained in the L.H.S of (14) representing the Symmetric K-L distance between $\hat{p}_i(t_{k,m}^j)$ and $\hat{q}_i(t_{k,m}^j)$ is calculated and set to zero. By simplifying the either derivatives, an optimum value of λ can be obtained which is adaptive to each subband of a frame.

$$\lambda(k) = \sigma_N(k) \sqrt{2(\gamma_k + \gamma_k^2) \ln \left(\sqrt{1 + \frac{1}{\gamma_k}} \right)} \tag{16}$$

where, σ_N is the variance of noise in each subband], k is the sub-band index and γ_k is the segmental SNR calculated as,

$$\gamma_k = \sigma_S^2(k) / \sigma_N^2(k) \tag{17}$$

We calculate the second order derivation of the L.H.S of (14) with respect to the obtained threshold to demonstrate that the calculated threshold minimize (14). As the second order derivation of the L.H.S of (14) is nonnegative, the obtained thresholds are valid. In order to have smaller threshold for higher input SNR values, we have to adjust the threshold obtained by (16). Since the variance of the noise is inversely proportional to the input SNR, we can modify (16) as,

$$\lambda(k) = \left[\sigma_N(k) / \sqrt{\gamma_k} \right] \sqrt{2(\gamma_k + \gamma_k^2) \ln \left(\sqrt{1 + \frac{1}{\gamma_k}} \right)} \quad (18)$$

Since in the silent segment of a noisy speech, only noise exists, a threshold value different than that used in the speech segment should be selected in order to eliminate the noise completely. The Symmetric K-L divergence between the $t_{k,m}^j$ of the noisy speech and that of the $t_{k,m}^j$ of the noise is nearly zero in the non speech subbands. Exploiting this idea a suitable time adaptive threshold value λ' can be obtained as,

$$\lambda'(k) = \begin{cases} \max(t_{k,m}^j), & \text{SKL}(\hat{q}_i(t_{k,m}^j), \hat{q}_i(t_{k,m}^j)) \approx 0 \\ \lambda(k), & \text{Otherwise.} \end{cases} \quad (19)$$

3.5. Denoising by Thresholding

Removing noise components by thresholding operation of the WP coefficients is based on the fact that for many signals (such as speech), the energy is mostly concentrated in a low frequency region that corresponds to a small number of lower WP coefficients. So, by thresholding the WP coefficients, we can reduce the effect of the high frequency noise components on the speech signal components. We employ hard thresholding for denoising purpose.

Hard thresholding sets zero to the noisy speech WP coefficients whose absolute value is below the threshold. Noting the threshold determined by (19) as $\lambda_1(k)$ and using it, the hard thresholding function can be applied on the m -th WP coefficients of the k -th subband $V_{k,m}^j$ as,

$$\hat{V}_{k,m}^j = \begin{cases} V_{k,m}^j & |V_{k,m}^j| \geq \lambda_1(k) \\ 0, & |V_{k,m}^j| < \lambda_1(k). \end{cases} \quad (20)$$

Here, $\hat{V}_{k,m}^j$ stands for the m -th WP coefficients of the k -th subband after the hard thresholding operation.

3.6. Inverse Wavelet Transform

The enhanced speech frame is synthesized by performing the inverse WP transformation WP^{-1} on the resulting thresholded WP coefficients, $\hat{V}_{k,m}^j$

$$\hat{S}[n] = WP^{-1}(\hat{V}_{k,m}^j) \quad (21)$$

where, $\hat{S}[n]$ represents the enhanced speech frame. The final enhanced speech signal is reconstructed by using the standard overlap-and-add method.

4. SIMULATION

In this Section, a number of simulations is carried out to evaluate the performance of the proposed method.

4.1. Simulation Conditions

Real speech sentences from the *NOIZEUS* database are employed for the experiments, where the speech data is sampled at 8 KHz. To imitate a noisy environment, noise sequence is added to the clean speech samples at different signal to noise ratio (SNR) levels ranging from 15 dB to -15 dB. Four different types of noises, such as, white, car, and pink are adopted from the *NOISEX92* [20] and *NOIZEUS* databases.

In order to obtain overlapping analysis frames, hamming windowing operation is performed, where the size of each of the frame is 512 samples with 50% overlap between successive frames. A 3-level WP decomposition tree with db10 bases function is applied on the noisy speech frames and the Teager energy operation is performed on the resulting WP coefficients. By computing the threshold from (19), a hard thresholding function is developed and applied on the WP coefficients of the noisy speech using (20).

4.2. Comparison Metrics

Standard Objective metrics [48], namely, overall SNR improvement in dB, Perceptual Evaluation of Speech Quality (PESQ) and Weighted Spectral Slope (WSS) are used for the evaluation of the proposed method. The proposed method is subjectively evaluated in terms of the spectrogram representations of the clean speech, noisy speech and enhanced speech. Informal listening tests are also carried out in order to find the analogy between the objective metrics and subjective sound quality. The performance of our method is compared with some of the existing thresholding based speech enhancement methods, such as, Universal [30], WTHSKL [29] and TAT [28] in both objective and subjective senses.

4.3. Objective Evaluation

4.3.1. Results on White Noise-corrupted Speech

The results in terms of all the objective metrics, such as, SNR improvement in dB, PESQ and WSS obtained by using the Universal, WTHSKL, TAT, and proposed methods for white noise-corrupted speech are presented in Fig. 5 through Fig. 6 and in Table 1.

Fig. 5 shows the SNR improvement in dB obtained by using different methods employing hard thresholding function in the presence of white noise, where the SNR varies from 15 dB to -15 dB. It is seen from this figure that in the SNR range under consideration, the improvement in SNR in dB is comparable for all the comparison methods, but they show comparatively lower values relative to the proposed method at all the levels of SNR.

The PESQ scores vs SNR obtained by using different methods are portrayed in Fig. 6. This figure shows that the proposed method using the hard thresholding function is capable of producing enhanced speech with better quality as it gives larger scores of PESQ for a wide range of SNR levels whereas, the PESQ scores resulting from all other methods are comparable and relatively lower even at a high SNR of 15 dB. It is also seen from Fig. 6 that the difference in PESQ scores of the proposed method and that of the other methods increases as SNR decreases, thus indicating the effectiveness of the proposed method using hard thresholding function in enhancing speech even in a severe noisy environment.

The WSS values obtained by using different methods are summarized in Table 1. for varying SNR of 15 dB to -15 dB. For a particular method in Table 1, the WSS increases as SNR decreases. At a particular SNR, such as -15 dB, the proposed method using hard function is superior in a sense that it gives the lowest WSS value, whereas the other methods produce comparatively higher values of WSS.

4.3.2. Results on Car Noise-corrupted Speech

Now, we present the results in terms of all the objective metrics as mentioned above obtained by using the Universal, WTHSKL, TAT, and the proposed methods in Table 2 and in Fig. 7 through Fig. 8 for car noise-corrupted speech.

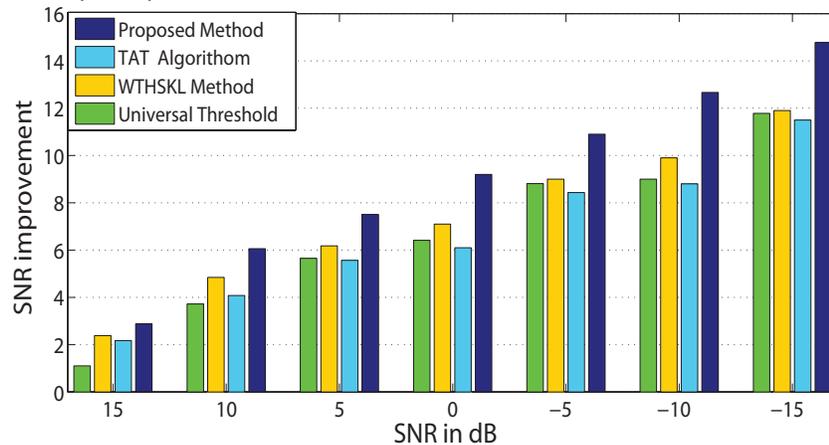


FIGURE 5: Performance comparison of different methods using hard thresholding function in terms of SNR Improvement in dB for white noise corrupted speech.

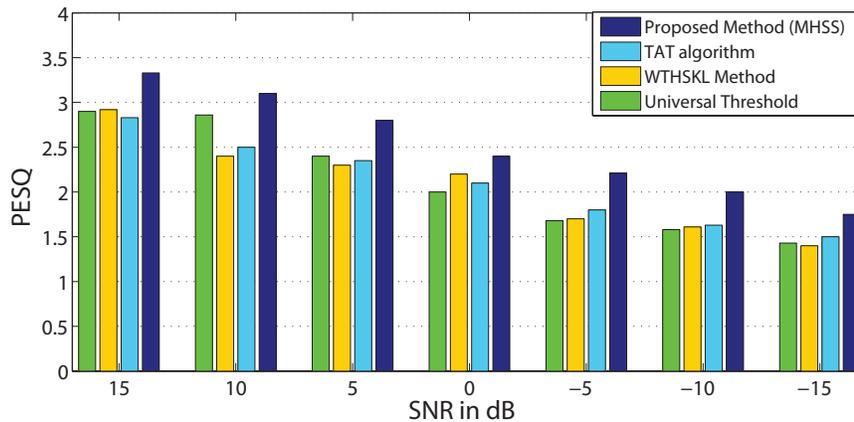


FIGURE 6: Performance comparison of different methods using hard thresholding function in terms of PESQ for white noise corrupted speech.

SNR [dB]	Universal	TAT	WTHSKL	Proposed Method
15	34.84	27.2115	27.1	22.7
10	42.1	34.5904	33.2337	29.1
5	51.3	43.2807	42.7895	37.3
0	62.5	55.1436	56.4079	46.82
-5	73.13	67.1598	66.8581	59.1
-10	90	85.43	86.1	78.1
-15	100.6	100.63	92.77	84.8

TABLE 1: Performane comparison of WSS for different methods in the presence of white noise.

In Table 2, the performance of the proposed method using hard thresholding function is compared with that of the other methods at different levels of SNR. For a method in Table 2, the SNR improvement in dB increases as SNR decreases. At a low SNR of -15 dB, the proposed method yields the highest SNR improvement in dB. Such larger values of SNR improvement in dB at a low level of SNR attest the capability of the proposed method in producing the enhanced speech with better quality even for car noise-corrupted speech.

SNR [dB]	Universal	TAT	WTHSKL	Proposed Method
15	1.9	3.4	0.6	3
10	3.8	4.01	2.9	4.51
5	4.9	5	4	6.8
0	6.7	5.7	5.5	8.1
-5	7	7.23	8.1	9.14
-10	8.9	9.6	9.2	10.78
-15	10.1	10.63	11.6	12.7

TABLE 2: Performance comparison of SNR improvement in dB for different methods in the presence of car noise.

In the presence of car noise, the PESQ scores at different SNR levels resulted by using the other methods are compared with respect to the proposed method employing hard thresholding function in Fig. 7. It can be seen from the figure that at a high level of SNR, such as 15 dB, Universal, WTHSKL and TAT methods show lower values of PESQ scores, whereas the PESQ score is much higher, as expected, for the proposed method. The proposed method also yields larger PESQ scores compared to that of the other methods at lower levels of SNR. Since, at a particular SNR, a higher PESQ score

indicates a better speech quality, the proposed method is indeed better in performance even in the presence of a car noise.

Fig. 8 represents the WSS values as a function of SNR for the proposed method employing hard thresholding function and that for the other methods. As shown in the figure that the WSS values resulting from all other methods are comparable and relatively larger for a wide range of SNR levels, whereas the proposed method is capable of producing enhanced speech with better quality as it gives lower values of WSS at a low SNR of -15 dB.

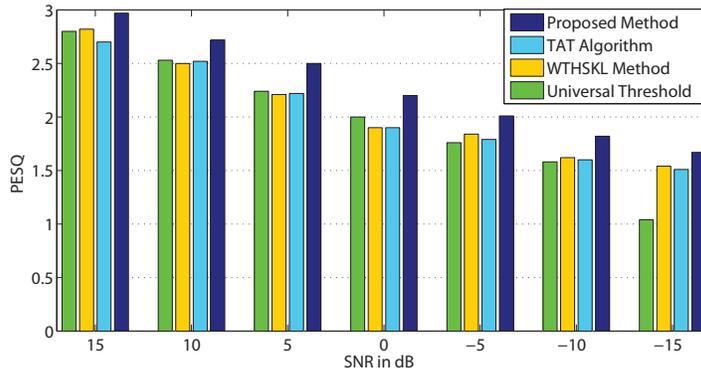


FIGURE 7: Performance comparison of different methods using hard thresholding function in terms of PESQ scores for car noise corrupted speech.

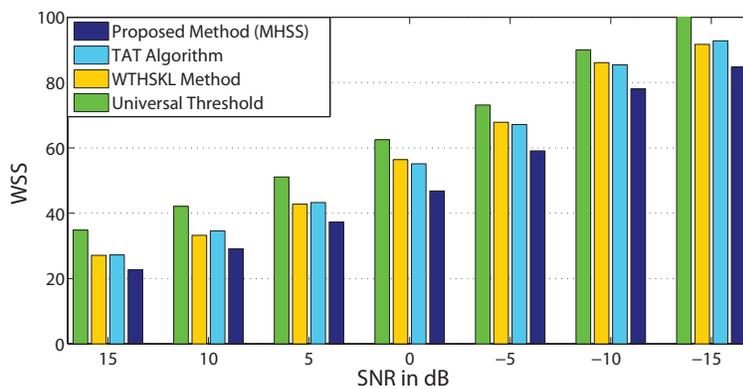


FIGURE 8: Performance comparison of different methods using hard thresholding function in terms of WSS for car noise corrupted speech.

4.3.3. Results on Pink Noise-corrupted Signal

All the objective metrics for evaluating the performance of the proposed method relative to the other methods for pink noise-corrupted speech are computed and depicted in Fig. 9 through Fig. 10 and in Table 3.

The SNR improvement in dB resulted by using different methods are summarized in Fig. 9. It is vivid from this figure that the other methods produce comparatively lower improvement in SNR in dB in the whole SNR range, while the proposed method using hard thresholding function continues to remain superior in a sense that it gives the highest improvement in SNR in dB even at an SNR as low as -15 dB of pink noise.

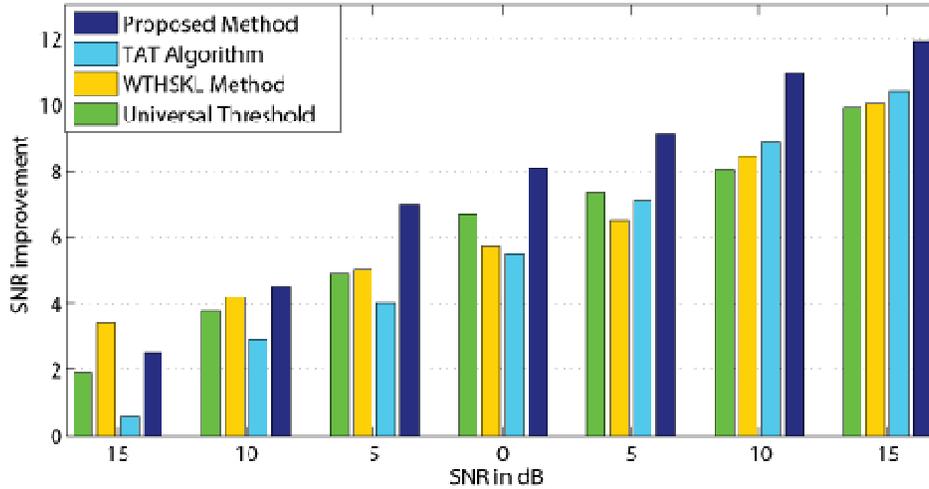


FIGURE 9: Performance comparison of the SNR Improvement in dB for different methods in the presence of pink noise.

The PESQ scores of the proposed method and that obtained by using different comparison methods are shown in Table 3 with respect to SNR levels varying from high (15 dB) to low (-15 dB). It is clear from the table that the other methods continue to provide lower PESQ scores, while the proposed method maintain comparatively higher PESQ scores even in the presence of severe pink noise of -15 dB.

The variation of the output WSS with respect to SNR levels for different methods and that for the proposed method using hard thresholding function is portrayed in Fig. 10. It is evident from analyzing each of these figures that, in the whole SNR range, the other methods continue to produce much higher WSS values with respect to the proposed method using hard thresholding function. Note that, the propose method performs the best in a sense that it yields the lowest WSS values almost at different SNR levels.

SNR [dB]	Universal	TAT	WTHSKL	Proposed Method
15	2.8	2.7	2.82	2.97
10	2.53	2.52	2.5	2.71
5	2.24	2.22	2.21	2.5
0	2	1.9	1.87	2.1
-5	1.55	1.65	1.67	1.9
-10	1.4	1.47	1.52	1.7
-15	1.3	1.39	1.43	1.5

TABLE 3: Performance comparison of PESQ scores for different methods in the presence of pink noise.

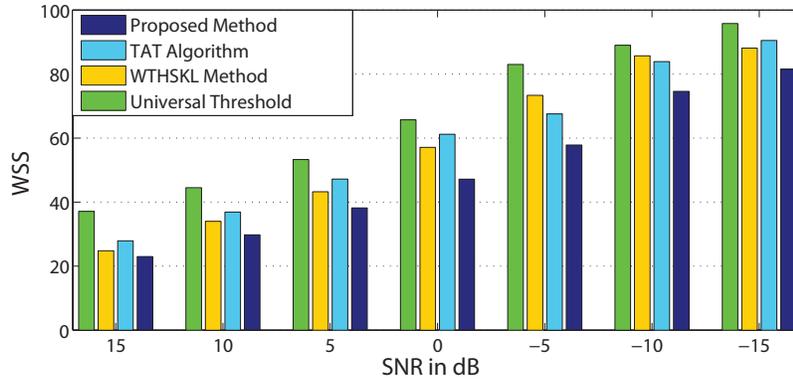


FIGURE 10. Performance comparison of different methods using hard thresholding function in terms of WSS values for pink noise corrupted speech.

4.3.4. Results on Multi-talker Babble Noise-corrupted Speech

The results obtained from the multi-talker babble noise-corrupted speech in terms of the SNR improvement in dB, PESQ scores, and WSS values for the proposed method using hard thresholding function and that for the other methods are depicted in Fig. 11 through Fig. 13 at particular SNR levels of 15 dB, 0 dB and -15 dB. It is noticeable from these figures that the performance of all the methods degrades in the presence of multi-talker babble noise compared to that in the pink or car or white noise, but the proposed method retains its superiority with respect to all the levels of SNRs.

Fig. 11 provides plot for the SNR improvement in dB for all the methods for babble noise-corrupted speech. It is seen that the proposed method maintains better performance at all the SNR levels considered. Also the proposed method still remains the best thus showing higher capability of producing enhanced speech with better quality at a very low level of SNR of 0 dB or even lower than that.

In similar babble noisy condition, the PESQ scores resulting from using the speech enhancement methods under consideration are shown in Fig. 12. As seen, the proposed method continues to provide better results for the low levels of SNR, such as -15 dB.

Also, the WSS values obtained from all the methods as a function of SNR are plotted in Fig. 13 for babble noise-corrupted speech. This figure illustrates that, as expected, the WSS values of the proposed method are somewhat increased in comparison to the other noisy cases, but its performance still remains better than that provided by the other methods for a wide range of SNR values from 15 dB to -15 dB.

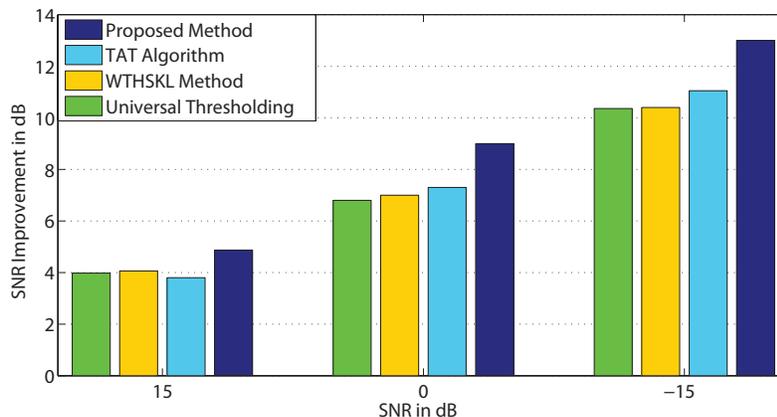


FIGURE 11. Performance comparison of different methods using hard thresholding function in terms of SNR improvement in dB for babble noise corrupted speech.

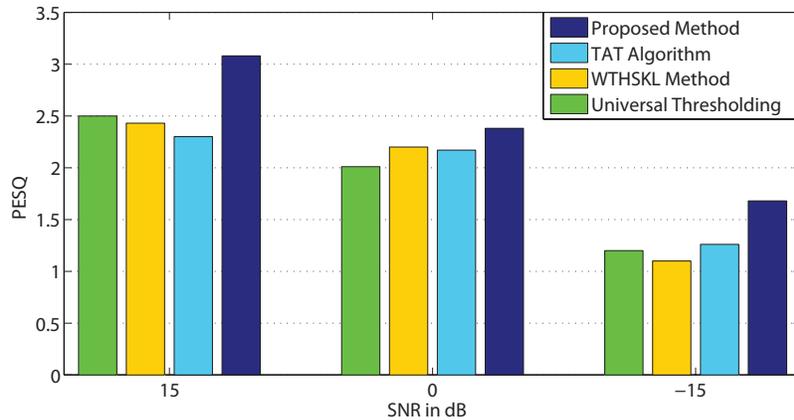


FIGURE 12. Performance comparison of different methods using hard thresholding function in terms of PESQ scores for babble noise corrupted speech.

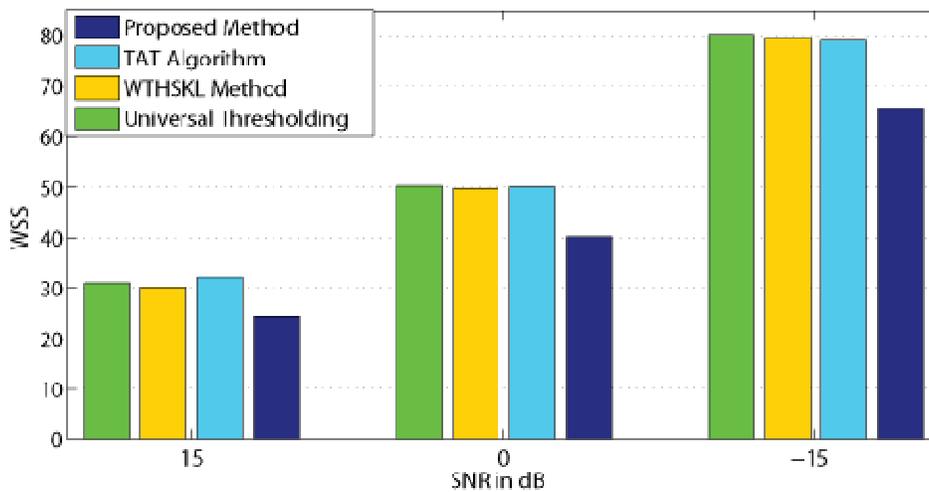


FIGURE 13. Performance comparison of different methods using hard thresholding function in terms of WSS values for babble noise corrupted speech.

