iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 4 **Issue:** XII **Month of publication:** December 2016

DOI:

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

Volume 4 Issue XII, December 2016
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

Noise Estimation of Non-Stationary Voiced Speech Using Chirp Model

Omana. B.[1], Bini Palas. P.[2]
[1,2]*Dept.Of ECE, Easwari Engineering College*

Abstract— Speech Enhancement algorithmic models removes or reduces the noise and improve one or more perceptual aspects of noisy speech most notably quality and intelligibility. The speech signal is considered as Non-stationary components, for instance time varying amplitudes and frequencies, which may change quickly over short time intervals by harmonic chirp model. Minimum Variance (MV) filter and Amplitude EStimation (AES) filter are derived and used in removal for noise, enhancement of speech, and parameter estimation to measure the quality of clean speech. In other words, the filtering approaches proposed herein provide full parameterizations of periodic signals through the use of filters based on chirp model. Noise covariance matrix based on MV and AES filters is generated to measure the noise power and reduce the noise present in speech signals. MUSIC algorithm is used to calculate the pseudo-spectral noise power in signals applied. Performance of MV and AES filters and the quality of speech enhanced are measured in terms of Speech Intelligibility Index (SII) and Mutual Information (MI). Simulation on noise corrupted speech signal result in a denoised clean speech with the SII value of 0.8 and MI of 0.17, and the values are compared with the input speech values. Results shows that the SII value and MI value of clean speech are increased by 0.3 and 0.5 from the input speech signal values.
Keywords— Chirp model; Non-stationary speech; Mutual information; Speech Intelligibility Index; MUSIC algorithm; Speech enhancement

I. INTRODUCTION

Speech is a natural way for communication to convey messages. Depending on the character, speech frequency varies from 3 KHz to 4 KHz. Noise shrinking algorithm aims to improve the performance of the communication system where signals are corrupted by noise. Speech enhancement is carried based on two approaches one is the information on desired signal and the other on noise contaminating the signal. Speech enhancement is widely used in applications such as hearing aids and mobile phones. Based on speech production, models of voiced speech is generated which when corrupted by unwanted noise can be processed after it is recorded using microphone in single channel or multichannel model. Speech enhancement algorithm can be applied over speech where filtering techniques by which the noise is removed in frequency domain.

The objective of speech enhancement is to improve the perceptual aspects of speech [2]. Speech is measured in terms of quality that indicates the naturalness in speech and Intelligibility that indicates the correlation relation of input speech with clean enhanced speech [3]. Various methods based on Spectral subtraction [1] and Wiener filtering requires power spectral density (PSD) [8] estimates of the noise signal. The problem of estimating the noise present in speech using PSD in frequency domain using non stationary speech is over helmed by generating noise covariance matrix based on chirp model [4].

Voiced model of speech is generated based on chirp model [7], which allows the frequency of speech signal to vary linearly within each time frame. In the chirp model, Minimum Variance (MV) and Amplitude Estimation Filter (AES) are used in filtering the noise corrupted signal [4, 6]. The novelty of the filters is to pass the undistorted desired signal, minimizing the noise. The filters are derived using chirp model, and their performance is evaluated by means of a noise covariance matrix, which deals with the Non Stationarity of speech in estimating the noise value.

In Section II, the presented model is introduced. In Section III, MV and AES filters along with noise covariance matrix are derived. In Section IV, performance of MV and AES filters are measured in terms of parameters measured. In section V, simulation results are presented. In section VI, analysis and comparison of results among input and clean speech in terms of measured parameters. In Section VII, the presented work is concluded.

II. PROPOSED SYSTEM

A voiced speech is modelled as a signal with a periodic component, consisting of a weighted sum of sinusoidal having frequencies that are integral multiples of fundamental frequency and speech shaped noise signal. Chirp model assumes the Non-Stationarity of speech, whose frequencies over time is varied linearly. The phase values of the speech harmonic is also given as the integer multiple of fundamental frequency.