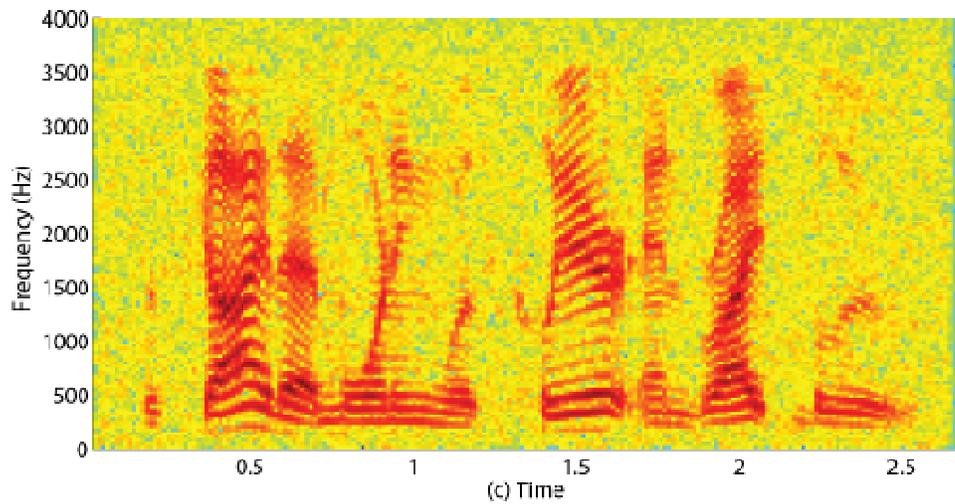
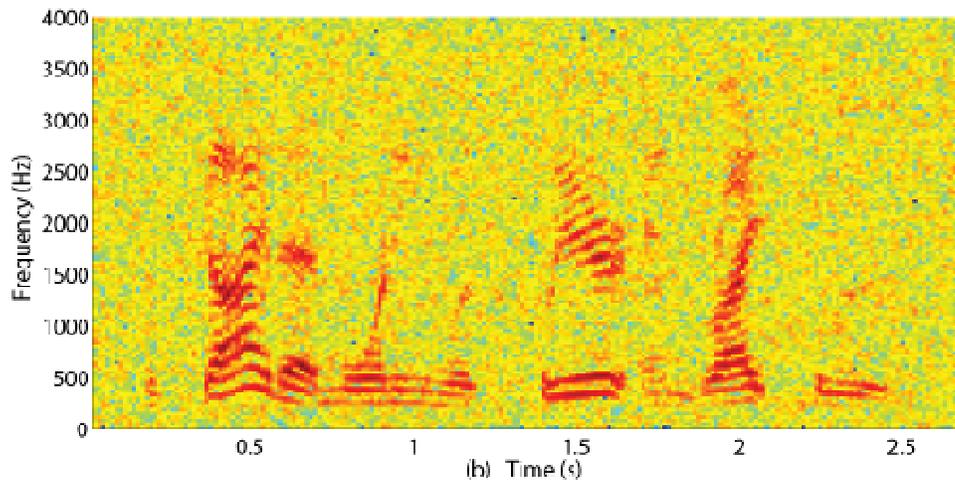
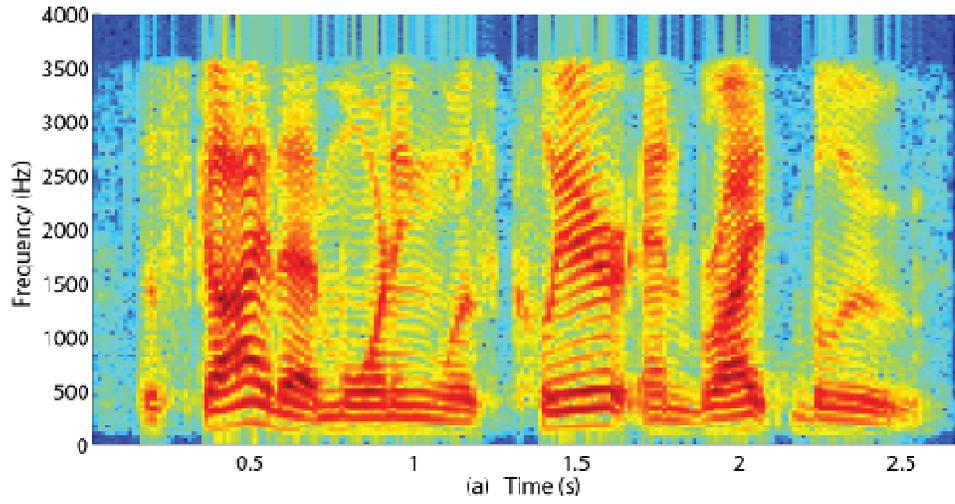
4.4. Subjective Evaluation

In order to evaluate the subjective observation of the enhanced speech obtained by using the proposed method, spectrograms of the clean speech, the noisy speech, and the enhanced speech signals obtained by using all the methods are presented in Fig. 14 and 15 for white noise corrupted speech at an SNR of 5 dB and car noise corrupted speech at an SNR of -5 dB, respectively. It is evident from these figures that the harmonics are preserved and the amount of distortion is greatly reduced in the proposed method no matter the speech is corrupted by white or car noise regardless of its level. Thus, the spectrogram observations with lower distortion also validate our claim of better speech quality as obtained in our objective evaluations in terms of higher SNR improvement in dB, higher PESQ score and lower WSS in comparison to the other methods.

Informal listening tests are also conducted, where the listeners are allowed and arranged to perceptually evaluate the clean speech, noisy speech, and the enhanced speech signals. It is found that the subjective sound quality of the proposed method possess the highest correlation with the objective evaluation in comparison to that of the other methods in case of all the noises considered at different levels of SNR.

5. CONCLUSION

An improved wavelet-based approach to solve the problems of speech enhancement using the Probability distribution of Teager Energy Operated wavelet Packet coefficients has been presented in this paper. We incorporated a statistical model-based technique with teager energy operator of the wavelet packet coefficients to obtain a suitable threshold using symmetric K-L divergence. For solving the equation of pdf's, we choose Gaussian distribution as an acceptable pdf for noisy speech, clean speech and noise TEO coefficients in each sub-band. Unlike the unique threshold based method, the threshold value here is adapted based on the speech and silence segments. Then, by employing hard thresholding function the WP coefficients of the noisy speech are thresholded in order to obtain a cleaner speech. Simulation results show that the proposed method yields consistently better results in the sense of higher output SNR in dB, higher output PESQ, and lower WSS values than those of the existing thresholding based methods, hence results in a better enhanced speech than the existing thresholding methods.



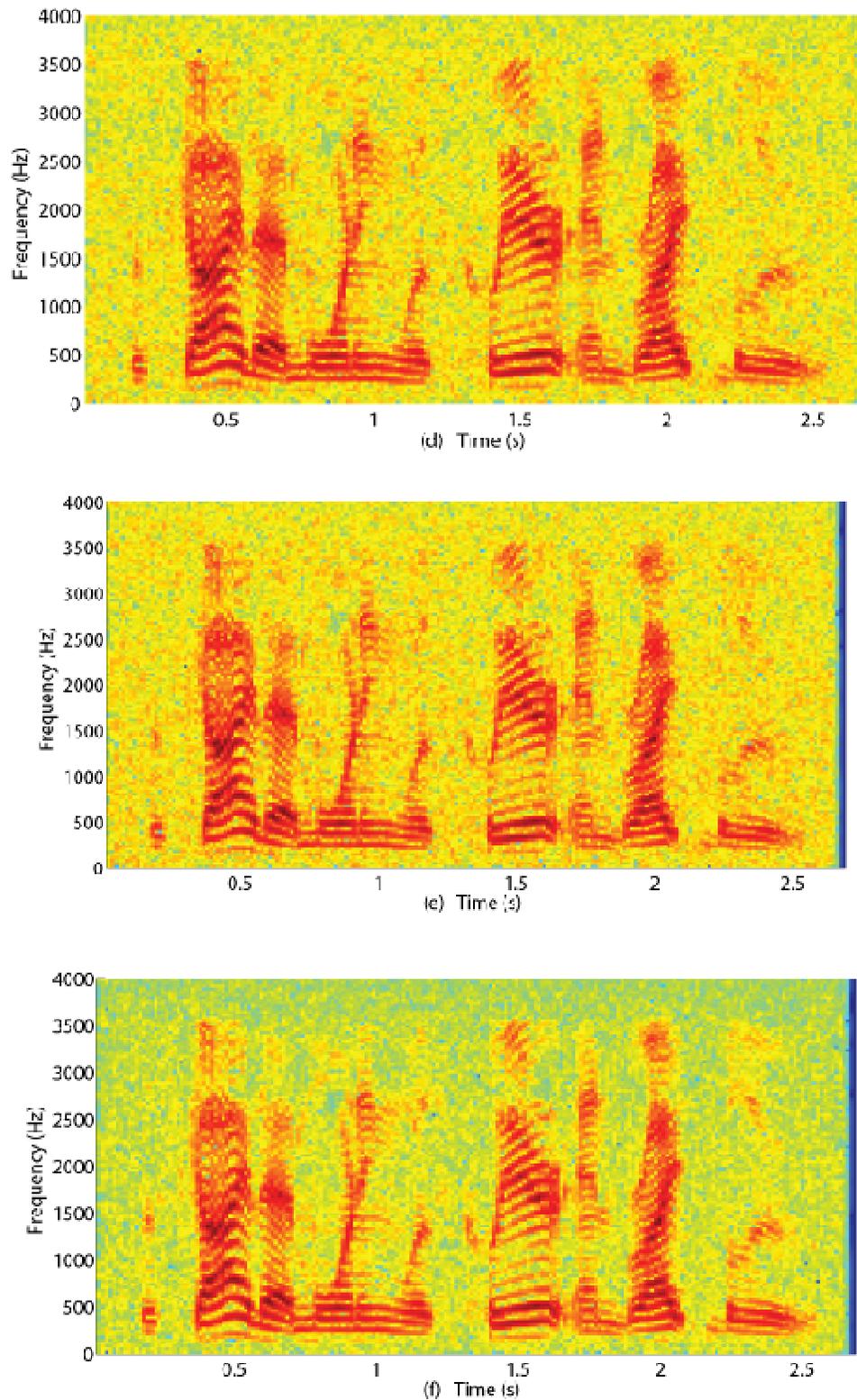
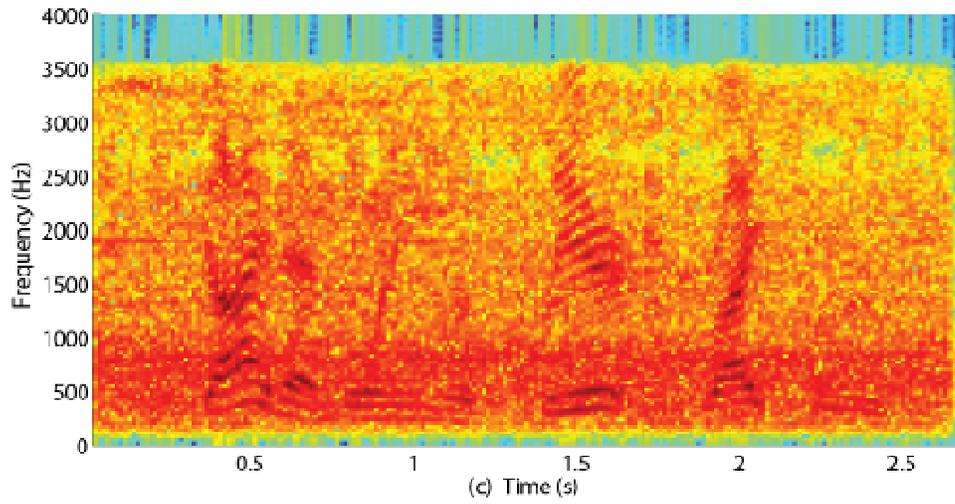
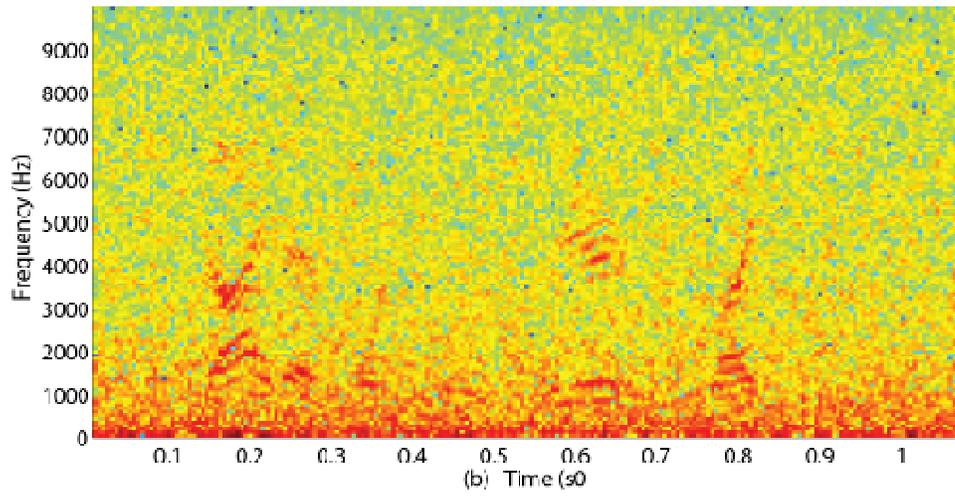
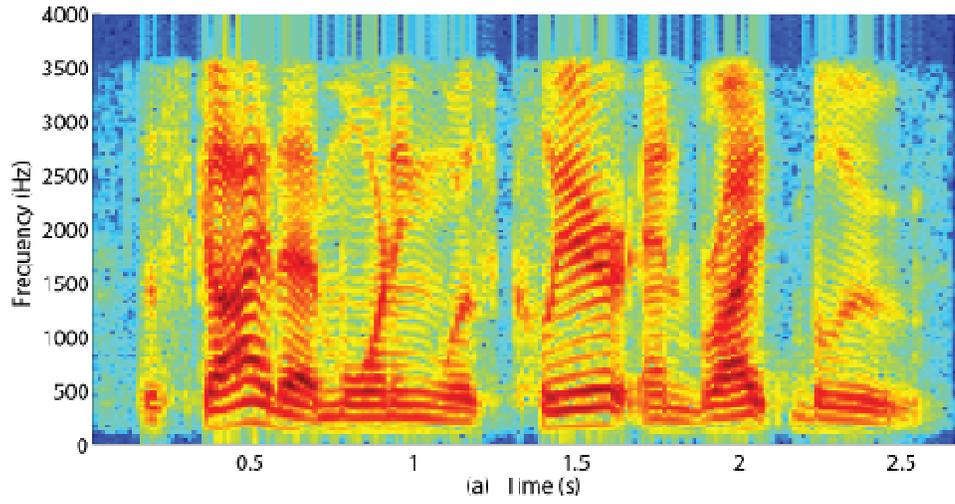


FIGURE 14. Spectrogram of sp10.wav utterance by a male speaker from the NOIZEUS database: (a) Clean speech, (b) Noisy speech (white noise from NOISEX92 database of SNR 5 dB), (c), (d), (e), (f)- enhanced speech signals obtained by using the Universal, TAT, WTHSKL, and the proposed methods, respectively.



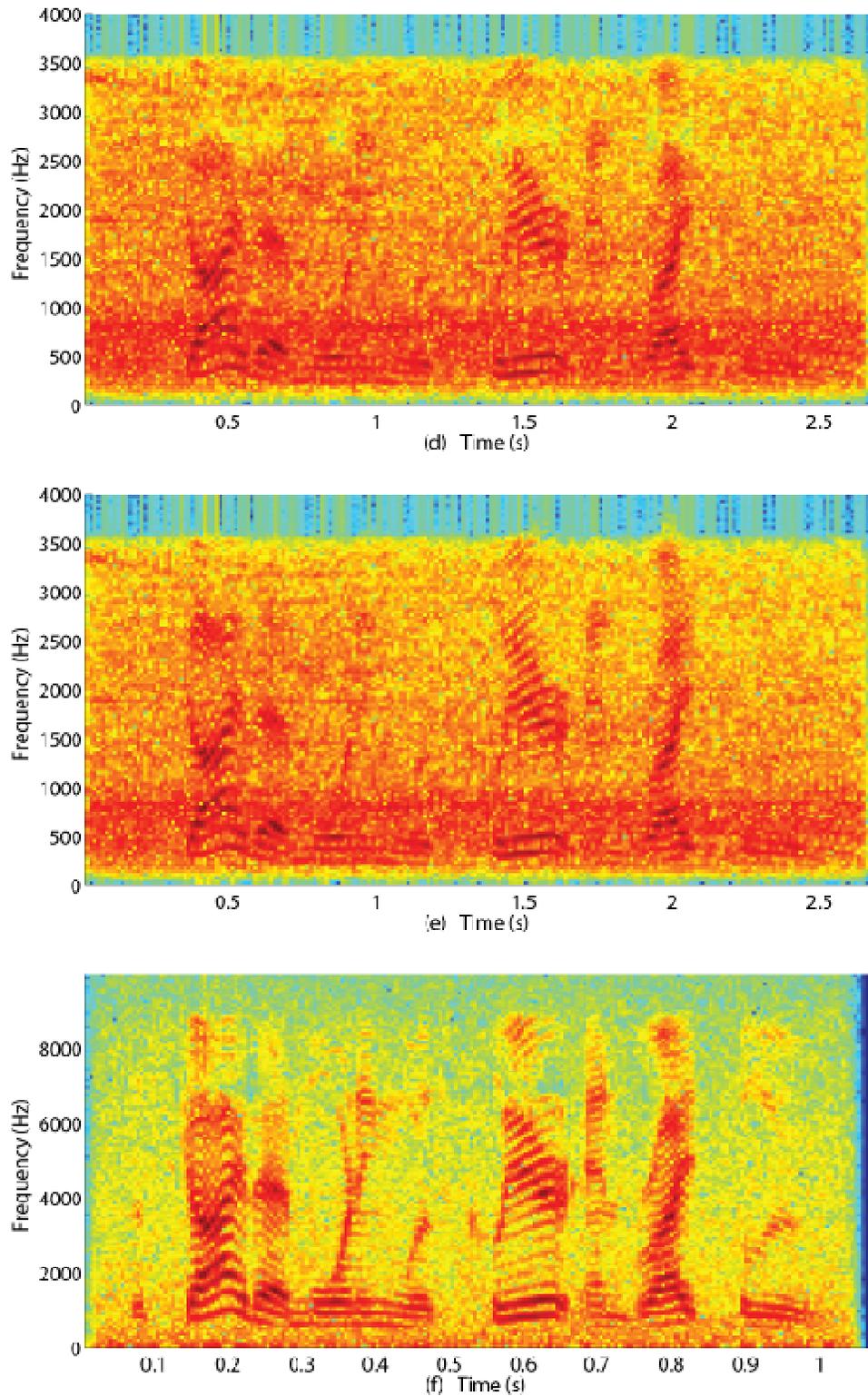


FIGURE 15. Spectrogram of sp01.wav utterance by a male speaker from the NOIZEUS database: (a) Clean speech, (b) Noisy speech (car noise from NOIZEUS database of SNR -5 dB), (c), (d), (e), (f)-enhanced speech signals obtained by using the Universal, TAT, WTHSKL, and the proposed methods, respectively.

REFERENCE

- [1] O'Shaughnessy, D., "Speech Communications: Human and Machine", 2nd Edition, *Wiley-IEEE Press*, 1999.
- [2] Loizou, P. C., "Speech Enhancement: Theory and Practice," *Boca Raton: CRC Press*, 2007.
- [3] Deller, J. Jr., Hansen, J. and Proakis, J., "Discrete-Time Processing of Speech Signals," *NY: IEEE Press*, 2000.
- [4] Lim, J., and Oppenheim, A., "Enhancement and bandwidth compression of noisy speech" *Proc. IEEE*, vol. 67, No. 12, pp. 221-239, Dec. 1979.
- [5] Virag, N., "Single channel speech enhancement based on masking properties of the human auditory system," *IEEE Transactions on Speech and Audio Processing*, volume 7, no. 2, pp. 126-137, Mar 1999.
- [6] Ephraim, Y. and Van Trees, H. L., "A signal subspace approach for speech enhancement," *IEEE Trans. Speech Audio Processing*, volume 3, pp. 251-266, 1995.
- [7] Mittal, U., and Phamdo, N., "Signal/noise KLT based approach for enhancing speech degraded by colored noise," *IEEE Trans. Speech Audio Processing*, volume 8, pp. 159-167, March 2000.
- [8] Y. Hu and P. C. Loizou, "A generalized subspace approach for enhancing speech corrupted by colored noise" *IEEE Trans. Speech, Audio Process*, volume. 11, pp. 334– 341, Jul. 2003.
- [9] Jabloun, F. and Champagne, B., "Incorporating the human hearing properties in the signal subspace approach for speech enhancement," *IEEE Transactions on Speech and Audio Processing*, volume. 11, pp. 700-708, 2003.
- [10] You, C. H., Koh, S. N., and Rahardja, S.; , "An invertible frequency eigen domain transformation for masking-based subspace speech enhancement," *IEEE Signal Processing Letters*, volume.12, no.6, pp. 461- 464, June 2005.
- [11] Chang, J.-H., "Warped discrete cosine transform-based noisy speech enhancement," *IEEE Trans. Circuits and Systems II: Express Briefs*, volume 52, pp. 535 – 539, 2005.
- [12] Gustafsson, H., Nordholm, S.E., and Claesson, I., "Spectral subtraction using reduced delay convolution and adaptive averaging" *Speech and Audio Processing, IEEE Transactions on*, vol.9, no.8, pp.799-807, Nov 2001.
- [14] Kamath, S., and Loizou, P.; "A multi-band spectral subtraction method for enhancing speech corrupted by colored noise," *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 4, pp. IV-4164, 13-17, May 2002.
- [15] Yamashita, K. and Shimamura, T., "Nonstationary noise estimation using low-frequency regions for spectral subtractio," *Signal Processing Letters*, volume. 12, pp. 465-468, 2005.
- [16] Boll, S., "Suppression of acoustic noise in speech using spectral subtraction," *IEEE Trans. Acoust., Speech, Signal Process.*, volume.27, pp. 113-120, Apr. 1979.
- [17] Chen, B. and Loizou, P. C., "A Laplacian-based (MMSE) estimator for speech enhancemen," *Speech Communication*, volume. 49, pp. 134-143, 2007.
- [18] P. C. Loizou, "Speech enhancement based on perceptually motivated bayesian estimators of the magnitude spectrum," *IEEE Trans. Speech, Audio Process*, volume. 13, pp. 857–869, Sep. 2005.
- [19] Ephraim, Y. and Malah, D., "Speech enhancement using a minimum mean-square error log-spectral amplitude estimator," *IEEE Transactions on Acoustics, Speech and Signal Processing*, volume.33, no.2, pp. 443- 445, Apr 1985.

- [20] Sameti, H., and Sheikhzadeh, H., "L. Deng, R. Brennan, HMM-based strategies for enhancement of speech signals embedded in nonstationary noise," *IEEE Trans. Speech and Audio Processing*, volume 6, pp. 445–455, 1998.
- [21] Hansen, J. H. L., Radhakrishnan, V., and Arehart, K.H., "Speech Enhancement Based on Generalized Minimum Mean Square Error Estimators and Masking Properties of the Auditory System," *IEEE Transactions on Audio, Speech, and Language Processing*, volume. 14, no.6, pp. 2049-2063, Nov. 2006.
- [22] Almajai, I., and Milner, B., "Visually Derived Wiener Filters for Speech Enhancement," *IEEE Transactions on Audio, Speech, and Language Processing*, volume. 19, no.6, pp. 1642-1651, Aug. 2011.
- [23] Ben Jebara, S., "A Perceptual Approach to Reduce Musical Noise Phenomenon with Wiener Denoising Technique," *IEEE International Conference on Acoustics, Speech and Signal Processing, (ICASSP)*, volume 3, pp. 14-19, May 2006.
- [24] Martin, R., "Speech Enhancement Based on Minimum Mean-Square Error Estimation and Super gaussian Priors," *IEEE Transactions on Speech and Audio Processing*, volume. 13, no.5, pp. 845- 856, Sept. 2005.
- [25] Papoulis, A. and Pillai, S. U., "Probability, Random Variables and Stochastic Processes," 4th Edition, *McGraw-Hill*, 2002.
- [26] Chang, S., Kwon, Y., Yang, S. I., and Kim, I. J., "Speech enhancement for non-stationary noise environment by adaptive wavelet packet," *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, volume. 1, pp. 1-561 -1-564, 2002.
- [27] Yi, H. and Loizou, P.C., "Speech enhancement based on wavelet thresholding the multitaper spectrum," *IEEE Signal Processing Letters*, volume. 12, pp. 59-67, 2004.
- [28] Tabibian, S. and Akbari, A. and Nasersharif, B., "A new wavelet thresholding method for speech enhancement based on symmetric Kullback-Leibler divergence," *14th International CSI Computer Conference (CSICC)*, pp. 495-500, 2009.
- [29] Bahoura, M. and Rouat, J., "Wavelet speech enhancement based on the Teager energy operator," *IEEE Signal Processing Letters*, volume. 8, pp. 10-12, 2001.
- [30] Donoho, D.L., "De-noising by soft-thresholding," *IEEE Transactions on Information Theory*, volume. 41, pp. 613-627, 1995.
- [31] Ghanbari, "A new approach for speech enhancement based on the adaptive thresholding of the wavelet packets," *Speech Communication*, volume 48, pp. 927 – 940, 2006.
- [32] Sheikhzadeh, H., and Abutaleb, H. R., "An improved wavelet-based speech enhancement system," *EUROSPEECH*, pp. 1855–1858, 2001.
- [33] Shao, Y., Chang, C. H., "A Generalized Time–Frequency Subtraction Method for Robust Speech Enhancement Based on Wavelet Filter Banks Modeling of Human Auditory System," *IEEE Transactions on Systems, Man, and Cybernetics*, volume. 37, no.4, pp.877-889, Aug. 2007.
- [34] Johnson, M. T., Yuan, X., and Ren, Y., "Speech signal enhancement through adaptive wavelet thresholding", *Speech Communication*, 2007.
- [35] Kaiser, J., "Some useful properties of teager's energy operators," *IEEE Int. Conf. Acoustics, Speech, and Signal Processing, (ICASSP)*, volume 3, pp. 149 –152.
- [36] Kaiser, J., "Some useful properties of teager's energy operators," *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 3, pp.149 –152.

- [37] Lallouani, A., Gabrea, M., and Gargour, C., "Wavelet based speech enhancement using two different threshold-based denoising algorithms," *Canadian Conference on Electrical and Computer Engineering*, volume 1, pp. 315 – 318, 2004.
- [38] Malayeri, A., "Noise speech wavelet analyzing in special time ranges," *Int. Conf. Advanced Communication Technology (ICACT)*, volume 1, pp. 525 –528.
- [39] Ephraim, Y.; Malah, D.;; "Speech enhancement using a minimum-mean square error short-time spectral amplitude estimator," *IEEE Transactions on Acoustics, Speech and Signal Processing*, volume. 32, no.6, pp. 1109- 1121, Dec 1984.
- [40] Erkelens, J.S.; Hendriks, R.C.; Heusdens, R.; Jensen, J.; , "Minimum Mean-Square Error Estimation of Discrete Fourier Coefficients With Generalized Gamma Priors," *IEEE Transactions on Audio, Speech, and Language Processing*, volume. 15, no.6, pp.1741-1752, Aug. 2007.
- [41] Benesty, J., Sondhi, M., and Huang, Y., "Handbook of Speech Processing," *Springer*, 2008.
- [42] Gazor, S.; Wei Zhang; , "Speech enhancement employing Laplacian-Gaussian mixture," *IEEE Transactions on Speech and Audio Processing*, volume. 13, no.5, pp. 896- 904, Sept. 2005.
- [43] Plourde, E., "Bayesian short-time spectral amplitude estimators for single channel speech enhancement," *Ph.D. thesis*, Montreal, Que., Canada, Canada, 2009.
- [44] Makovoz, D., "Noise Variance Estimation In Signal Processing," *IEEE International Symposium on Signal Processing and Information Technology*, pp.364-369, Aug. 2006.
- [45] Babaie-Zadeh, M., and Jutten, C., "A general approach for mutual information minimization and its application to blind source separation," *IEEE Signal Processing Letters*, volume 85, pp. 975– 995, 2005.
- [46] De Souza, P., "A statistical approach to the design of an adaptive self normalizing silence detector, Acoustics," *IEEE Transactions on Speech and Signal Processing*, volume. 31, pp. 678 – 684, 1983.
- [47] Busso, C.; Sungbok Lee; Narayanan, S.; , "Analysis of Emotionally Salient Aspects of Fundamental Frequency for Emotion Detection," *IEEE Transactions on Audio, Speech, and Language Processing*, vol.17, no.4, pp.582-596, May 2009.
- [48] Hu, Y., and Loizou, P., "Evaluation of objective quality measures for speech enhancement," *IEEE Transactions on Audio, Speech, and Language Processing*, volume. 16, pp. 229 –238, 2008.

DSP Based Speech Operated Home Appliances Using Zero Crossing Features

Deepali Y Loni

Assistant Professor

Department of Electronics

D.K.T.Es' Textile & Engineering Institute,

Ichalkaranji. Post code:416115, India.

deepaliloni@rediffmail.com

Abstract

The main idea of this paper is to build a simple speech recognition system using Digital Signal Processor (DSP) that controls home appliances (i.e. turning on/off) by processing the spoken word. The method used is simple, involving a plain count of the frequency of zero crossings. Two features of zero-crossing are used namely: maxima & running sum that increases the accuracy of recognition. The DSP calculates the zero crossings of the spoken words and accordingly generates different analog signals at its output. These analog signals are further processed so as to operate the appliances. The words chosen for recognition are 'ONE', 'TWO' and 'THREE'.

The paper includes two approaches for implementation of speech recognition into DSP, using Matlab Simulink approach and secondly using Code Composer Studio (CCS). Moreover the first approach performs offline processing and the other performs real time processing of words. The results at the end describe the efficiency of the system.

Keywords: Speech Recognition, Matlab, Simulink, Zero Crossing.

1. INTRODUCTION

Speech recognition is a vast topic of interest and is looked upon as a complex problem. In a practical sense, speech recognition solves problems, improves productivity, and changes the way we run our lives. Reliable speech recognition is a hard problem, requiring a combination of many techniques; however modern methods have been able to achieve an impressive degree of accuracy [1]. Real-time digital signal processing made considerable advances after the introduction of specialized DSP processors. Suitable DSP Starter Kits, with specific DSP processor and related software tools such as assemblers, simulators and debuggers are available to make system design and application development easier. Digital Signal Processor TMS320C6713 enables to design a system with very high computational power and large memory space with minimal count of components what saves printed circuit board space and simplifies design [2,3].

Talking to appliances in a home has been a science fiction staple for decades. This paper proposes solution for providing speech interface to electronic devices in a home. The TMS320C6713 digital signal processor (DSP) and Microcontroller are used in this work. The DSP is used for implementing speaker dependent speech recognition system for capturing vocal commands for operating the appliances and the Microcontroller serves as the main interface between DSP and the control circuit handling the appliances.

2. THE PROPOSED SYSTEM MODEL

The system model of **Figure 1** consists of two parts. First part informs about interfacing and processing of the spoken word using floating point DSP Starter kit (DSK) TMS320C6713. DSK is used for the research module because it provides an efficient and stable DSP development

environment and it is a robust, low-cost and easily available DSK in both universities and industry [4].

The second part consists of control circuitry that consists of hardware that operates the appropriate appliance through the command recognized by DSK.

The system model consists of:

2.1 The TMS320C6713 Processor for Speech Recognition

The system model uses Texas Instrument TMS320C6713 DSP Processor to perform the task of speech recognition. The Processor first filters the noise from input speech. It then processes and identifies the spoken word. For the identified word, the Processor generates an analog sinusoidal signal of certain frequency. For example, for the audio command ‘ONE’, the Processor generates a sinusoidal signal of frequency 400Hz. Thus sine waves of different frequencies are generated at output of Processor for each spoken word ‘ONE’, ‘TWO’ and ‘THREE’ as shown in the **Table 1**.

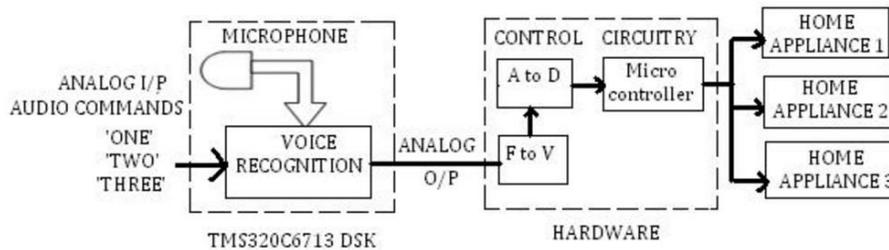


FIGURE 1: The proposed system model

Input audio command	Frequency of the analog signal generated at the output of processor
ONE	400Hz
TWO	800Hz
THREE	2000Hz

TABLE 1: Output of Processor

2.2 The Hardware Control Circuitry

The control circuitry consists of Frequency to Voltage (F to V) converter, Analog to Digital (A to D) converter, AT89C52 Microcontroller, relay drivers and relays.

The analog signals at the output of Processor are passed to F to V converter. The audio commands ‘ONE’, ‘TWO’ and ‘THREE’ are mapped into three different voltage levels by F to V. These voltage levels are then digitalized by ADC0804. AT89C52 Microcontroller operates the appropriate relays scanning its digital input. These relays then in turn operate the corresponding appliances like; lamps, fan, dishwashers, clothes washers, dryers, microwaves, refrigerators, freezers, etc., switching them either ON or OFF. For example if the spoken word is ‘ONE’, then home appliance one will toggle (i.e. if it is ON then it will turn OFF and vice versa), if say the spoken word is ‘TWO’ then home appliance two will toggle.

3. PREPARATION OF DATABASE

In the first part of this work, Matlab is used to prepare database for offline processing of the words. These words are first recorded by means of microphone using wavrecord command in Matlab. The Matlab code is discussed below:

```
fs = 8000;  
Y=wavrecord (5632, fs, 1, 16);  
wavwrite(y, fs,16,'E:\database\three\three10');
```

It records 5632 samples of an audio signal, sampled at a rate of 8000 Hz using channel number 1 of input channels from the audio device each of 16 bits. The speech is recorded for $5632 / 8000 = 0.704$ seconds which is enough time to say a complete word. The Matlab command wavwrite is used to write the recorded data to .wav file. Ten sample words are stored for each audio command 'ONE', 'TWO' and 'THREE'.

4. USING ZERO-CROSSING FEATURE

The feature used to differentiate between the audio commands; 'ONE', 'TWO' and 'THREE' is zero crossings in the words. Zero-crossing rate is a measure of the number of time in a given time interval that the amplitude of the speech signal passes through a value of zero [1]. It means, a zero-crossing is said to have occurred in a signal when its waveform crosses the time axis or changes its algebraic sign. This feature has been used heavily in both speech recognition and music information retrieval.

5. METHODS ADOPTED FOR IMPLEMENTATION OF SPEECH RECOGNITION

The speech recognition using zero crossing features is implemented in this paper using two approaches.

5.1 Approach I: Implementation of Speech Recognition Using Matlab Simulink.

The Mathwork's Simulink is used to implement speech recognition system. There are many advantages in programming DSP algorithms using Matlab. These include ease of coding, able to use a powerful set of inbuilt functions and seamless link between Matlab and Simulink [5]. The use of Simulink enables the creation of sophisticated algorithms in an intuitive top-level design [6].

Figure 2 shows the complete Simulink model, build for the recognition of the three audio commands: 'ONE', 'TWO' and 'THREE'. The model recognizes the sample words from database prepared.

The block From Wave File reads data from stored sample words at a rate of 256 samples per frame. As the audio is recorded at 8 KHz, each sample word gets broken into 22 frames.

The Digital Filter Design block implements a Low Pass FIR Filter to filter the noise from the input signal. An offset of value 0.05 is added to the input signal before computing the zero crossing, in order to avoid the zero crossings due to signal noise. The zero crossing block outputs the number of times the signal crosses zero at its output port. The model further computes two features from the zero crossing block. They are: maximum value of zero crossings count from all 22 frames and the sum of zero crossings count in all 22 frames of the spoken word, using the Maximum and Running Sum blocksets. These computed features differ for every word and hence help to exactly identify the spoken word.

The Simulink model to understand the output of Maximum and Running Sum blockset for the word 'ONE' is shown in **Figure 3** and **4**. The Maximum block finds out frame with maximum zero crossings. Hence **Figure 3** displays the highest zero crossings from the 22 frames of the word. The running sum block computes the zero crossings in entire word by summing the zero

crossings in all 22 frames. Hence **Figure 4** displays the sum of the zero crossings in the word 'ONE'.

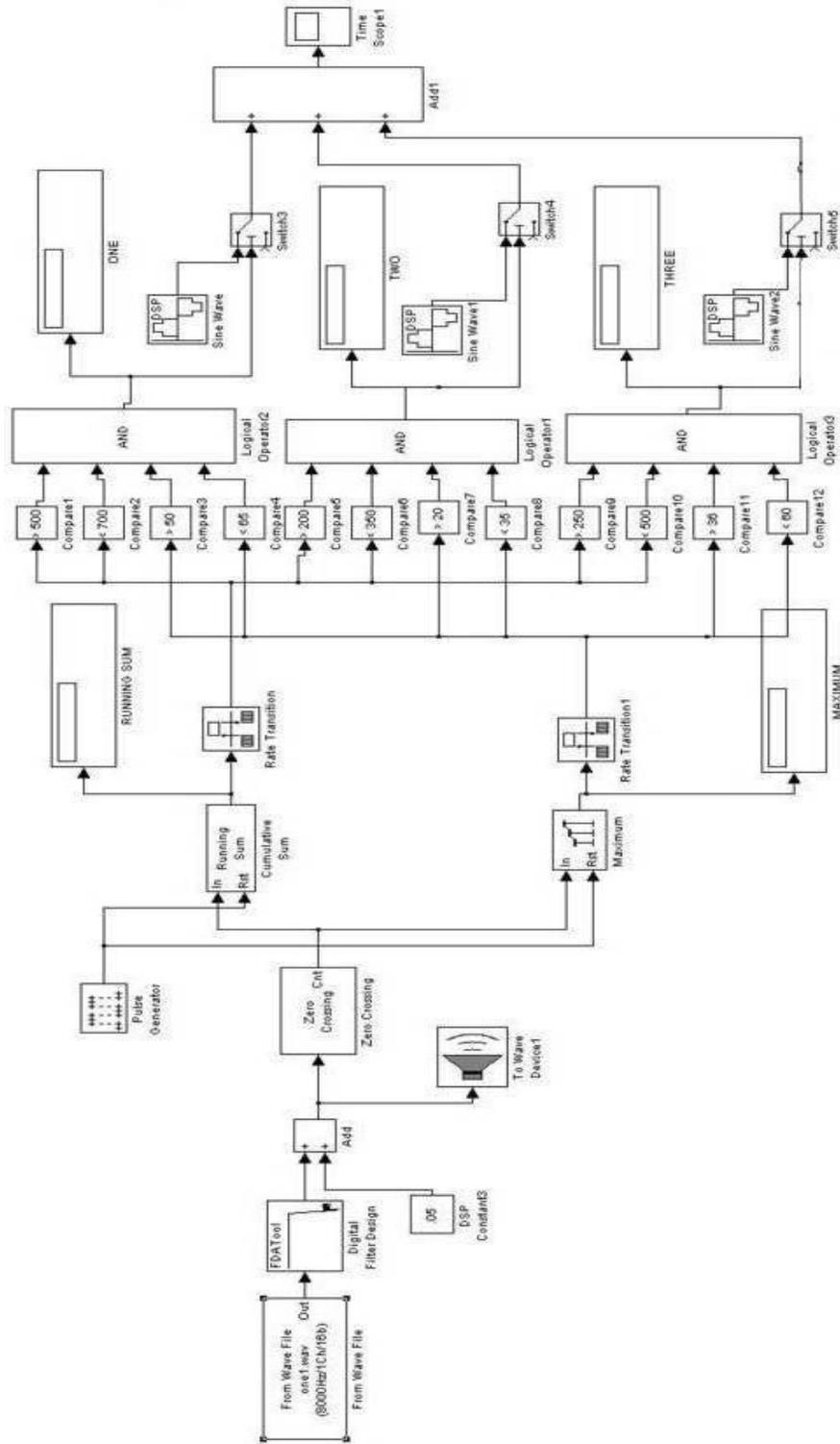


FIGURE 2: The Simulink Model recognizing commands 'ONE', 'TWO', & 'THREE'.

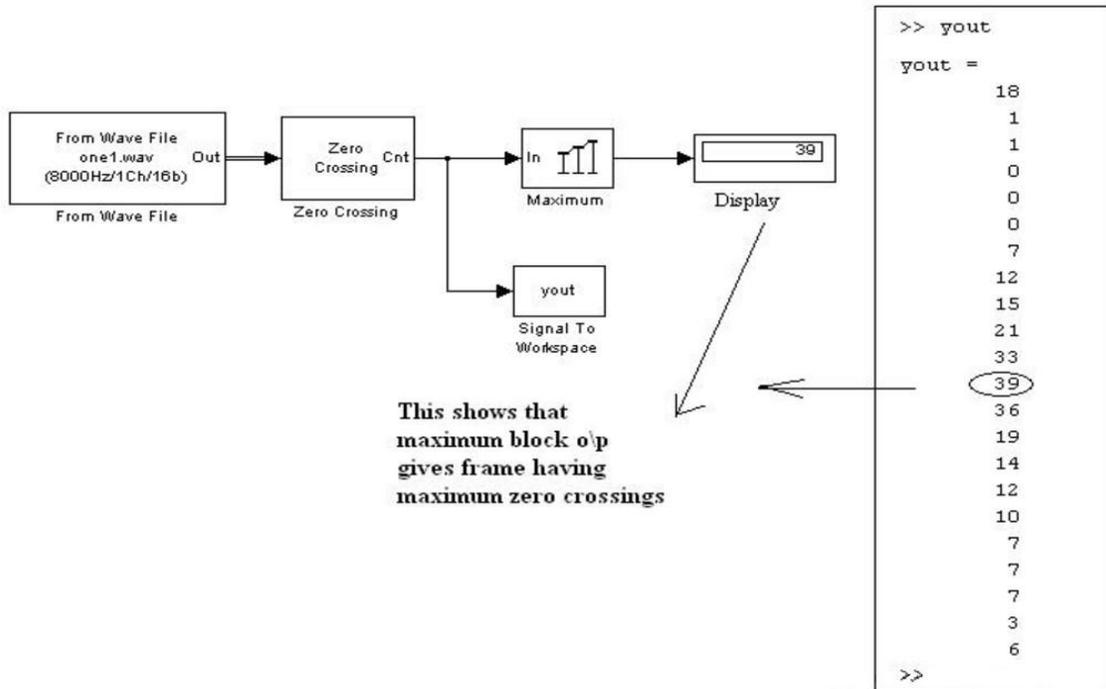


FIGURE 3: Maximum Block output of the Simulink Model

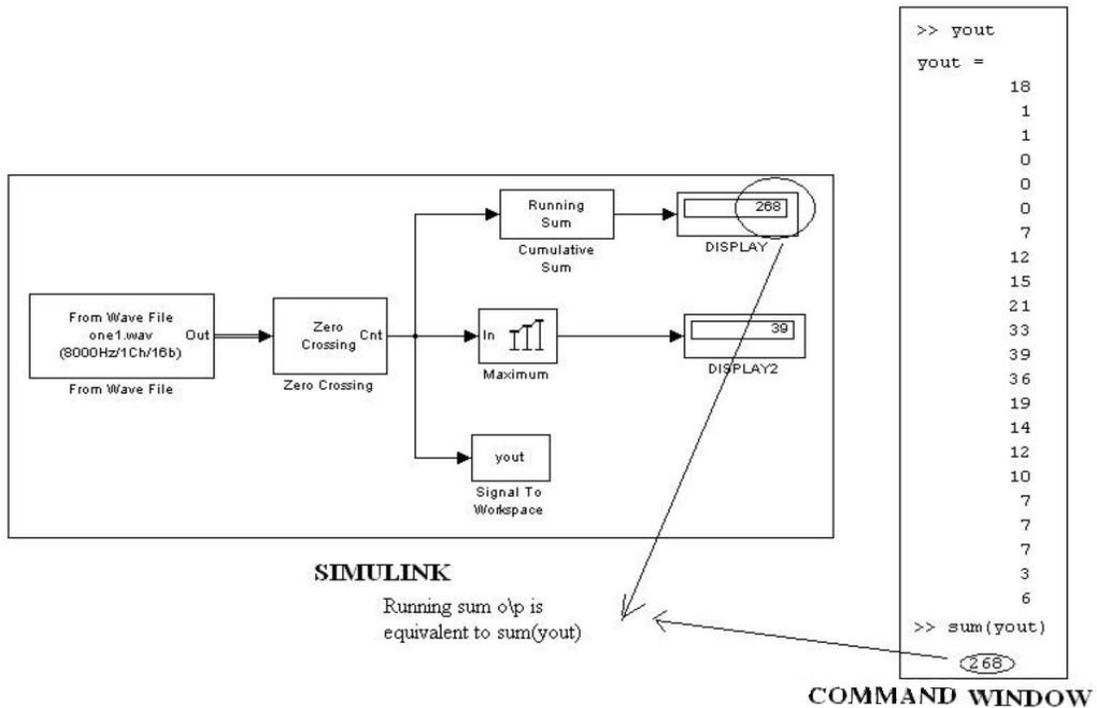


FIGURE 4: Running Sum Block output of the Simulink Model

Audio command	Simulink Blocksets			
	Maximum		Running sum	
	Min	Max	Min	Max
ONE	50	65	500	800
TWO	20	35	200	350
THREE	36	60	250	500

TABLE 2: Minimum & Maximum values of Maximum and Running Sum Blocks

These features are calculated for all the recorded commands in the database and it is observed that, each audio command lie within certain range of minima and maxima as shown in **Table 2**.

Table 2 shows that, the total zero crossings for audio command 'ONE' as computed by running sum block lies between 500-800 and maximum zero crossing observed in a frame lie between 50- 65. While for audio command 'TWO' and 'THREE' the cumulative zero crossings is between 200-350 and 250-500, and maximum zero crossing observed in a frame lie between 20-35 and 36-60 respectively. In addition **Table 2** also reveals that the maximum zero crossing value for the audio command 'ONE' & 'THREE' have a common range, but their running sum ranges differ. Accordingly, the Simulink model of **Figure 2** is designed that check all states and accurately identify the command 'ONE'.

5.2 Approach II: Implementation of Speech Recognition Using Code Composer Studio.

The Code Composer Studio (CCS) software comes with the DSP Starter kit (DSK) and is used to download programs into DSK [8, 9]. CCS includes tools for code generation, such as a C compiler, an assembler, and a linker. The C compiler compiles a C source program with extension .c to produce an assembly source file with extension .asm. The assembler assembles an .asm source file to produce a machine language object file with extension .obj. The linker combines object executable file that can be loaded and run directly on C6713 DSP.

CCS supports C-language coding. The C program for implementation of speech recognition algorithm in CCS is explained in the form of flowchart in **Figure 5**.

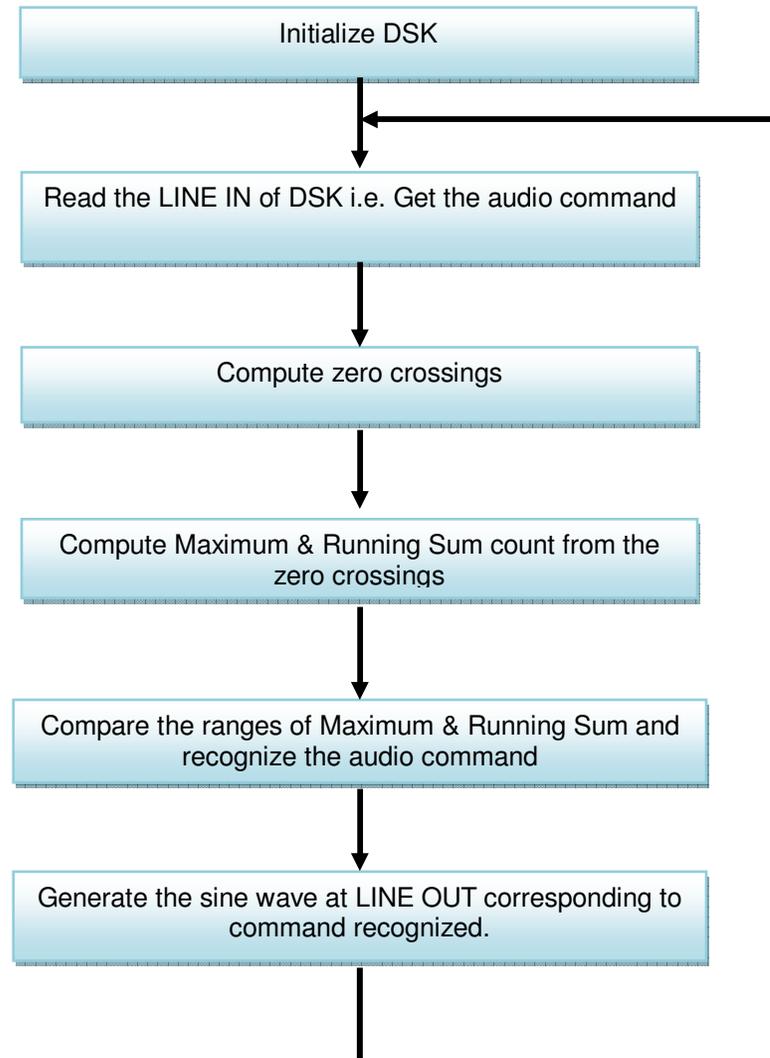


FIGURE 5: Flowchart of C program for implementation of speech recognition algorithm in CCS

6. EMBEDDING SPEECH RECOGNITION ALGORITHM INTO TMS320C6713 DSK

For standalone working of DSK TMS320C6713, the speech recognition algorithm must be embedded in its flash memory. As the paper proposes two approaches of speech recognition implementation, the methods of interfacing these algorithms to DSK are discussed below:

6.1 Embedding Matlab Simulink Speech Recognition Algorithm Into DSK.

With the advent of Matlab's Real-Time Workshop (RTW) it is possible to compile, load, and execute graphically designed Simulink models on an actual DSP platform.

Simulink uses graphical block diagrams to create models for real-time implementation of applications and then use Real-Time Workshop to generate C code targeted to the TI DSP board by mean Code Composer Studio (CCS IDE) [7]. For downloading this Simulink model of **Figure. 2** into DSK, it is essential to include C6713 DSK Board Support Library Blocks in the model so as to establish a communication with the codec of target C6713 DSK [10-13].

Steps involved in building and executing the Simulink model on C6713 DSK:

- 1) Configure Simulink parameters and configure Real-Time Workshop.
- 2) Run the CCS software in background.
- 3) Open the Simulink model of speech recognition and press CNTRL + B to build an equivalent 'C' language code in CCS. Simulink starts communicating with CCS and generates the C code that can be run by the C6713 DSP on DSK.
- 4) To stop model execution, click the Reset DSK block or use the Halt option in CCS.

6.2 Embedding the CCS Speech Recognition Algorithm Into DSK.

The approach using Simulink performs processing on stored audio words. For real time processing of audio words, this approach includes Real Time Data Exchange (RTDX). RTDX provides real-time, continuous visibility into the way target applications operates in the real world. RTDX allows transfer of data between the host computer and DSP devices without stopping target application.

Figure 6 illustrates how realtime audio signal can be passed on to DSK using RTDX. The audio command is recorded with PC-based audio input device i.e. microphone using wavrecord command of Matlab (running on the host PC) in real time. The array of recorded data is then sent to C6713 processor through RTDX. The C program for implementation of speech recognition algorithm (**Figure 5**) is transformed to executable file and is loaded on the C6713 DSP using CCS tools. The C source program (running on the DSK) then calculates the frame wise zero crossings and running sum of zero crossings. Depending upon its value, the DSK recognizes the spoken word 'ONE', 'TWO' or 'THREE' and sends the respective sine wave of particular frequency on LINE OUT port of DSK.

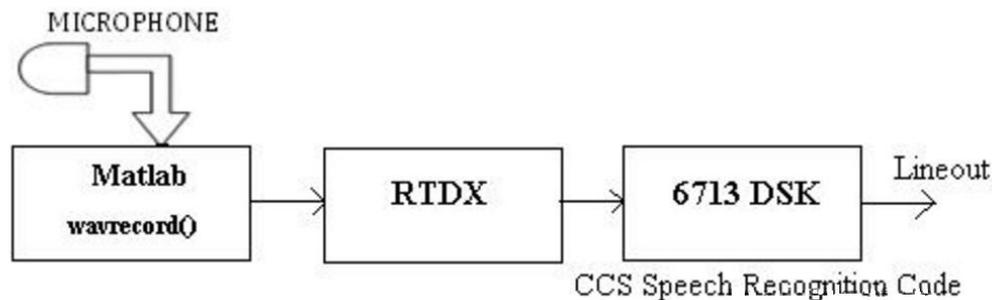


FIGURE 6: Interface between real time audio command and DSK using RTDX.

7. HARDWARE INTERFACE TO C6713 DSK

Once the DSK recognizes the spoken words, it generates a sinusoidal signal of a fixed frequency assigned to each word from the LINE OUT port. The hardware circuit interfaced to DSK operates the home appliances accordingly. The Block Diagram of the hardware circuit is shown in **Figure 7**.

As DSK outputs sinusoidal signals of fixed frequency, the first block in the hardware circuit is an F to V converter. As per the digital signals received at its input port, Microcontroller operates the relay. As the Microcontroller cannot provide sufficient current to drive relays, a relay driver is essential interface between the two. The relays in turn operate the home appliances.

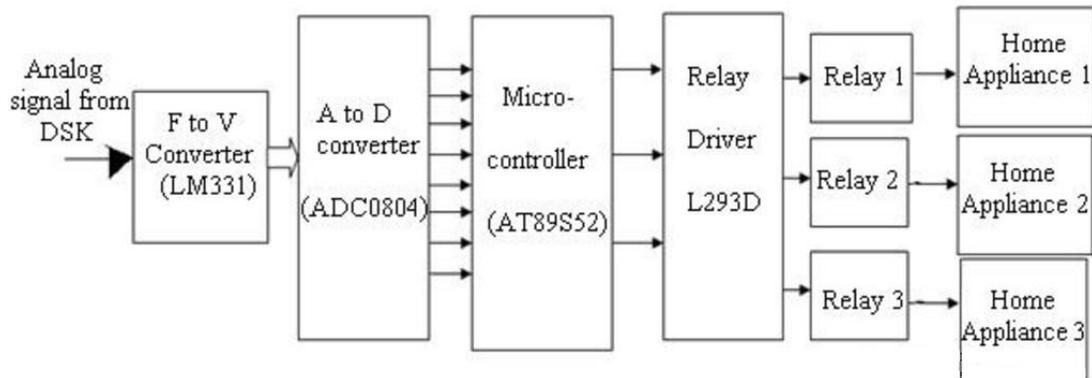


FIGURE 7: Block diagram of Hardware circuit interfaced to DSK

8. RESULTS

8.1 Simulink Result

The Simulink result for the audio command 'THREE' is shown in **Figure 8**. In **Figure 8**, the model processes the audio command 'THREE', generating the Running Sum value of 450 and Maximum value of 40. These values exactly lie in the ranges mentioned in **Table 2** that rightly identifies the audio command 'THREE'.

8.2 The DSK Result

The DSK generates sinusoidal signals at its LINE OUT port for every word recognized. The sinusoidal signals of 400Hz and 2000Hz obtained at the LINE OUT port for the audio command 'ONE' and 'THREE' are shown in **Figure 9**.

8.3 Results Obtained From Hardware Circuit

Table 3 shows the output of F-V converter and the ADC ranges generated by the hardware in response to the word recognized by the DSK. The audio commands 'ONE', 'TWO' and 'THREE' are mapped into three different voltage levels by F to V converter. Accordingly, the ADC values lie in separate ranges. The Microcontroller operates the correct relay, which in turn switches the corresponding appliance.

9. CONCLUSION

The paper presents two approaches for implementing speech recognition algorithm; using Matlab Simulink approach and secondly using Code Composer Studio (CCS) using DSK TMS320C6713. The algorithms are based on zero crossing feature. The simulations as well as experimental results of the hardware circuit are included. These results indicate that the home appliances can be operated reliably with voice commands. The proposed method also finds promising applications in robot control and helpful for industries where there is immense danger in operating the system manually.

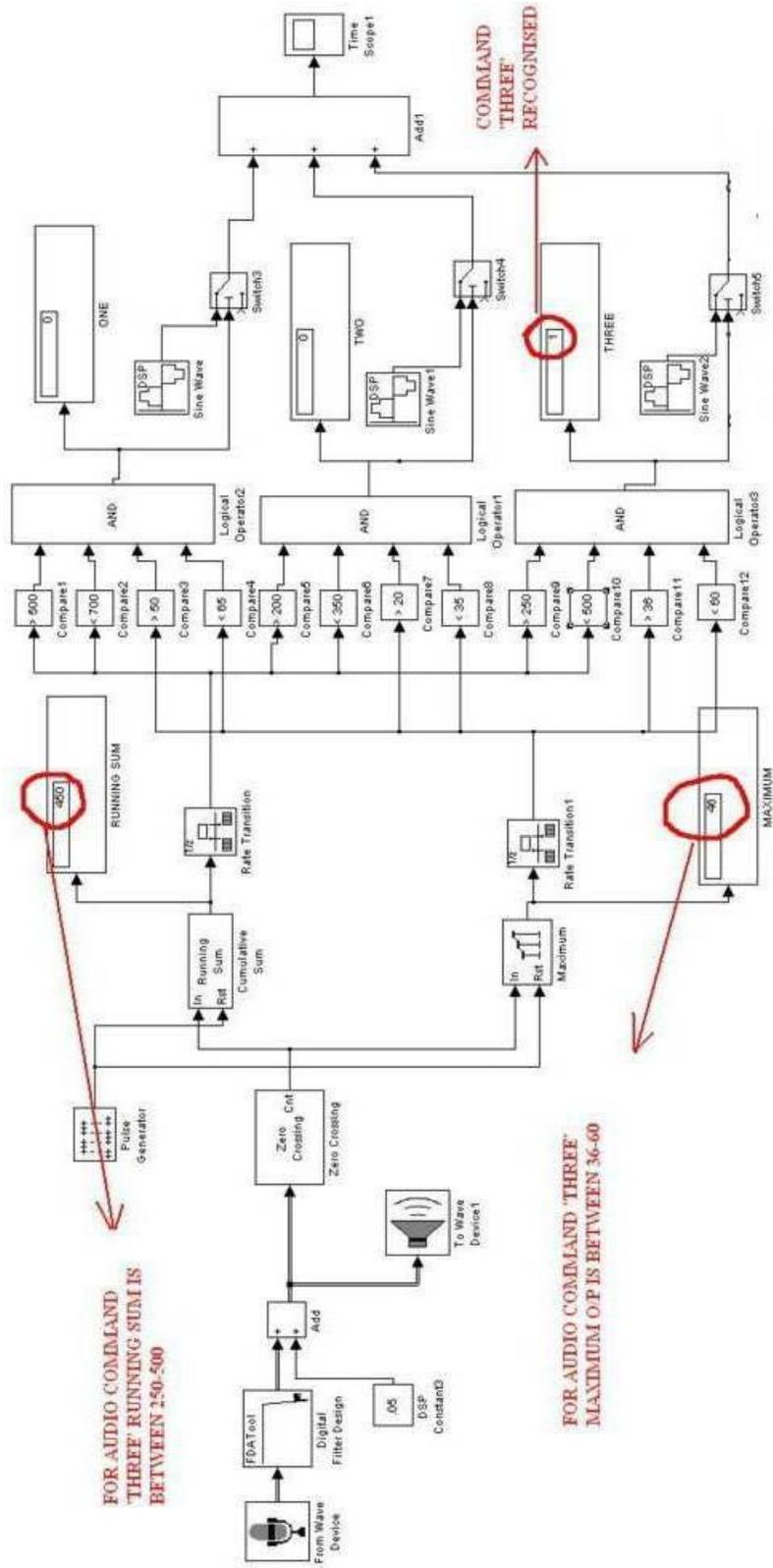
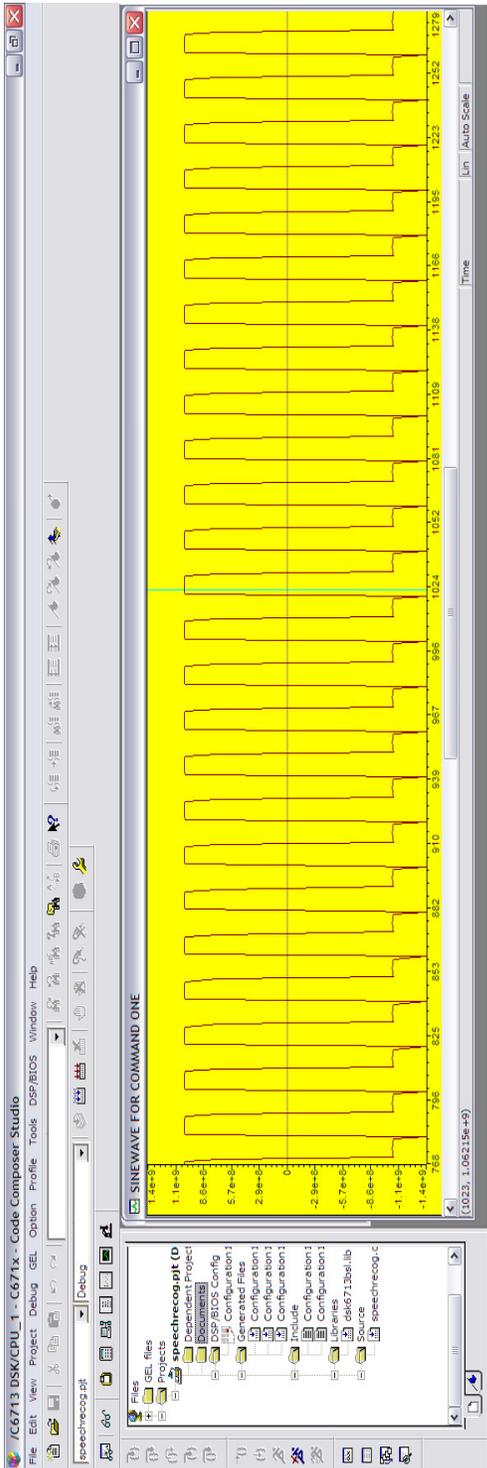
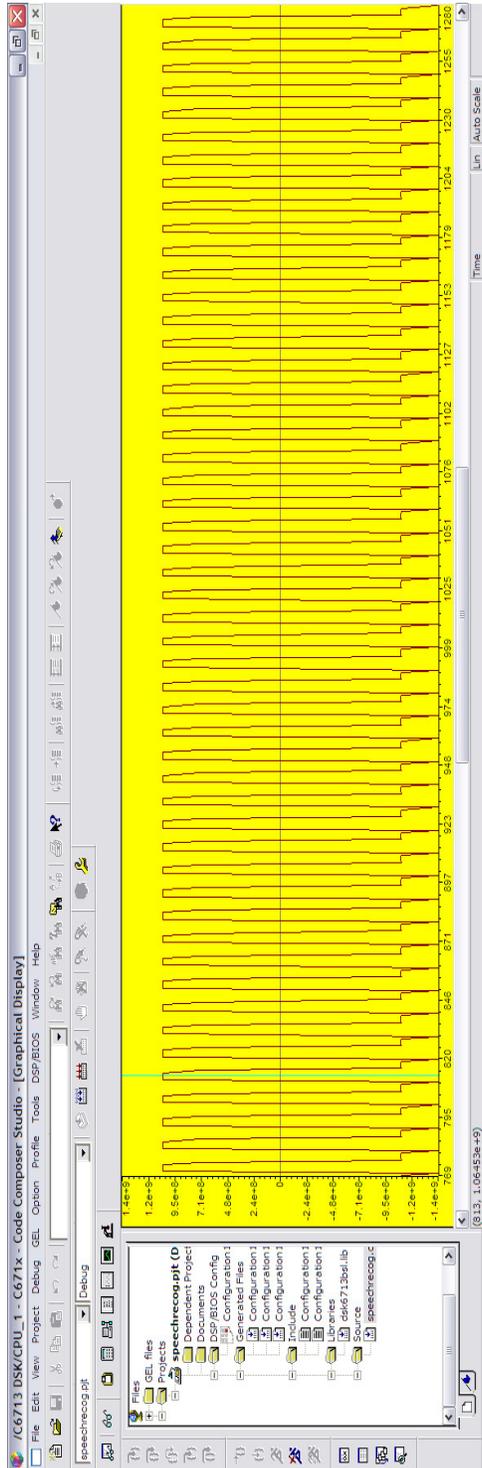


FIGURE 8: Result of Simulink model for the audio word 'THREE'



(a)



(b)

FIGURE 9: Plots of sinusoidal signals generated at the output of DSK. . (a) For audio command 'ONE'. (b) For audio command 'THREE'.

Audio command	Sine Wave Frequency	F to V Converter output	ADC Range
ONE	400Hz	1.10V	0x30-0x3F
TWO	800Hz	2.65V	0x80-0x8F
THREE	2000Hz	4.04V	0xC0-0xDF

TABLE 3: Results obtained from hardware circuit.

10. REFERENCES

- [1] S. K. Hasnain, A. Beg and S. Awan, “*Frequency Analysis of Urdu Spoken Numbers Using MATLAB and Simulink*” Journal of Science & Technology PAF KIET ISSN 1994- 862x, Karachi, Dec. 2007.
- [2] F.Wallam, M. Asif, “*Dynamic Finger Movement Tracking and Voice Commands Based Smart Wheelchair*”, International Journal of Computer and Electrical Engineering, Vol. 3, No. 4, August 2011,pp.497 502.
- [3] H. Chin, J.Kim, I.Kim, Y.Kwon, K.Lee, S. Yang, “Realization of Speech Recognition Using DSP(Digital Signal Processor)”, IEEE International Symposium On Industrial Electronics Proceedings 2001, pp.508-512.
- [4] T. J. Ahmad, H. Ali, M. A. Ajaz, S. A. Khan, “A DSK Based Simplified Speech Processing Module for Cochlear Implant”, IEEE International Conference on Acoustics, Speech and Signal Processing, 2009. <http://dx.doi.org/10.1109/ICASSP.2009.4959597>
- [5] W. S. Gan¹ and S. M. Kuo², “*Transition from Simulink to MATLAB in Real-Time Digital Signal Processing Education*”, Int. Journal Engineering Education Vol. 21, No. 4, pp. 587±595, 2005.
- [6] S. Gannot,V. Avrin. “A Simulink And Texas Instruments C6713 Based Digital Signal Processing Laboratory”, 14th European Signal Processing Conference, Italy, September 4-8, 2006.
- [7] J. Zapata,R. Ruiz, “Rapid Development of Real-Time Applications UsingMATLAB/Simulink on TI C6000-based DSP”, Proceedings of the 5th WSEAS International Conference on Education and Educational Technology, Tenerife, Canary Islands, Spain, December 16-18, 2006,pp.104-110.
- [8] M. T. Qadri, S. A. Ahmed,”Voice Controlled Wheelchair Using DSK TMS320C6711”, IEEE Computer Society, International Conference on Signal Acquisition and Processing 2009, pp. 217-220. <http://dx.doi.org/10.1109/ICSAP.2009.48>
- [9] R. Chassaing, “*Digital Signal Processing and Applications with the C6713 and C6416 DSK*”, New Jersey, A John Wiley & Sons, Inc., Publication, 2005.
- [10] The MathWorks, Signal Processing Toolbox User’s Guide, Version 6, 2003.
- [11] The MathWorks, DSP Blockset User's Guide, Version 5, 2003.
- [12] The MathWorks, Embedded Target for the TI TMS320C6000™ DSP Platform, Version1, 2002.
- [13] Texas Instruments, C6713 DSP Starter Kit. Available at: http://focus.ti.com/docs/toolsw/folders/print/tmds_20006713.html#220

Peak-to-Average Power Ratio Reduction in NC-OFDM based Cognitive Radio

Mohammad Zavid Parvez

*Signal Processing
Blekinge Institute of Technology
Karlskrona, 37179, Sweden*

zavidparvez@hotmail.com

Md. Abdullah Al Baki

*Telecommunication Systems
Blekinge Institute of Technology
Karlskrona, 37179, Sweden*

bakiswd@hotmail.com

Mohammad Hossain

*Telecommunication Systems
Blekinge Institute of Technology
Karlskrona, 37179, Sweden*

reganmh8@gmail.com

Abstract

This paper presents a novel technique for reducing the peak-to-average power ratio (PAPR) in non-contiguous bands spectrum of Orthogonal Frequency Division Multiplexing (OFDM) based Cognitive Radio (CR). The proposed system exposed is to carry the earlier period channel information as well as the spectrum sensing to utilize the radio spectrum to achieve an appropriate PAPR reduction. It is maintaining end-to-end throughput performance by using a set of approaches in the current CR environment. The simulation results for PAPR reduction has shown that higher constellation modulation schemes are better compared to lower constellation modulation schemes.

Keywords: OFDM, NC-OFDM, PAPR, Cognitive Radio, Spectrum Sensing.

1. INTRODUCTION

High data rate wireless communications are demanded by many applications. On average, more bandwidth is required for high data rate transmission in most of the systems. With potential technology and increasing wireless devices, the spectrum is becoming scarcer in each time. In this case, efficient transmission of spectrum by Orthogonal Frequency Division Multiplexing (OFDM) and Cognitive Radio (CR) is an alternative solution.

OFDM is followed by spectrum efficient multicarrier modulation where the spectrum is divided into subcarriers with each subcarrier containing a low rate data stream. OFDM has added a remarkable interest in recent years because of its high spectral efficiency, robustness in the presence of severe multipath channel conditions with simple equalization, Inter-symbol Interference (ISI), multipath fading, etc. However, OFDM has is a major drawback of high Peak-to-Average Power Ratio (PAPR). OFDM contains lot of independent modulated subcarriers that carry to create high peak because the independent phases of the subcarriers often combined constructively.

Nowadays, CR is defined as an intelligent wireless system that continually aware about its surrounding environment during sensing and able to dynamically adjust its radio spectrum parameters. Physical layer (PHY) of CR needs to be adaptable and flexible.

The OFDM along with CR is an attractive candidate for flexibility and adaptability of spectral. This paper proposes a novel non-contiguous OFDM (NC-OFDM) technique, where the system achieves high data rates of non-contiguous subcarriers while concurrently avoids interference.

There are different techniques have been proposed to reduce PAPR [4], [5], [6], [7], [8], [9], [10], [11], [12]. An attractive solution to the PAPR problem has proposed by J. Armstrong, where clipping and filtering techniques were used for PAPR reduction in the transmitter [1]. Another paper describes the End-to-End QoS Maintenance in Non-Contiguous OFDM (NC-OFDM) based CRs [2]. In this paper combined a clipping and filtering technique and the end-to-end QoS maintenance in Non-Contiguous OFDM (NC-OFDM) based CRs to reduce PAPR and calculated performance in non-contiguous bands spectrum of OFDM based CR system.

2. SYSTEM MODEL

The system model is defined as signal model, channel model and PAPR reduction technique. The system model can be described as below:

2.1 Signal Model

In a basic OFDM system the input data symbols are supplied into a channel encoder where data are mapped onto BPSK/QPSK/QAM constellation.

The data symbols are converted into serial to parallel before using Inverse Fast Fourier Transform (IFFT) block to get the time domain representation of OFDM symbols. Time domain symbols can be represented as:

$$x_n = \text{IFFT} \{X_k\}$$

$$= \frac{1}{N} \sum_{k=0}^{N-1} X_k e^{j \frac{2\pi n k n}{N}} \quad 0 \leq n \leq N-1 \quad (1)$$

where,

X_k is the transmitted symbol of the k^{th} subcarriers
 N is the number of subcarriers.

The time domain signal is regularly extended to prevent Inter Symbol Interference (ISI) from the former OFDM symbol using cyclic prefix (CP).

The D/A Converter performed to convert the baseband digital signal into analog signal. This operation is executed in D/A block shown in Figure 1. Then, the analog signal proceeded to the Radio Frequency (RF) frontend. The RF frontend start to operations after receiving the analog signal. The signal is converted to RF frequencies using mixer and amplified by Power Amplifier (PAs) and then transmitted throughout antennas. At the receiver side, the received signal is converted to base band signal while passing throughout by the RF block.

The analog signal is re-sampled and digitized by the A/D Converter and CR is removed from the signal. The received signal in the frequency domain obtained from the Fast Fourier Transform (FFT) block is represented as:

$$Y(k) = H(k)X(k) + W(k) \quad (2)$$

where, $Y(k)$ is the received signal of the k^{th} subcarriers, $H(k)$ is the frequency response of the channel and $W(k)$ is the additive noise which is generally assumed to be Gaussian random variable with zero mean and variance of σ_w^2 . After this, FFT signals are de-interleaved and decoded to turn into the original signal.

In OFDM system, the realization of large number of non-contiguous subcarriers by collective procedure for high data rate transmission is referred to as a Non-Contiguous OFDM (NC-OFDM) [3]. NC-OFDM is provided the necessary of agile spectrum usage for the target licensed spectrum

if spectrum is occupied by primary and secondary users. The spectrum sensing has right to deactivate the spectrum for the secondary user during spectrum occupied by primary user. Moreover, dynamic spectrum sensing can be retrieved the information while the active subcarriers are located in the vacant spectrum bands.

Fundamentally, the NC-OFDM and OFDM are quite similar in the case of transmission and reception. However, an NC-OFDM technique is offered very significant improvement for growing scarcity of the large contiguous frequency spectrum, i.e. dynamic spectrum pooling for high data rate transmissions.

2.2 Channel Model

The communication channel is the physical medium connecting the transmitter with the receiver. This paper introduced the Additive White Gaussian Noise (AWGN) channel and Rayleigh fading channel in the proposed system.

The simplest channel model in wireless communication is the well known Additive White Gaussian Noise (AWGN) model which is presented as follows:

$$Y(t) = X(t) + N(t), \quad (3)$$

where, $X(t)$ is the transmitted signal and $N(t)$ is the AWGN.

Multipath is the propagation phenomenon that results in radio signal reaching the receiver antennas via multiple propagation paths. Rayleigh fading model performs as reasonable channel model when there are many objects (such as building and mountain) in the propagation environment which scatter the radio signal before it arrives at the receiver.

Rayleigh distribution is given by:

$$P(r) = \begin{cases} \frac{r}{\sigma^2} \exp\left[-\frac{r^2}{2\sigma^2}\right], & 0 \leq r \leq \infty \\ 0, & r < 0 \end{cases} \quad (4)$$

where, σ^2 is the variance of the Rayleigh distributed variable.

Peak-to-Average Power Ratio

The PAPR of the OFDM signal can be written as:

$$\text{PAPR}\{s(t), T\} = \frac{\max_{t \in T} [s(t)]^2}{E\{[s(t)]^2\}} \quad (5)$$

where,

$s(t)$ is the original signal

T is the time interval

$\max_{t \in T} [s(t)]^2$ is the peak signal power

$E\{[s(t)]^2\}$ is the average signal power

3. PROPOSED SYSTEM

This work described the NC-OFDM based CR architecture's block diagram. In order to reduce the high peak power ratio are introducing repeated clipping and frequency domain filtering are introduce which is demonstrated in Figure 1.

In this paper, we utilized previous channel information in Non-Contiguous OFDM (NC-OFDM) based CRs under dynamic spectrum sharing environments. In the conventional OFDM system has the resource allocation problem where the allocated transmission spectrum is fixed. In the CR system, the operating bandwidth is not always fixed and spectrum is also co-shared. The channel and power status will be tracked by this system and provided reliable response from

channel state information to the transmitter. In this paper, NC-OFDM is considered and calculated the SNRs of sub-channels under total power constraints. It has been adaptively preferred high PAPR reduction approach during the operation by repeated clipping and filtering.

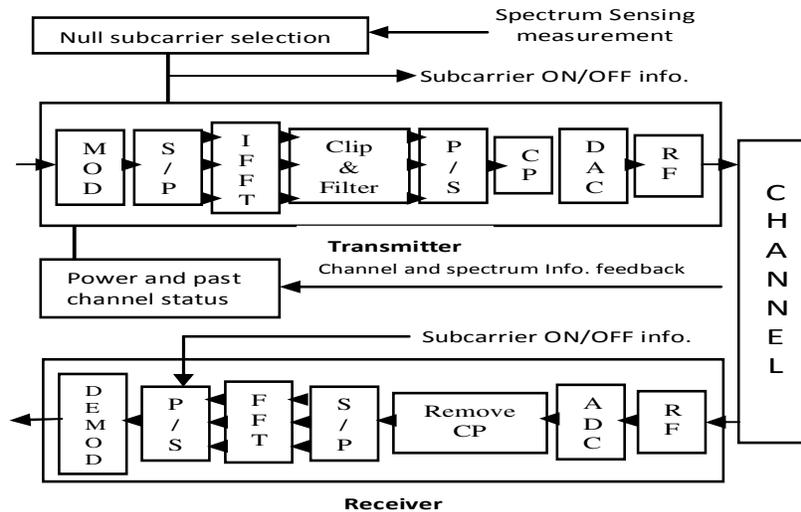


FIGURE 1: NC-OFDM based Cognitive Radio.

4. THE SIMULATION RESULTS

It is considered an NC-OFDM transceiver employing 10 MHz bandwidth, 256 FFT block size, and clipping ratio is 4. There are three modulation techniques (e.g. BPSK, QPSK, and QAM16 for clipping and filtering) as well as two channel models (e.g. AWGN and Rayleigh fading channel) used in the simulations. Finally, the two modes of output of the proposed system are shown: (1) PAPR reduction, and (2) BER calculation.

The complementary cumulative distribution functions (CCDF) of the PAPR for the transmitted signal are plotted in Figure 2, 3, and 4, where the PAPR techniques are being employed by the clipping and filtering. It is evident from these results that the PAPR can be improved by using repeated clipping and filtering.

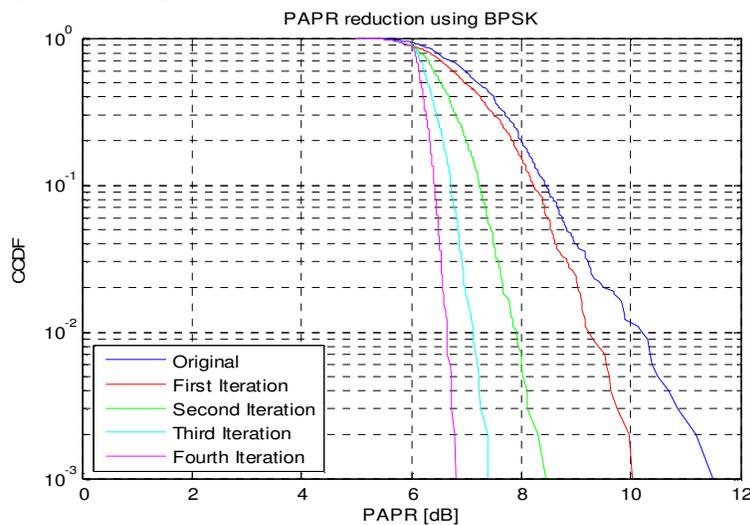


FIGURE 2: PAPR reduction using BPSK with clipping and filtering.

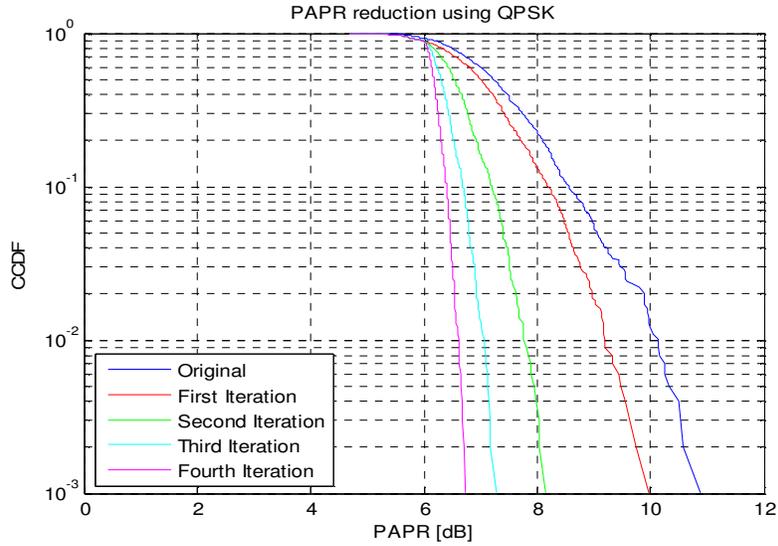


FIGURE 3: PAPR reduction using QPSK with clipping and filtering.

In Figure 2-4, by increasing PAPR for BPSK, QPSK, and QAM16, it is shown that CCDF is 10^{-3} at 11.5, 10.9, and 10.7 dB, but during next iteration, same 10^{-3} is achieved at less PAPR, i.e., 10.05, 9.95, and 9.7 dB. At fourth iteration, PAPR is 6.8, 6.7, and 6.65 dB.

Repeated clipping and filtering are significantly reduce PAPR, where modulation schemas are BPSK, QPSK, and QAM16. PAPR of QAM16 is low compared to BPSK and QPSK which gives better result shown in Figure 4. According to the fourth iteration, PAPR are given respectively 6.8, 6.7, and 6.65 dB for BPSK, QPSK, and QAM16.

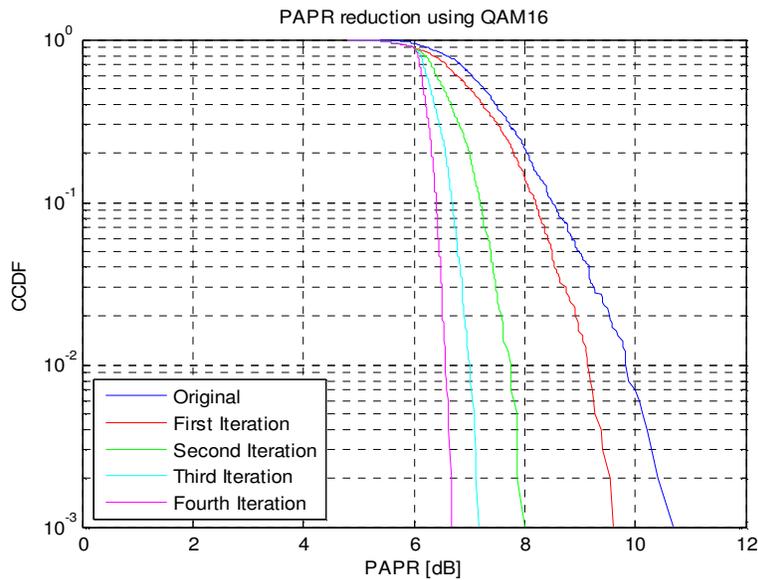


FIGURE 4: PAPR reduction using QAM16 with clipping and filtering.

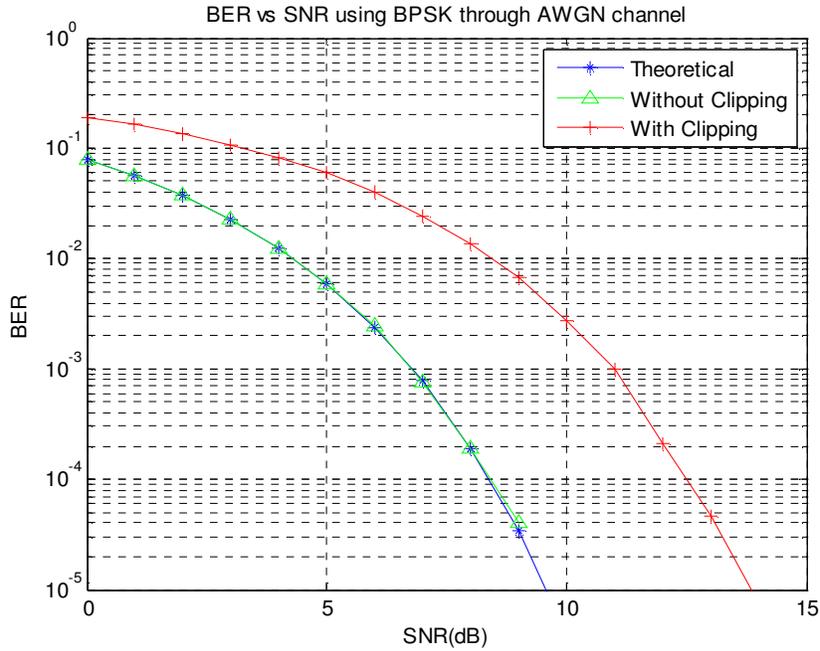


FIGURE 5: BER Vs SNR using BPSK through AWGN channel.

The performance of the system is measured by measuring BER using different channel models with different modulation schemes. The performance of NC-OFDM transceiver is shown in terms of curves representing BER against SNR values and are compared with theoretical, without clipping, and with clipping for every channel models.

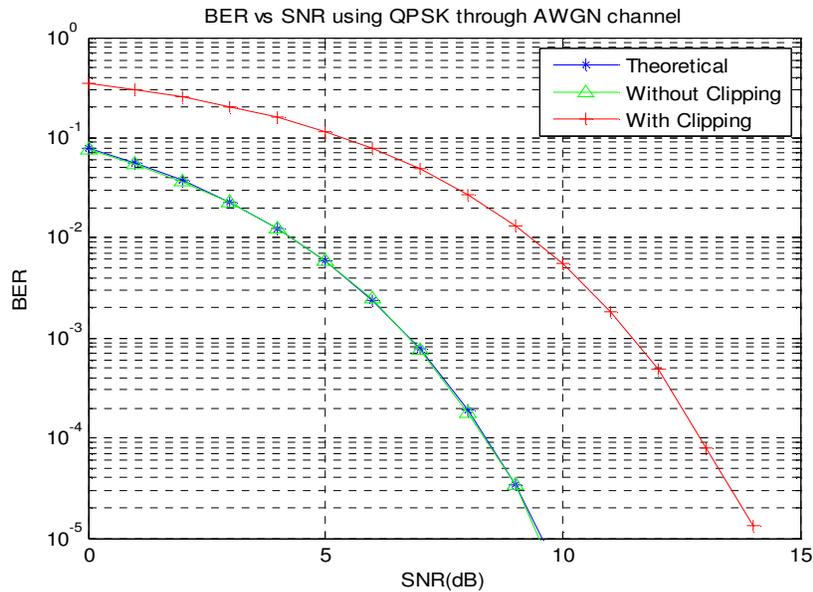


FIGURE 6: BER Vs SNR using QPSK through AWGN channel.

Figure 5 and 6 illustrated BER versus SNR for AWGN channel which is employed with BPSK and QPSK modulation schemes. In these figures, SNR for BPSK and QPSK, it can be seen that BER is 10^{-3} at 7 dB for theoretical and without clipping. On the other hand, SNR for BPSK and QPSK, BER is 10^{-3} at 11.2 and 11.5 dB with clipping.

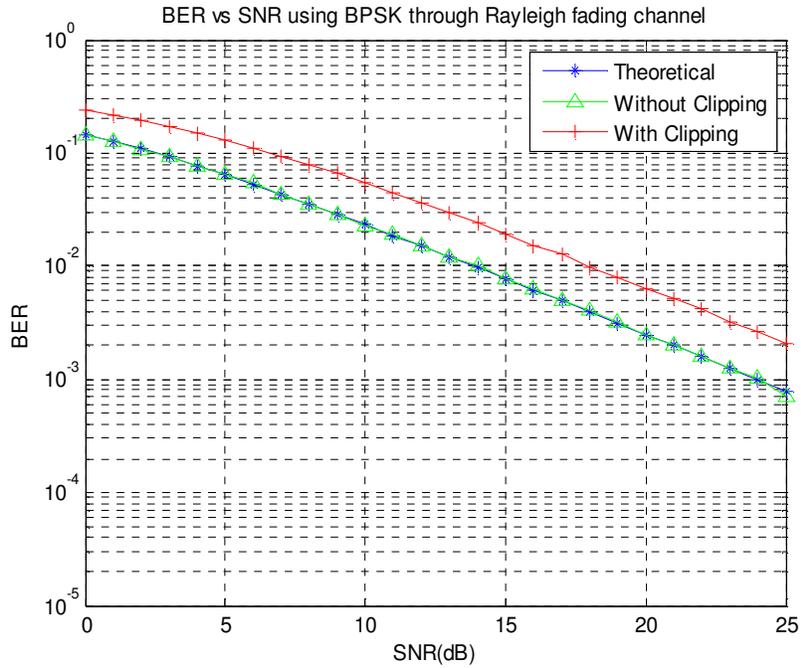


FIGURE 7: BER Vs SNR using BPSK through Rayleigh fading channel.

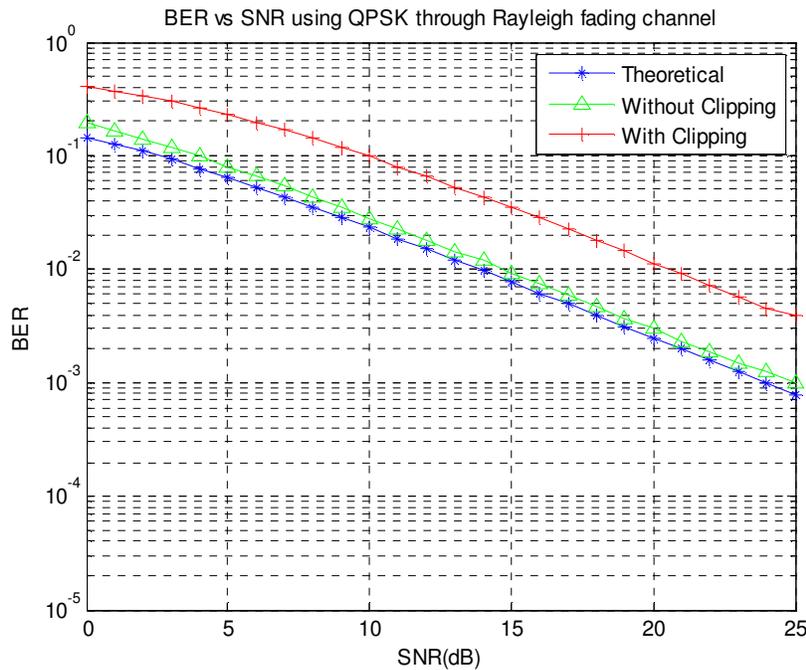


FIGURE 8: BER Vs SNR using QPSK through Rayleigh fading channel.

Figure 7 and 8 shown BER versus SNR for Rayleigh fading channel is using BPSK and QPSK modulation schemes. In these figures, SNR for BPSK and QPSK, it can be seen that BER is 10^{-2} at 14 dB for theoretical and without clipping. On the other hand, SNR for BPSK and QPSK, BER is 10^{-2} at 18 and 20.05 dB with clipping.

5. CONCLUSION

This paper presents an overview of Orthogonal Frequency Division Multiplexing (OFDM) and Cognitive Radio (CR). The OFDM is a smart candidate for CR systems for its flexibility and adaptability characteristics. This paper possesses a novel non-contiguous OFDM (NC-OFDM) technique, where the system achieves high data rate of non-contiguous subcarriers while simultaneously avoiding interference to the transmissions. The main goal of this paper work is to investigate PAPR reduction techniques for non-contiguous bands of OFDM based CR system. Simulation results of PAPR reduction has shown that the performance of QAM16 is good compared to other modulation schemes such as BPSK, QPSK. According to the BER results, we need higher value of SNR to achieve the same BER as compared to lower constellation modulation schemes.

ACKNOWLEDGEMENT

We would like to give our sincere gratitude to honorable Professor Abbas Mohammed, Signal Processing, Blekinge Institute of Technology (BTH), Sweden for his assistance and good guidance time after time which made our work become more precise and attractive.

REFERENCES

- [1] Armstrong, J, "New OFDM Peak-to-average power reduction scheme", vehicular technology conference, vol-1, 6-9 May 2001, pp 756-760.
- [2] W.M. Joseph, B.L. Khaled, and C. Zhigang, " Robust end-to-end QoS maintenance in non-contiguous OFDM based Cognitive Radio", Dept. of Electronic and Computer Engineering, Hong Kong University of Science and Technology, Hong Kong.
- [3] Rakesh Rajbanshi, Alexander M. Wyglinski, Gary J. Minden, "Adaptive-Mode Peak-to-Average Power Ratio Reduction Algorithm for OFDM-based Cognitive Radio," The University of Kansas, 2006.
- [4] Wilkison, T. A. and Jones A. E., "Minimization of the Peak to mean Envelope Power Ratio of Multicarrier Transmission Schemes by Block Coding," IEEE, Vehicular Conference, Vol.2, Jul. 1995.
- [5] Ahn, H., Shin, Y. m and Im, S., "A Block Coding Scheme for Peak to Average Power Ratio Reduction in an Orthogonal Frequency Division Multiplexing System ," IEEE Vehicular Conference Proceedings, Vol.1, May 2000.
- [6] Bauml , R. W. , Fischer, R. F. H. , and Huber, J. B., "Reducing the Peak to Average Power Ratio of Multicarrier Modulation by Selective Mapping," IEEE Electronics Letters, Vol.32, Oct. 1996.
- [7] Jayalath, A. D. S. , and Tellambura C., "The use of Interleaving to Reduce the Peak to Average Power Ratio of an OFDM Signals ," IEEE Global telecommunications conference, Vol.1, Nov. 2000.
- [8] J. Tellado-mourelo, "Peak to Average Power reduction of OFDM signals using Peak Reduction Carriers," James Cook University Signal Processing Research, 22-25 August, 1999.
- [9] Muller , S. H. , and Huber, J. B., "OFDM Reduced Peak to Average Power Ratio by optimum combination of Partial Transmit Sequences," IEEE Electronics letters, Vol.33, Feb. 1997.
- [10] Van Nee, R. , and Wild, A., "Reducing the Peak to Average Power Ratio of OFDM," in IEEE Vehicular Technology Conference, May 1998, Vol.3.

- [11] Foomooljareon , P and Fernando, W. A. C., "Input Sequence Envelope Scaling in PAPR Reduction of OFDM," in IEEE 5th International Symposium on Wireless Personal Multimedia Communications, , Vol.1, Oct 2002.
- [12] Tan, C. E. Wassell I. J., "Data bearing peak reduction carriers for OFDM systems," in IEEE Proceedings of the 2003 Joint Conference of the Fourth International Conference of Information, Communication Signal Processing and Fourth Pacific Rim Conference On Multimedia, , Vol.2, Dec 2003.

A Simple Design to Mitigate Problems of Conventional Digital Phase Locked Loop

M. Saber

*Department of Informatics
Kyushu University
744 Motoooka, Nishi-ku, Fukuoka-shi, 89-0395, Japan*

mohsaber@kairo.csce.kyushu-u.ac.jp

Y. Jitsumatsu

*Department of Informatics
Kyushu University
744 Motoooka, Nishi-ku, Fukuoka-shi, 89-0395, Japan*

jitsumatsu@inf.kyushu-u.ac.jp

M. T. A. Khan

*Ritsumeikan Asia Pacific University, College of Asia Pacific Studies
1-1 Jumonjibaru, Beppu, Oita, 874-8577, Japan*

tahir@apu.ac.jp

Abstract

This paper presents a method which can estimate frequency, phase and power of received signal corrupted with additive white Gaussian noise (AWGN) in large frequency offset environment. Proposed method consists of two loops, each loop is similar to a phase-locked loop (PLL) structure. The proposed structure solves the problems of conventional PLL such as limited estimation range, long settling time, overshoot, high frequency ripples and instability. Traditional inability of PLL to synchronize signals with large frequency offset is also removed in this method. Furthermore, proposed architecture along with providing stability, ensures fast tracking of any changes in input frequency. Proposed method is also implemented using field programmable gate array (FPGA), it consumes 201 mW and works at 197 MHz.

Keywords: Digital Phase-Locked Loop (DPLL), Field Programmable Gate Array (FPGA).

1. INTRODUCTION

In communication, the received signal has to be synchronized first before being demodulated. Synchronization includes carrier recovery and timing recovery. Carrier recovery is the estimation and compensation of carrier frequency and phase. Carrier frequency offset occurs in communication systems due to different reasons such as frequency mismatch between oscillators in transmitter and receiver, Doppler shift as a result of motion between transmitter and receiver and the channel noise. The estimation of frequency and phase in the presence of noise is a critical issue in signal processing applications [1: 3]. PLL is used in synchronization of frequency and phase of received signal. A PLL is a feedback control circuit; it operates by trying to lock the phase and frequency of input signal through the use of its negative feedback. Conventional PLL consists of three fundamental blocks namely a phase detector (usually multiplier), loop filter which is a low pass filter (LPF) and a voltage controlled oscillator (VCO) as shown in Fig. 1 [4, 5].

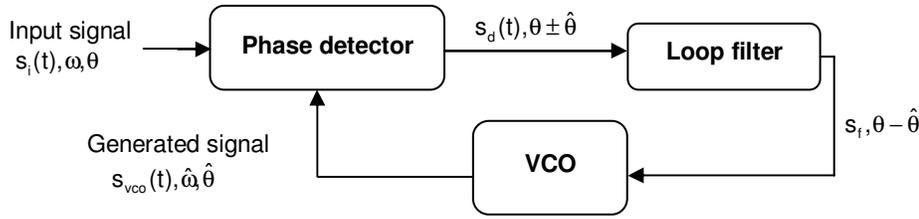


FIGURE 1: Phase Locked Loop.

Phase detector compares the phase of output signal to the phase of input signal. If there is a phase difference between the two signals, it generates an output voltage, which is proportional to the phase error of two signals. This output voltage passes through the loop filter and then as an input to the VCO. Due to this self-correcting technique, the output signal will be in phase with the input signal. When both signals are synchronized the PLL is said to be in lock condition. The bandwidth of a PLL depends on the characteristics of the phase detector, LPF and VCO [6:8]. The performance of PLL in estimating and tracking frequency has some limitations such as:

- 1) Higher order of LPF is needed to remove high frequency ripples inside the loop. Unfortunately this leads to unstable PLL [9: 11].
- 2) All PLLs will have ringing or settling time [12, 13].
- 3) The tracking range of PLL is very narrow and must be near the center frequency of VCO. As the order of LPF increases tracking range of PLL decreases [14, 15].

We propose a design which can not only estimate frequency and phase of the input signal but can also provide estimate of power of the received signal in the presence of background noise. Proposed structure consists of two loops, the first loop has the ability to estimate frequencies up to half the sampling frequency. The second loop is a low power digital phase-locked loop (DPLL). Using the frequency estimator's information, the second loop estimates phase of the received signal. The estimator is modeled with VHDL and implemented using FPGA. Proposed structure's hardware implementation consumes 201 mW and works at a frequency of 197 MHz.

2. PROPOSED ESTIMATOR

The main idea of proposed structure is to first estimate the large frequency offset using frequency estimator which has the ability to estimate frequencies with large range from zero to half sampling frequency with high accuracy. The second step is to use the output of frequency estimator to control a phase estimator which is a low noise PLL. Proposed architecture is shown in Fig. 2 and in Fig. 3 in which the frequency estimator block receives quadrature signal; the output of this block is the estimated power and the estimated frequency. Estimated frequency is then used in phase estimator block as the center frequency of VCO. In this manner PLL, will compensate only the difference in phase between the input signal and generated signal of VCO not in both phase and frequency. In the next subsections, we will provide a detailed explanation of the operation of each block in the proposed architecture.

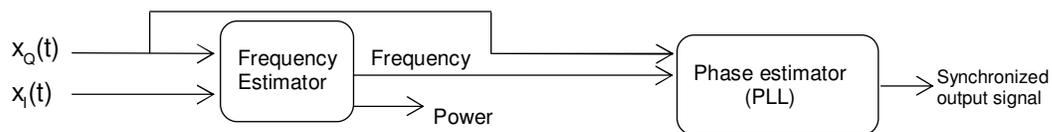


FIGURE 2: Proposed architecture.

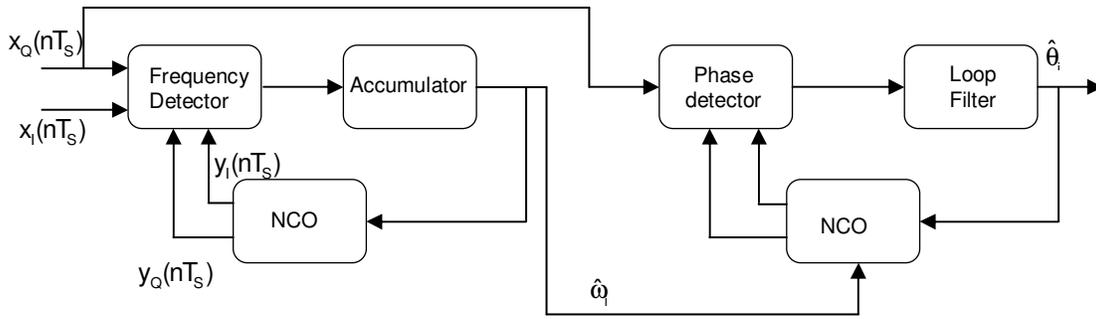


FIGURE 3: Operation of proposed architecture.

A. Frequency Estimator

Block diagram of the suggested frequency estimator is given in Fig. 4. It is like a PLL structure except that the phase detector is changed to frequency detector, LPF is changed to an accumulator and VCO is changed to quadrature numerically controlled oscillator (NCO) [16]. The quadrature input signal enters the frequency detector component which consists of multipliers, adders, subtractors, differentiator and squaring circuit. Fig. 4 explains discrete time implementation of the proposed estimator structure. We would also provide here the mathematical details of frequency and power estimation and stability of proposed architecture. Suppose the input quadrature signal and the quadrature signal generated by the oscillator are

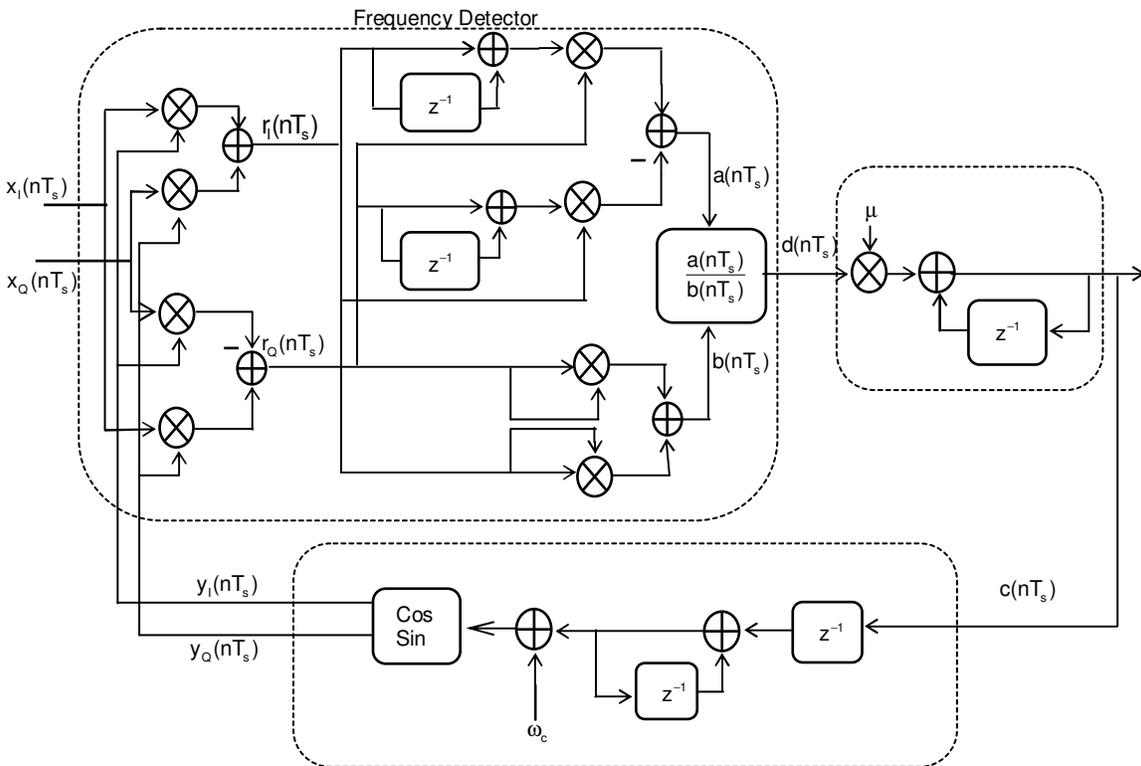


FIGURE 4: Discrete time model of proposed estimator.

1) Frequency and power estimation

We assume the input quadrature signal and the quadrature signal generated by the oscillator are

$$x_i(nT_s) = \sqrt{2p(nT_s)} \cos(\omega_i nT_s + \theta_i). \quad (1)$$

$$x_Q(nT_s) = \sqrt{2p(nT_s)} \sin(\omega_i nT_s + \theta_i). \quad (2)$$

$$y_i(nT_s) = \cos(\omega_o nT_s + \theta_o). \quad (3)$$

$$y_Q(nT_s) = \sin(\omega_o nT_s + \theta_o). \quad (4)$$

Where $p(nT_s)$ is the power of input signal. $\omega_i, \omega_o, \theta_i$ and θ_o are the input and output radian frequencies and phases respectively, T_s is the sampling time. Two input signals are multiplied with the two signals generated by oscillator. This results in two quadrature signals ($r_i(n), r_Q(n)$) given as:

$$r_i(nT_s) = x_i(nT_s)y_i(nT_s) + x_Q(nT_s)y_Q(nT_s) = \sqrt{2p(nT_s)} \cos(\Delta\omega nT_s + \Delta\theta), \quad (5)$$

$$r_Q(nT_s) = x_i(nT_s)y_Q(nT_s) - x_Q(nT_s)y_i(nT_s) = -\sqrt{2p(nT_s)} \sin(\Delta\omega nT_s + \Delta\theta). \quad (6)$$

Where $\Delta\omega = \omega_i - \omega_o, \Delta\theta = \theta_i - \theta_o$. The quadrature signals ($r_i(n), r_Q(n)$) pass through a differentiator circuit.

$$\begin{aligned} \frac{d}{dt}r_i(t) \Big|_{t=nT_s} &\approx \{r_i(nT_s) - r_i(n-1)T_s\} \times \frac{1}{T_s} \\ &= -\sqrt{2p(nT_s)}\Delta\omega \sin(\Delta\omega nT_s + \Delta\theta) + \left(\frac{d}{dt}\sqrt{2p(t)}\right) \Big|_{t=nT_s} \times \cos(\Delta\omega nT_s + \Delta\theta), \end{aligned} \quad (7)$$

$$\begin{aligned} \frac{d}{dt}r_Q(t) \Big|_{t=nT_s} &\approx \{r_Q(nT_s) - r_Q(n-1)T_s\} \times \frac{1}{T_s} \\ &= -\sqrt{2p(nT_s)}\Delta\omega \cos(\Delta\omega nT_s + \Delta\theta) - \left(\frac{d}{dt}\sqrt{2p(t)}\right) \Big|_{t=nT_s} \times \sin(\Delta\omega nT_s + \Delta\theta). \end{aligned} \quad (8)$$

The output of the differentiator is

$$\begin{aligned} a(nT_s) &= r_Q(nT_s) \times \frac{d}{dt}r_i(t) \Big|_{t=nT_s} - r_i(nT_s) \times \frac{d}{dt}r_Q(t) \Big|_{t=nT_s} \\ &= 2p(nT_s)\Delta\omega [\sin^2(\Delta\omega nT_s + \Delta\theta) + \cos^2(\Delta\omega nT_s)] \\ &= 2p(nT_s)\Delta\omega \end{aligned} \quad (9)$$

The signals ($r_i(n), r_Q(n)$) also at the same time passes through squaring circuit. The output of squaring circuit can be considered as power of input signal.

$$\begin{aligned} b(nT_s) &= r_i^2(nT_s) + r_Q^2(nT_s) \\ &= 2p(nT_s)[\sin^2(\Delta\omega nT_s + \Delta\theta) + \cos^2(\Delta\omega nT_s)] \\ &= 2p(nT_s). \end{aligned} \quad (10)$$

The output of divider is

$$d(nT_s) = \frac{a(nT_s)}{b(nT_s)} = \Delta\omega. \quad (11)$$

The accumulator starts to accumulate frequency difference until the generated frequency of oscillator is the same as the input frequency. The accumulator equation is

$$\mu d(nT_s) = c(nT_s) - c((n-1)T_s). \quad (12)$$

Where “ μ ” is a variable which determines the speed of locking and stability of the system. NCO output is adjusted according to the output of accumulator. Feedback loop of the estimator continues until there is no difference between the input frequency and NCO generated frequency. NCO equation is

$$\omega_o(nT_s) = \omega_c + \sum_{k=0}^{n-1} c(kT_s). \quad (13)$$

Where ω_c is the center frequency. When the generated frequency from NCO is equal to the input frequency, the accumulator saturates. Thus at $\omega_i(nT_s) = \omega_o(nT_s)$

$$c(nT_s) = \omega_i T_s = 2\pi f_i T_s \quad (14)$$

2) Stability

Linear model (z-model) of the system is shown in Fig. 5. Following equations are obtained from Fig. 5

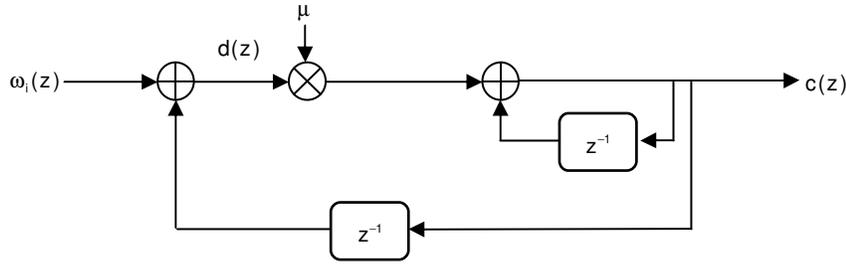


FIGURE 5: Z-model of proposed estimator.

$$c(z) = \omega_i(z) - z^{-1}c(z), \quad (15)$$

$$\mu d(z) = c(z)(1 - z^{-1}), \quad (16)$$

$$\omega_i(z) = c(z) \left(z^{-1} + \frac{1 - z^{-1}}{\mu} \right). \quad (17)$$

Therefore, transfer function of the proposed estimator

$$T(z) = \frac{c(z)}{\omega_i(z)} = \frac{\mu z}{z - (1 - \mu)}. \quad (18)$$

The variable μ controls stability of the system and its range must be $0 \leq \mu < 1$. For small values of μ the system is more stable but takes long settle time, for greater values of μ settling time decreases.

B. Phase Estimator

The phase estimator used in this paper is a low noise digital phase-locked loop (DPLL) [17]. Conventional DPLL cannot completely remove the high frequency ripples present in the signal. These ripples come from the phase detector (multiplier) which multiplies the input signal with the signal generated by NCO. The result of multiplication is a signal with two terms; high frequency term and low frequency term. In conventional PLL, it assumed that the LPF will remove the high frequency term completely, but first order LPF is not able to remove all ripples. One solution to this problem is to use higher order LPF but this leads to instability of PLL. The other solution could be lowering the cut-off frequency of LPF, but also this will lower the bandwidth of PLL. Our proposed structure of DPLL aims to use a first order LPF to provide stability for the system besides the ability to remove the ripples. Fig. 6 shows the proposed design of DPLL.

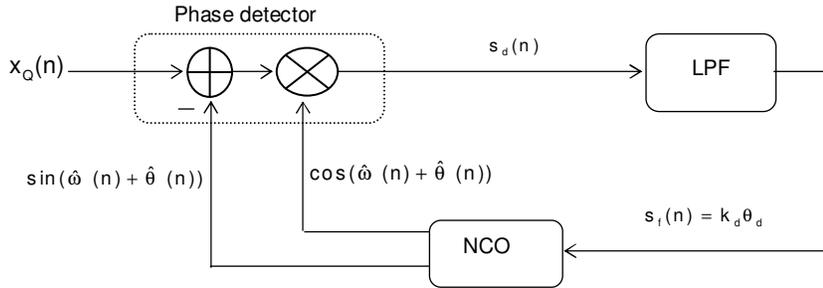


FIGURE 6: Block diagram of proposed PLL.

$$s_d(nT_s) = k_d \{A_i \sin(\omega_i nT_s + \theta_i) - A_o \sin(\hat{\omega}_i nT_s + \hat{\theta})\} \times A_o \cos(\hat{\omega}_i nT_s + \hat{\theta}). \quad (19)$$

Where $A_i = \sqrt{2p(nT_s)}$, A_o , $\hat{\omega}_i$, $\hat{\theta}$ are amplitude, radian frequency and phase of NCO generated signal respectively, k_d is the gain of multiplier. The first stage of propose architecture estimates the frequency so at $\omega_i = \hat{\omega}_i$

$$s_d = \frac{A_i A_o k_d}{2} \{\sin(2\omega_i nT_s + \theta_i + \hat{\theta}) + \sin(\theta_i - \hat{\theta}) - \sin(2\omega_i nT_s + 2\hat{\theta})\}. \quad (20)$$

From (20) a new term is subtracted from the high frequency term before passing to LPF. Therefore, in this case LPF just have to remove the residual of subtraction instead of removing the entire high frequency term in conventional PLL. As such, a first order LPF is able to remove the residual and stability of DPLL is enhanced. After LPF the resulting signal will be

$$s_i(nT_s) = \frac{A_i A_o k_d}{2} \sin(\theta_i - \hat{\theta}). \quad (21)$$

At $(\theta_i - \hat{\theta}) \ll 1$, then

$$s_i(n) = \frac{A_i A_o k_d}{2} (\theta_i - \hat{\theta}). \quad (22)$$

This signal is used to control the phase of generated signal from NCO, the DPLL continues to vary the phase until locking occurs and $\theta_i = \hat{\theta}$.

3. SIMULATION RESULTS

We have also carried out computer simulations to compare working of our proposed method with conventional DPLL having both 1st and 2nd order LPF. We have investigated the frequency and phase estimation and tracking performances in both with and without AWGN.

A. Frequency Estimation

First of all we provide frequency estimation comparison between proposed and conventional DPLL. An input signal with frequency of 10 kHz is applied. In these simulations we used a conventional DPLL with both 1st order LPF and 2nd order LPF. Each LPF has cut-off frequency of 1 kHz. The parameters of DPLL are: NCO gain=1024 rad/ v. s, $f_i=10.5$ kHz, $f_{nco}=10$ kHz, $f_s=100$ kHz, while the parameters of proposed estimator are $f_{nco}=100$ Hz, $f_s=100$ kHz.

Results of simulations are shown in Fig. 7. Performance of a conventional DPLL with a first order LPF is shown in Fig. 7(a). It is evident that the LPF cannot eliminate the double frequency component completely and overshoot occurs. The settling time, which is time taken by the PLL to reach the frequency estimate is a very important measure of PLL performance. In this case settling time is around 1 ms. In case of DPLL, to remove the higher frequency ripple, a higher order LPF is employed. We have also performed simulations of PLL using 2nd order LPF and results are shown in Fig. 7(b). It can be seen that the high frequency ripple has been suppressed better than a first order LPF. However, the overshoot and ringing still exist and settling time has increased. Furthermore, estimation range is reduced and stability becomes critical in this case. In Fig. 7(c) the proposed method estimates the input frequency in 0.5 ms with no ripples (although no loop filter is used in the system), the estimation range is wide from 0 up to half the sampling frequency. The system is also more stable than conventional PLL.

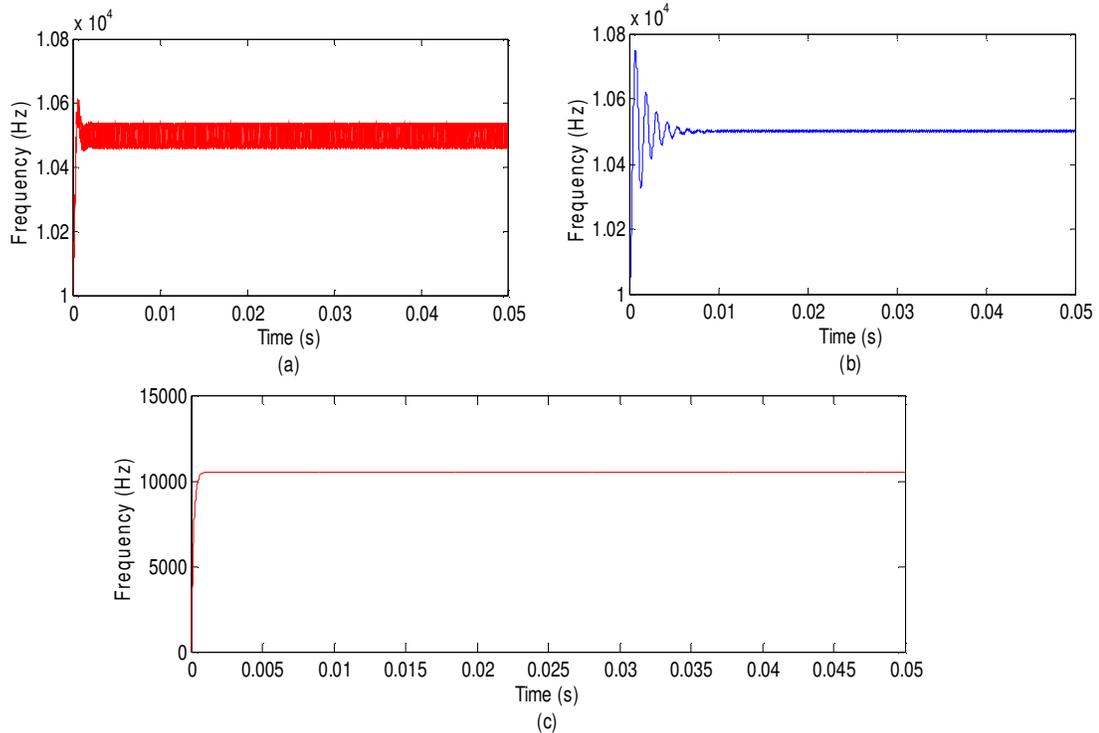


FIGURE 7: Frequency estimation curves of (a) Conventional DPLL with 1st order LPF, (b) Conventional DPLL with 2nd order LPF, (c) Proposed Estimator.

B. Frequency and Phase Estimation

The aim of this simulation is to determine the locking time for both frequency and phase. The estimator receives input sinusoidal signal $x_Q(n) = \sin(10000n + \pi/4)$, the center frequency of the NCO is $f_{\text{nco}} = 100 \text{ Hz}$. The proposed estimator estimates the input frequency as shown in Fig. 8 while phase estimation is shown in Fig. 9. It's noticed that the frequency estimator locked the frequency after a very short time ($3\mu\text{s}$), phase estimator locked the phase after (0.1 ms) as shown in Fig. 8.

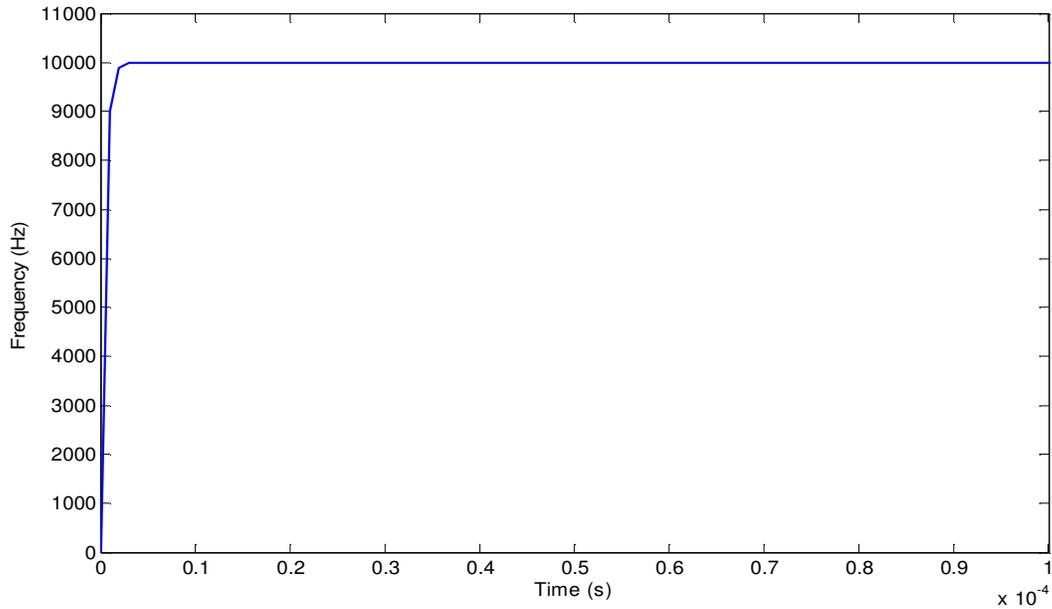


FIGURE 8: Estimated Frequency.

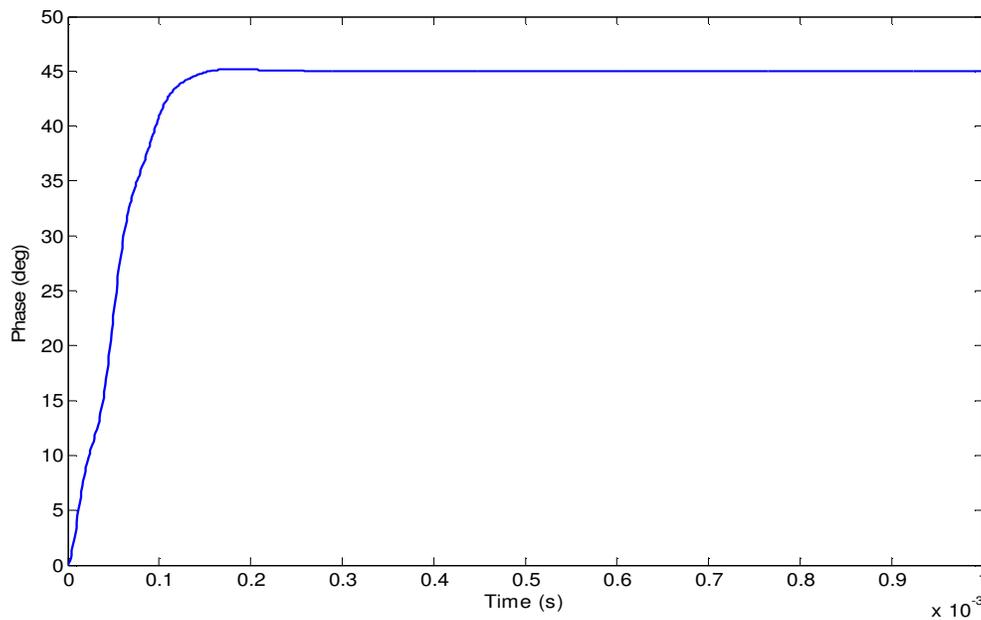


FIGURE 9: Estimated Phase.

C. Frequency and Phase Estimation with AWGN

When the input signal is corrupted with AWGN, two moving average filters (MAF) with length of 10 stages are used for I and Q phase input signals. The MAF is a low pass finite impulse response (FIR) filter commonly used for smoothing an array of sampled data/signal. The following simulations are performed to investigate the effect of AWGN on both conventional DPLL with 1st order LPF and 2nd order LPF and proposed estimator. MAF with the same lengths are added to both architectures. An input signal is applied to three architectures with three different frequencies 50.5, 51.5, 52.5 kHz, all input signals have input phase $\pi/8$ and AWGN with SNR= 10 dB at sampling frequency of 10 MHz. Fig. 10 shows estimations of the three frequencies with conventional DPLL with 1st order LPF, 2nd order LPF and proposed estimator respectively. DPLL with 1st order DPLL estimates the first frequency with high noise level and noise variance is 3269.7. While DPLL with DPLL with 2nd order LPF estimates the second frequency with a reduction in noise level causes the variance to be 1618. Proposed estimator estimates the third frequency with a higher reduction in noise level causes the variance to further reduce to 999. It can be seen that, the proposed architecture produces the smallest values for each estimated frequency although loop filter is not used.

Fig. 11 shows the estimates of the input phase for the three architectures. DPLL with 1st order LPF, estimates the input phase with variance 128.5932. DPLL with 2nd order DPLL estimates the phase with variance 118.7396. The proposed estimator estimates the input phase with variance 99.2743. As the proposed estimator estimates the input frequency with further reduction in variance, it also estimates the input phase with reduction in variance compared to conventional DPLL.

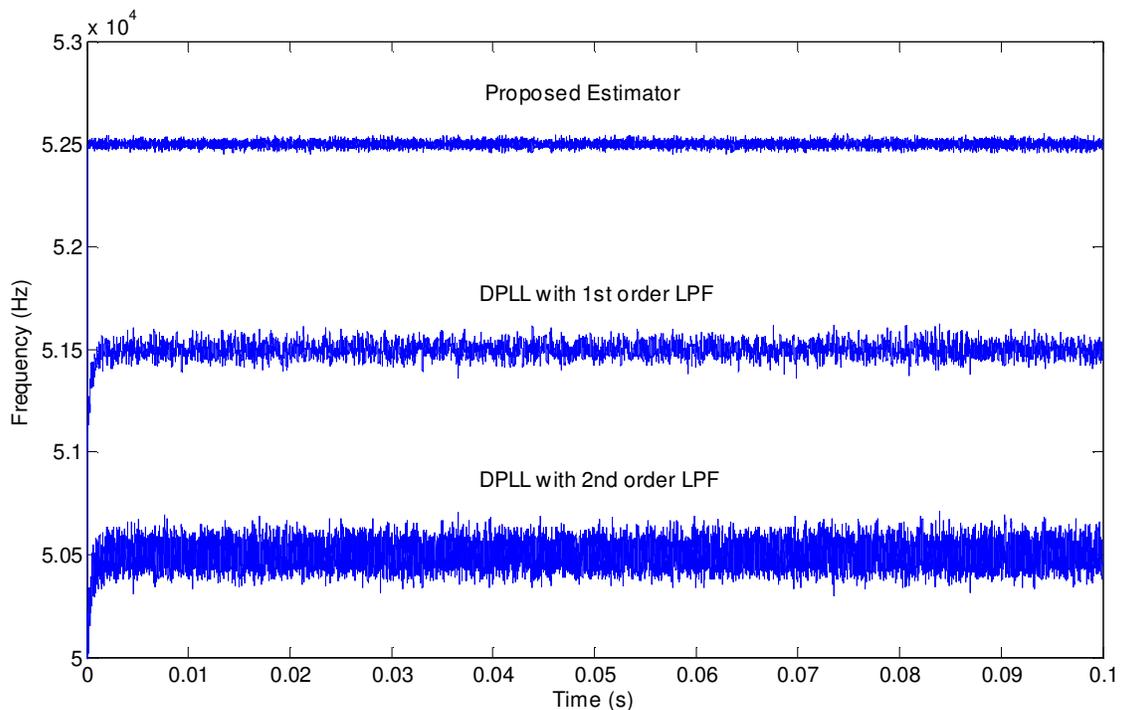


FIGURE 10: Frequency Estimation Comparison.

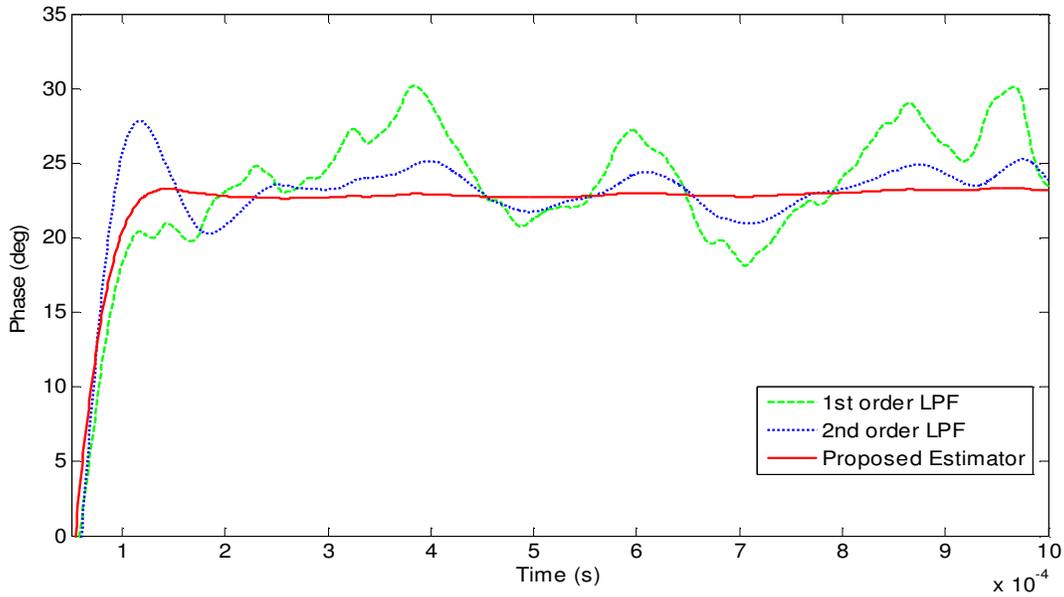


FIGURE 11: Phase Estimation Comparison.

D. Frequency, Power and Phase tracking

We also checked performance of the proposed method in tracking frequency changes and estimating amplitude/power variations. Frequency of the input sinusoidal signal is varied randomly within the range 10 kHz: 60 kHz and phases within the range $\pi/8 : \pi/2$ rad. The input signal power is also randomly varied within the range 0.06125:0.21125 watt.

The simulations are performed with SNR 10 dB and sampling frequency of 1 MHz. Fig. 12 shows the tracking of the input frequency, while Fig. 13 gives estimation of power. It can be seen that the estimator tracks changes in input frequency and amplitude very fast. Fig. 14 shows phase estimation, it is seen from the figure that the locking time for each phase is greater than the frequency locking time because phase estimation begins after frequency estimation is finished this also explains the gaps between phases.

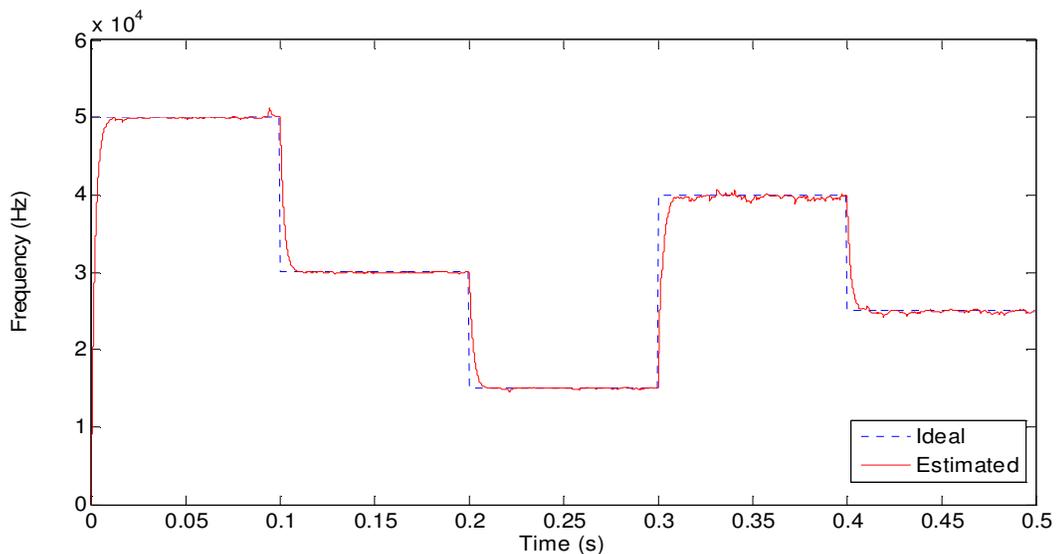


FIGURE 12: Frequency tracking.

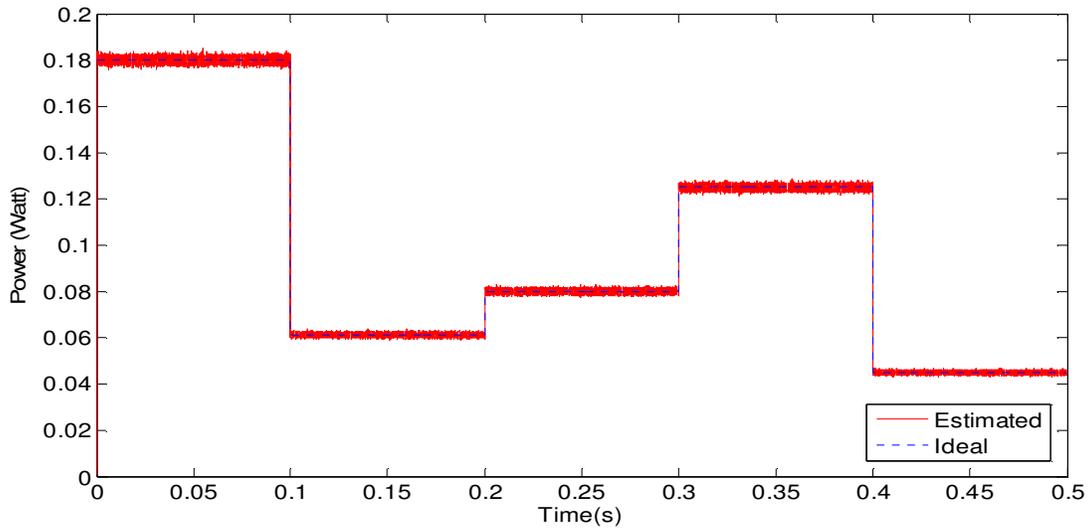


FIGURE 13: Power estimation.

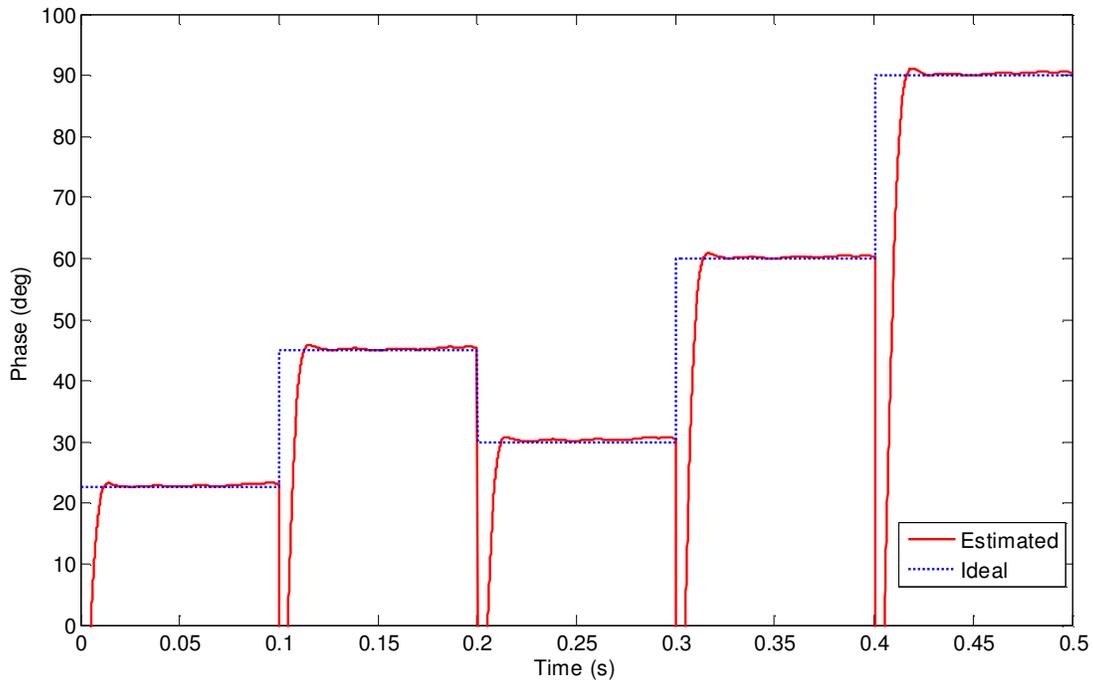


FIGURE 14: Phase tracking.

4. FPGA IMPLEMENTATION

The estimator is modeled using XILINX system generator tool. VHDL code is used to describe the proposed estimator [18:21]. All signals of the estimator model are fixed point signals with 16 bits. The estimator was implemented using XCSD1800A-4FG676C Spartan-3A DSP board. The hardware model simulation results agree with the simulations in locking time and stability. FPGA resources utilization for proposed architecture, which indicates how many hardware components are used by the model, is provided in Table 1. It can be seen that the proposed method requires fewer hardware resources. Implementation results indicate that the proposed design provides fast

operation and low power consumption. Furthermore, it can be observed that the method has a simple design and avoids complexities.

TABLE 1: FPGA resource utilization

Component	Used	Available	Utilization
Number of slice flip flops	180	33280	1%
Number of 4 input LUTs	1100	33280	3%
Number of occupied slices	594	16640	3%
Number of DSP48As	13	84	7%
Number of bonded IOBs	97	519	18%
Maximum frequency	201 MHz		
Power consumption	197 mW		

5. CONCLUSION

Conventional DPLL used for carrier synchronization is unable to synchronize signals with large frequency shift. DPLL has other drawbacks such as overshoot, ringing, limited tracking range and instability. This paper presented a new structure which can solve these problems. The proposed architecture divided the synchronization task to two steps. The first step is the estimation of the large frequency offset through a frequency estimator. The second step is to use the estimated frequency to estimate the phase through a low noise DPLL, which has the ability to estimate phase accurately by removing the internal noise of DPLL. Computer simulations performed showed significant improvements over conventional DPLL. The paper also discussed hardware implementation of the system designed and modeled using VHDL and implemented using FPGA circuit. Implementation results indicate that the estimator has a simple design, faster operation and low power consumption.

6. ACKNOWLEDGEMENT

This research is partially supported by Grant-in-Aid for Scientific Research (B) no.20360174, and the Aihara Project, the First program from JSPS, initiated by CSTP.

7. REFERENCES

- [1] J. G. Proakis, M. Salehi, Digital Communications, Mcgraw-Hill, 2008.
- [2] Y. Wang, Y. Zhang, Z. Yang, "Parallel Carrier Recovery in All-Digital Receiver," in Proc. WICOM, pp. 1-4, 2009.
- [3] G. Karam, J. Kervarec, H. Sari, P. Vandamme, "All-digital implementation of the carrier recovery loop in digital radio systems," International Conference on Communications, pp. 175-179, 1991.
- [4] M. Padua, S. Deckmann, G. Sperandio, F. Marafao, D. Colon "Comparative analysis of Synchronization Algorithms based on PLL, RDFT and Kalman Filter," IEEE International Symposium on Industrial Electronics, pp. 964-970, 2007.

- [5] Y. Linn, "Robust M-PSK phase detectors for carrier synchronization PLLs in coherent receivers: theory and simulations," *IEEE Transactions on Communications*, Vol. 57, pp. 1794-1805, 2009.
- [6] K. Lim, S. Choi, B. Kim, "Optimal loop bandwidth design for low noise PLL applications," *Proceedings of the ASP-DAC '97*, pp. 425-428, 1997.
- [7] J. Roche, W. Rahadjandrabey, L. Zady, G. Bracmard, D. Fronte, "A PLL with loop bandwidth enhancement for low-noise and fast-settling clock recovery," *In Proc. of Electronics, circuits and systems IEEE conference*, pp. 802-805, 2008.
- [8] W. Ping-Ying, Fu. Chia-Huang, "All digital modulation bandwidth extension technique for narrow bandwidth analog fractional-N PLL," *in Proc. of ESSCIR*, pp. 270-273, 2010.
- [9] A. Carlosena, A. Manuel- Lazaro, "Design of Higher-Order Phase-Lock Loops," *IEEE Transaction on Circuits and systems II*, vol. 54, pp. 9-13, 2007.
- [10] S. Changhong, C. Zhongze, Z. Lijun, L. Yong, "Design and Implementation of Bandwidth Adaptable Third-order All digital Phase-Locked Loops," *in Proc. WICOM*, pp.1-4, 2010.
- [11] X. hongbing, G. Peiyuan, L. Shiyi, "Modeling and simulation of Higher-order PLL rotation tracking system in GNSS receiver," *in Proc. ICMA*, 2009, pp. 1128-1133.
- [12] P. Hanumolu, "W Gu-Yeon, M. un-ku, A Wide- Tracking Range clock and Data recovery Circuit," *IEEE Transaction on Solid state circuits*, vol. 43, pp. 425-439, 2008.
- [13] D. Green. "Lock-In, Tracking, and acquisition of AGC-aided phase-locked loops," *IEEE Transaction on Circuits and Systems*, vol. 32, pp. 559-568, 1985.
- [14] L. Verrazzani, "Pull-In Time and Range of Any Order Generalized PLL." *IEEE Transaction on Aerospace and Electronic systems*, Vol. AES14, pp.329-333, 2007.
- [15] P. Hanumolu, "W Gu-Yeon, M. un-ku, A Wide- Tracking Range clock and Data recovery Circuit," *IEEE Transaction on Solid state circuits*, vol. 43, pp. 425-439, 2008.
- [16] M. Saber, M.T.A. Khan, Y. Jitsumatsu, "Frequency and Power Estimator for Digital Receivers in Doppler Shift environment." *Signal Processing: An international journal (SPIJ)*, vol. 5, pp. 192-208, Dec. 2011.
- [17] M. Saber, Y. Jitsumatsu, M.T.A. Khan, "Design and Implementation of Low Ripple Low Power Digital Phase-Locked Loop." *Signal Processing: An international journal (SPIJ)*, vol. 4, pp. 304-317, Feb. 2011.
- [18] P. Pong Chu. *RTL Hardware Design Using VHDL: Coding for Efficiency, Portability and Scalability*. Wiley-IEEE Press, 2006.
- [19] Xilinx Inc. *system generator for DSP user guide*. Xilinx, 2009.
- [20] W.Y. Yang. *Matlab/Simulink for digital communication*. A-Jin, 2009.
- [21] Pong p. chu. *FPGA Prototyping by VHDL Examples: Xilinx Spartan-3 Version*. Wiley-Interscience, 2008.

Adaptive Variable Step Size in LMS Algorithm Using Evolutionary Programming: VSSLMSEV

Ajjaiah H.B.M

*Research scholar
Jyothi institute of Technology
Bangalore, 560006, India*

hbmajay@gmail.com

Prabhakar V Hunagund

*Dept.of PG atudies and Research in applied electronics
Jnana Ganga, Gulbarga University
Gulbarga - 585106, India*

prabhakar_hunagund@yahoo.co.in

Manoj Kumar Singh

*Director
Manuro Tech Research
Bangalore, 560097, India*

mksingh@manuroresearch.com

P.V.Rao

*Dept.of ECE, RGIT,
Bangalore, India*

pachararao@rediffmail.com

Abstract

The Least Mean square (LMS) algorithm has been extensively used in many applications due to its simplicity and robustness. In practical application of the LMS algorithm, a key parameter is the step size. As the step size becomes large /small, the convergence rate of the LMS algorithm will be rapid and the steady-state mean square error (MSE) will increase/decrease. Thus, the step size provides a trade off between the convergence rate and the steady-state MSE of the LMS algorithm. An intuitive way to improve the performance of the LMS algorithm is to make the step size variable rather than fixed, that is, choose large step size values during the initial convergence of the LMS algorithm, and use small step size values when the system is close to its steady state, which results invariable step size Least Mean square (VSSLMS) algorithms. By utilizing such an approach, both a fast convergence rate and a small steady-state MSE can be obtained. Although many VSSLMS algorithmic methods perform well under certain conditions, noise can degrade their performance and having performance sensitivity over parameter setting. In this paper, a new concept is introduced to vary the step size based upon evolutionary programming (VSSLMSEV) algorithm is described. It has shown that the performance generated by this method is robust and does not require any presetting of involved parameters in solution based upon statistical characteristics of signal.

Keywords: Adaptive Equalization, LMS Algorithm, Step Size, Evolutionary Programming, MSE.

1. INTRODUCTION

The recent digital transmission systems impose the application of channel equalizers with short training time and high tracking rate. These requirements turn our attention to adaptive algorithms, which converge rapidly. One of the most important advantages of the digital transmission systems for voice, data and video communications is their higher reliability in noise environment in comparison with that of their analog counterparts. Unfortunately most often the digital transmission of information is accompanied with a phenomenon known as intersymbol interference (ISI). Briefly this means that the transmitted pulses are smeared out so that pulses that correspond to different symbols are not separable. Depending on the transmission media the main causes for ISI are: cable lines – the fact that they are band limited; cellular communications – multipath propagation. Obviously for a reliable digital transmission system it is crucial to reduce

the effects of ISI and it is where the adaptive equalizers come on the scene. Two of the most intensively developing areas of digital transmission, namely digital subscriber lines and cellular communications are strongly dependent on the realization of reliable channel equalizers. One of the possible solutions is the implementation of equalizer based on filter with finite impulse response (FIR) employing the well known LMS algorithm for adjusting its coefficients. The LMS algorithm is one of the most popular algorithms in adaptive signal processing. Due to its simplicity and robustness, it has been the focus of much study and its implementation in many applications. The popularity stems from its relatively low computational complexity, good numerical stability, simple structure, and ease of implementation in terms of hardware. The essence of LMS algorithm is to update the adaptive filter coefficients recursively along the negative gradient of estimate error surface. Conventional algorithm uses a fixed step-size to perform the iteration, and to get a compromise between the conflict of fast convergence and small steady-state MSE. A small step-size could ensure small MSE with a slow convergence, where as a large step-size will provide a faster convergence and better tracking capabilities at the cost of higher steady-state MSE. Therefore, fixed step-size LMS algorithm definitely cannot settle this contradiction. Consequently, many variable step-size algorithms were proposed to solve the problem. Though these algorithms could accelerate convergence and deduce steady-state MSE to some extent, they failed to analyze the optimality of variable step-size LMS further. LMS algorithm is described by the following equations:

$$e(n) = d(n) - X^T(n) * W(n) \quad \text{-----} \quad (1)$$

$$W(n+1) = W(n) + \mu e(n) X(n) \quad \text{-----} \quad (2)$$

where μ is learning step, $X(n)$ is the input vector at sampling time n , $W(n)$ is the coefficient vector of the adaptive filter, $d(n)$ is the expected output value $e(n)$ is the deviation error, dimension of $W(n)$ is the length of the adaptive filter.

2 RELATED WORK

In this work we introduced a novel method to obtain an optimal step-size and an algorithm for LMS. The algorithm runs iteratively and convergence to the equalizer coefficients by finding the optimal step-size which minimizes the steady-state error rate at each iteration. No initialization for the step-size value is required. Efficiency of the proposed algorithm is shown by making a performance comparison between some of the other LMS based algorithms and optimal step-size LMS algorithm [1]. A variation of gradient adaptive step-size LMS algorithms are presented. They propose a simplification to a class of the studied algorithms [2]. Adaption in the variable step size LMS proposed by [3] based on weighting coefficients bias/variance trade off. Authors in [4] examine the stability of VSLMS with uncorrelated stationary Gaussian data. Most VSLMS described in the literature use a data-dependent step-size, where the step-size either depends on the data before the current time (prior step-size rule) or through the current time (posterior step-size rule). It has often been assumed that VSLMS algorithms are stable (in the sense of mean-square bounded weights), provided that the step-size is constrained to lie within the corresponding stability region for the LMS algorithm. The analysis of these VSLMS algorithms in the literature typically proceeds in two steps [5], [6]. First, a rigorous stability analysis is attempted, apparently leading to conditions for MSE bounded weights and bounded MSE, and second, an approximate performance analysis is carried out, including convergence to and characterization of the asymptotic weight mean, covariance, and MSE. Thus one can at least guarantee stability (MS bounded weights) rigorously would seem to support the performance analysis. Two methods of variable step-size normalized least mean square (NLMS) and affine projection algorithms (APA) with variable smoothing factor have presented in [8]. With the Simulation results they have illustrated that the proposed algorithms have improvement in convergence rate and lower small adjustment error.

3. BASIC CRITERIA FOR PERFORMANCE

The performance of the LMS adaptive filter can be characterized in three important ways: i) the adequacy of the FIR filter Model (ii) the speed of convergence of the system and iii) the small

adjustment steady-state.

3.1 Speed of Convergence

The rate at which the coefficients approach their optimum values is called the speed of convergence. As per the previous analytical result obtained, there exists no one quantity that characterizes the speed of convergence, as it depends on the initial coefficient values, the amplitude and correlation statistics of the signals, the filter length L , and the step size $\mu(n)$.

However, we can make several qualitative statements relating the speed of convergence to both the step size and the filter length. All of these results assume that the desired response signal model is reasonable and the errors in the filter coefficients are uniformly distributed across the coefficients on average.

- The speed of convergence increases as the value of the step size is increased, up to step sizes near one-half the maximum value required for stable operation of the system. This result can be obtained from a careful analysis for different input signal types and correlation statistics. For typical signal scenarios, it is observed that the speed of convergence of the excess MSE actually decreases for large enough step size values.
- The maximum possible speed of convergence is limited by the largest step size that can be chosen for stability for moderately correlated input signals. In practice, the actual step size needed for stability of the LMS adaptive filter is smaller than one-half the maximum values when the input signal is moderately correlated. This effect is due to the actual statistical relationships between the current coefficient vector and the signals. Since the convergence speed increases as μ is increased over this allowable step size range, the maximum stable step size provides a practical limit on the speed of convergence of the system.

3.2 Choice of Step Size

Based on the previous result obtained, the speed of convergence as the step is increased. We have seen that the speed of convergence increases as the step size is increased, up to values that are roughly within a factor of 1/2 of the step size stability limits. Thus, if fast convergence is desired, one should choose a large step size according to the limits. However, we also observe that the small adjustment increases as the step size is increased. Therefore, if highly accurate estimates of the filter coefficients are desired, a small step size should be chosen. This classical trade off in convergence speed versus the level of error in steady state dominates the issue of step size selection in many estimation schemes. If the user knows that the relationship between input signal $x(n)$ and desired signal $d(n)$ is linear and time-invariant, then one possible solution to the above trade off is to choose a large step size initially to obtain fast convergence, and then switch to a smaller step size. The point to switch to a smaller step size is roughly when the excess MSE becomes a small fraction (approximately 1/10th) of the minimum MSE of the filter. This method of gear shifting, as it is commonly known, is part of a larger class of time-varying step size methods.

4. EVOLUTIONARY COMPUTATION

Evolutionary algorithms are stochastic search methods that mimic the metaphor of natural biological evolution. Evolutionary algorithms operate on a population of potential solutions applying the principle of survival of the fittest to produce better and better approximations to a solution. At each generation, a new set of approximations is created by the process of selecting individuals according to their level of fitness in the problem domain and breeding them together using operators borrowed from natural genetics. This process leads to the evolution of populations of individuals that are better suited to their environment than the individuals that they were created from, just as in natural adaptation. Evolutionary computation uses computational models of evolutionary processes as key elements in the design and implementation of computer-based problem solving systems. There are a variety of evolutionary computational models that have been proposed and studied (evolutionary algorithms). They share a common conceptual

base of simulating the evolution of individual structures via processes of selection and reproduction. These processes depend on the perceived performance (fitness) of the individual structures as defined by an environment. More precisely, evolutionary algorithms maintain a population of structures that evolve according to rules of selection and other operators such as recombination and mutation. Each individual in the population receives a measure of its fitness in the environment. Selection focuses attention on high fitness individuals, thus exploiting the available fitness information. Recombination and mutation perturb those individuals, providing general heuristics for exploration. Although simplistic from a biologist's viewpoint, these algorithms are sufficiently complex to provide robust and powerful adaptive search mechanisms.

```

Procedure EP; {
    t = 0;
    Initialize population P (t);
    Evaluate P (t);
    Until (done) {
        t = t + 1;
        Parent selection P (t);
        Mutate P (t);
        Evaluate P (t);
        Survive P (t);
    } }
    
```

Evolutionary algorithms (EA) differ substantially from more traditional search and optimization methods. The most significant differences of EA are:

- search a population of points in parallel, not just a single point.
- not require derivative information or other auxiliary knowledge; only the objective function and corresponding fitness levels influence the directions of search.
- use probabilistic transition rules, not deterministic ones.
- generally more straightforward to apply, because no restrictions for the definition of the objective function exist.
- Provide a number of potential solutions to a given problem. The final choice is left to the user. (Thus, in cases where the particular problem does not have one individual solution, for example a family of pareto-optimal solutions), as in the case of multi-objective optimization and scheduling problems, then the evolutionary algorithm is potentially useful for identifying these alternative solutions simultaneously.

4.1. Algorithmic View of EP in VSSLMSEV

Evolutionary programming (EP), developed by Fogel et al. traditionally has used representations that are tailored to the problem domain. For example, in real valued optimization problems, the individuals within the population are real-valued vectors. Similarly, ordered lists are used for traveling salesman problems, and graphs for applications with finite state machines. EP is often used as an optimizer, although it arose from the desire to generate machine intelligence. The outline of the evolutionary programming algorithm is shown below

(i). Generate initial population of μ individuals and set $k=1$, each individual is taken as a pair of real valued vectors $(p_i, \eta_i), \forall i=\{1, \dots, \mu\}$

(ii). Evaluate the fitness score of each individual $(p_i, \eta_i), \forall i=\{1, \dots, \mu\}$ of the population based on the objective function.

(iii). Each parent $(p_i, \eta_i), \forall i=\{1, \dots, \mu\}$ creates a single offspring (p'_i, η'_i) by:

$$p'_i(j) = p_i(j) + \eta_i(j) \cdot N(0,1) \quad \text{-----(3)}$$

$$\eta'_i(j) = \eta_i(j) \exp(\tau' N(0,1) + \tau N_j(0,1)) \quad \text{-----(4)}$$

for $j=1 \dots n$. where $N(0,1)$ is a Gaussian random variable with zero mean and unity standard deviation and $N_j(0,1)$ is generated a new for each value of j . $p_i(j)$, $p'_i(j)$, $\eta_i(j)$, $\eta'_i(j)$ denote the j th component of vector p_i , p'_i , η_i , η'_i respectively. The factor τ and τ' are commonly set to $(\sqrt{2\sqrt{n}})^{-1}$ and $(\sqrt{2n})^{-1}$.

(iv). Calculate the fitness of each offspring $(p_i, \eta_i), \forall i=\{1, \dots, \mu\}$

(v). From the union of parents (p_i, η_i) and offspring $(p_i, \eta_i), \forall i=\{1, \dots, \mu\}$ select the μ individuals have the maximum fitness to be parents of the next generation

5. SIMULATION

To see the performance of EVSSLMS, for channel, number of taps selected for equalizer is 11 taken. Input signal contains total 500 samples generated randomly through uniform distribution shown in fig1. Gaussian noise having zero mean and 0.01 standard deviation added with input signal as shown in fig2. channel characteristics is given by the vector:

[0.05 -0.063 0.088 -0.126 -0.25 0.9047 0.25 0 0.126 0.038 0.088]

5.1 Case1: LMS with fixed step size

A rule of thumb available in number of literatures applied for selecting the step size in order to ensure convergence and good tracking capability in slow varying channel.

$$\Delta = 1 / (5 * (2K + 1) * P_R) \quad \text{-----(5)}$$

Where P_R denotes the received signal plus noise power, which can be estimated from the received signal. In the taken input signal value of Δ from eq(5) is equal to 0.0088. Three different value of fixed step size applied with LMS to see the difference in performance: (a) 0.11 (b) 0.045 (c) 0.0088 and resulting performance has shown in the fig(3). From the result it is clear that it is very difficult to find the optimal step size.

5.2 Case 2: LMS With Adaptive Variable Step Size

To over the problem of optimal step size evolutionary computation as given above applied. For each iteration one generation created and fittest step size in that generation taken for that particular iteration. From previous generation a new generation created by mutation process defined in eq (3) and in eq (4), for next iteration and process will keep continue until all iteration are not completed.

Parameter setting: Initial population is defined by uniform distribution random variable in the range of [0 1] and η_i is taken as 0.000005, $\forall i=\{1, 2, \dots, \mu\}$ and fitness of solution is defined by $f=1/MSE$.

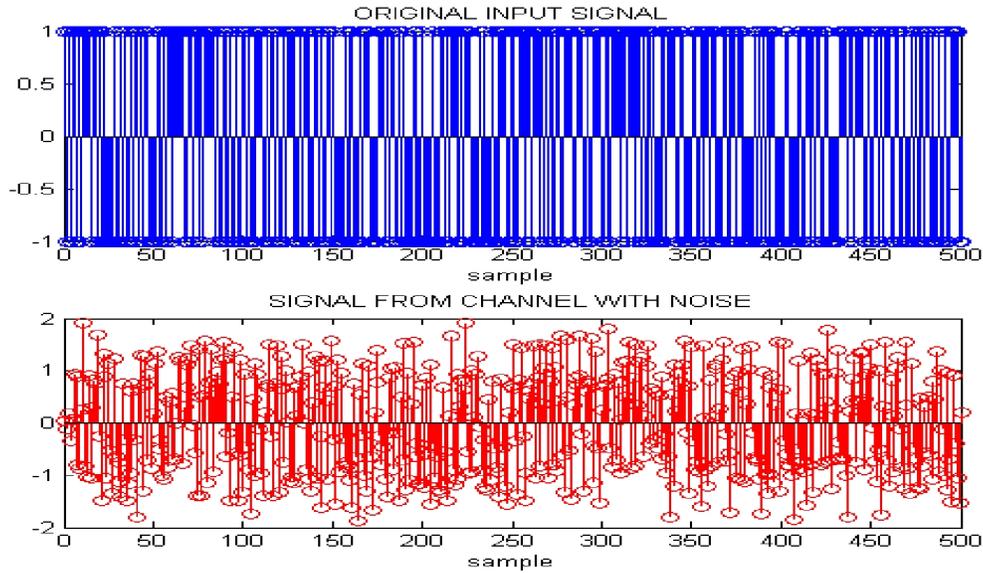


FIGURE 1: generated input signal and signal with noise from channel

To see the performance of EVSSLMS three different population size (i) 10 (ii) 20 (iii)50, taken for same input signal as in case(1) and outcome has shown in fig3.it is clear from result performance is vary very little with higher population size. Plot of adapted vary step size also shown in fig4.the requirement of initial higher value and later lower value of step size easily captured by evolutionary programming.

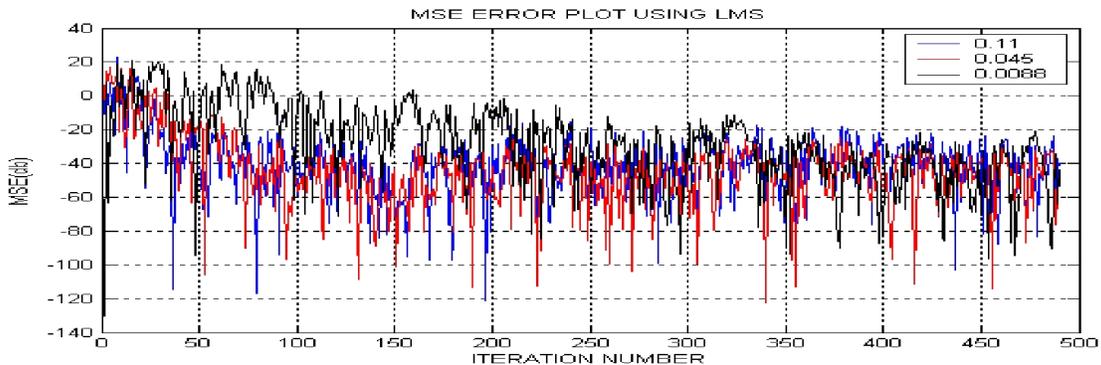


FIGURE 2: fixed step size performance of LMS with step size equal to 0.11, 0.045 and 0.0088

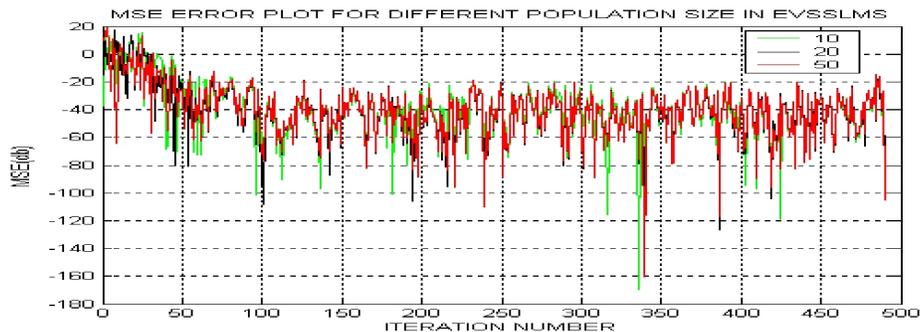


FIGURE 3: performance of EVSSLMS with different population size 10,20 and 50.

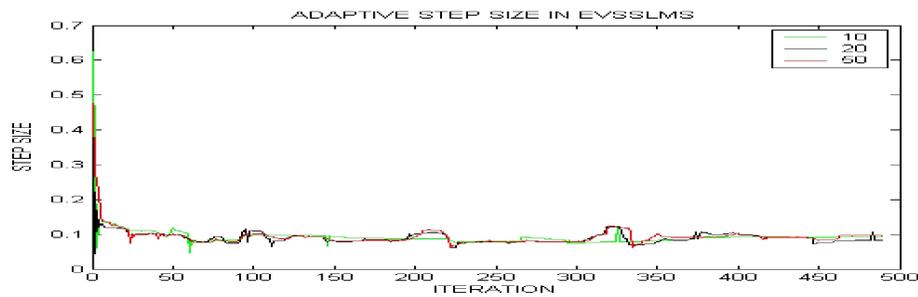


FIGURE 4: Defined step size by EVSSLMS for different population size

6. CONCLUSION

The problem of optimal variable step size integrated with LMS algorithm has solved with the involvement of evolutionary programming. Presented method is robust and does not require the statistical characteristics of input signal as in the case of other existing solutions. Very good convergence and tracking capability can be achieved automatically by presented method. Performance of proposed VSSLMSEV also checked with different population size and it has shown that with less population performance is also equally well and in result higher speed of solution.

Acknowledgement

This research is completed at Manuro Tech Research, Bangalore, India and one of author wanted to say thanks to the management and staff of Jyothi institute of technology, Bangalore, India for their support.

7. REFERENCE

- [1] Solmaz, C.O.; Oruc, O.; Kayran, A.H.; " Optimal step-size LMS equalizer algorithm "Signal Processing and Communications Applications (SIU), 2011 IEEE . April 2011, pp.853 - 856
- [2] Wee-PengAng Farhang-Boroujeny, B," A new class of gradient adaptive step size LMS algorithm" ,Signal Processing, IEEE Transactions on,2001, Volume: 49 Issue: 4 , pp: 805 – 810.
- [3] Krstajic, B. Stankovic, L.J. Uskokovic, Z. "an approach to variable step size LMS algorithm". :Electronics Letters , Aug 2002 ,Volume: 38 Issue: 16,PP: 927 - 928 .
- [4] Saul B. Gelfand, Yongbin Wei, James V. Krogmeier, ," The Stability of Variable Step-Size LMS Algorithms" IEEE transactions on signal processing, Vol. 47, NO. 12, December 1999.
- [5] R. Kwong and E. W. Johnston, "A variable step size LMS algorithm," IEEE Trans. Signal Processing, vol. 40, pp. 1633–1642, July 1992.
- [6] V. J. Mathews and Z. Xie, "A stochastic gradient adaptive filter with gradient adaptive step size," IEEE Trans. Signal Processing, vol. 41, pp. 2075–2087, June 1993.
- [7] Tao Dai; Shahrava, B.," Variable step-size NLMS and affine projection algorithms with variable smoothing factor "Circuits and Systems, 2005. 48th Midwest Symposium on, Aug. 2005, pp: 1530 - 1532 ,Vol. 2.
- [8] J. H. Husoy and M. S. E. Abadi "Unified approach to adaptive filters and their performance" IET Signal Processing, vol. 2 , No. 2 , pp. 97-109, 2008.

- [9] H. C . Shin and A. H. Sayed “Mean square performance of a family of affine projection algorithms” IEEE Trans. Signal Processing, vol. 52, pp. 90-102, 2004.

INSTRUCTIONS TO CONTRIBUTORS

The *International Journal of Signal Processing (SPIJ)* lays emphasis on all aspects of the theory and practice of signal processing (analogue and digital) in new and emerging technologies. It features original research work, review articles, and accounts of practical developments. It is intended for a rapid dissemination of knowledge and experience to engineers and scientists working in the research, development, practical application or design and analysis of signal processing, algorithms and architecture performance analysis (including measurement, modeling, and simulation) of signal processing systems.

As SPIJ is directed as much at the practicing engineer as at the academic researcher, we encourage practicing electronic, electrical, mechanical, systems, sensor, instrumentation, chemical engineers, researchers in advanced control systems and signal processing, applied mathematicians, computer scientists among others, to express their views and ideas on the current trends, challenges, implementation problems and state of the art technologies.

To build its International reputation, we are disseminating the publication information through Google Books, Google Scholar, Directory of Open Access Journals (DOAJ), Open J Gate, ScientificCommons, Docstoc and many more. Our International Editors are working on establishing ISI listing and a good impact factor for SPIJ.

The initial efforts helped to shape the editorial policy and to sharpen the focus of the journal. Starting with volume 6, 2012, SPIJ appears in more focused issues. Besides normal publications, SPIJ intend to organized special issues on more focused topics. Each special issue will have a designated editor (editors) – either member of the editorial board or another recognized specialist in the respective field.

We are open to contributions, proposals for any topic as well as for editors and reviewers. We understand that it is through the effort of volunteers that CSC Journals continues to grow and flourish.

SPIJ LIST OF TOPICS

The realm of Signal Processing: An International Journal (SPIJ) extends, but not limited, to the following:

- Biomedical Signal Processing
- Communication Signal Processing
- Detection and Estimation
- Earth Resources Signal Processing
- Industrial Applications
- Optical Signal Processing
- Radar Signal Processing
- Signal Filtering
- Signal Processing Technology
- Software Developments
- Spectral Analysis
- Stochastic Processes
- Acoustic and Vibration Signal Processing
- Data Processing
- Digital Signal Processing
- Geophysical and Astrophysical Signal Processing
- Multi-dimensional Signal Processing
- Pattern Recognition
- Remote Sensing
- Signal Processing Systems
- Signal Theory
- Sonar Signal Processing
- Speech Processing

CALL FOR PAPERS

Volume: 6 - Issue: 3- August 2012

i. Paper Submission: May 31, 2012 **ii. Author Notification:** July 15, 2012

iii. Issue Publication: August 2012

CONTACT INFORMATION

Computer Science Journals Sdn Bhd

B-5-8 Plaza Mont Kiara, Mont Kiara
50480, Kuala Lumpur, MALAYSIA

Phone: 006 03 6207 1607
006 03 2782 6991

Fax: 006 03 6207 1697

Email: cscpress@cscjournals.org

CSC PUBLISHERS © 2012
COMPUTER SCIENCE JOURNALS SDN BHD
M-3-19, PLAZA DAMAS
SRI HARTAMAS
50480, KUALA LUMPUR
MALAYSIA

PHONE: 006 03 6207 1607
006 03 2782 6991

FAX: 006 03 6207 1697
EMAIL: cscpress@cscjournals.org