©IJRASET: All Rights are Reserved

Volume 4 Issue XII, December 2016
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

The instant frequency of i'th harmonic, can be expressed by

$$W_i(n)= l(W_o+kn) \tag{1}$$

For time instant n=0 … N-1, where $W_i(n)$ is the normalized frequency and k is the chirp rate.
Consider the observed signal

$$Y(n) = X(n) + N(n) \tag{2}$$

Where Y(n) is the observed signal, given as the desired signal corrupted by noisy signal N(n).
The signals are assumed as real component and the filters are designed by transforming into complex valued signal. A real signal can be transformed to a complex value by Hermitian transpose, to ease the designing of optimal filters. The present work aims to recover the desired signal by applying a filtering technique with a transposed filter function. Because of the choice of complex valued signals, covariance matrix is derived based on the filters to give the estimated value for each signal.

A. Block diagram

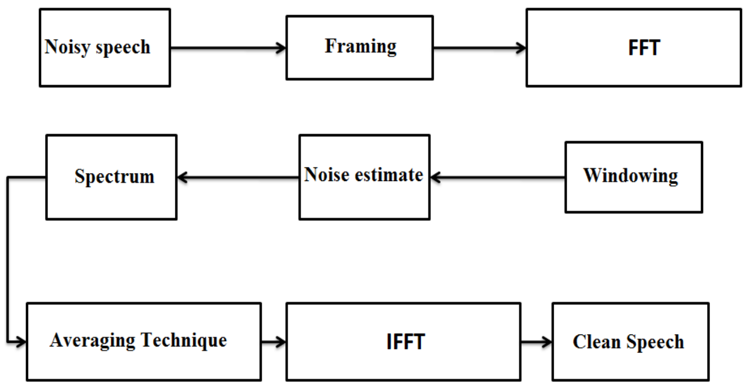


Fig.1 Block diagram of speech Enhancement

III. FILTERING TECHNIQUES

Minimum Variance (MV) and Amplitude Estimation filter (AES) filters are designed based on the principle that it passes the desired signal undistorted while the noise is reduced as much as possible. The MV and AES based filters are derived using the harmonic chirp model .These filters are designed to estimate the noise power for the signals applied.

A. Minimum Variance filter (MV)

Minimum Variance (MV) filter is used in extracting the harmonic signal there by producing undistorted desired signal with reduced noise power. A statistical expectation and variance of a desired signal is calculated, whose value is used in generating the covariance matrix. The covariance matrix is given as the expected mean of a desired signal vector multiplied with Hermitian matrix of a desired value vector component.
MV filter is generalized as the solution for noise power reduction problem based on variance estimate of observed signals. The chirp based MV filter derived over the estimated variance matrix, is given with the zero chirp rate of a signal. The design expression for MV filter is derived as

$$h_L= R_X^{-1}Z\,(Z^H R_X^{-1}Z)^{-1} \tag{3}$$

Where, R_x^{-1} is the inverse covariance matrix of the observed signal and Z is the Fourier sub vectors.

B. Amplitude estimation filter (AES)

Amplitude Estimation filter (AES) filter is used in providing solution to the problem of MSE between the filtered and the input

©IJRASET: All Rights Reserved

signal. As a part of filter design, noise variance matrix is generated by subtracting the desired signal from the matrix value of the observed signal. The design expression for AES filter derived as

$$h_A = Q_V^{-1} Z (Z^H Q_V^{-1} Z^H)^{-1} \qquad (4)$$

Where, Q_V^{-1} is the inverse covariance matrix of the noise signal and Z is the set of Fourier sub vectors.
The difference between the above derived filters is that the MV filter based covariance matrix is given by observed signal and AES filter based covariance matrix is given by the noise signal.

C. Multiple Signal Classification Algorithm

The multiple signal classification (MUSIC) method is a model-based spectral estimation method. The MUSIC method offers higher frequency resolution in the resulting power spectral density (PSD) than the fast Fourier transform (FFT)-based methods. However, the MUSIC method computes PSD magnitudes that generally are not proportional to the true PSD. Therefore, the PSD that the MUSIC method computes traditionally is known as the pseudo PSD.

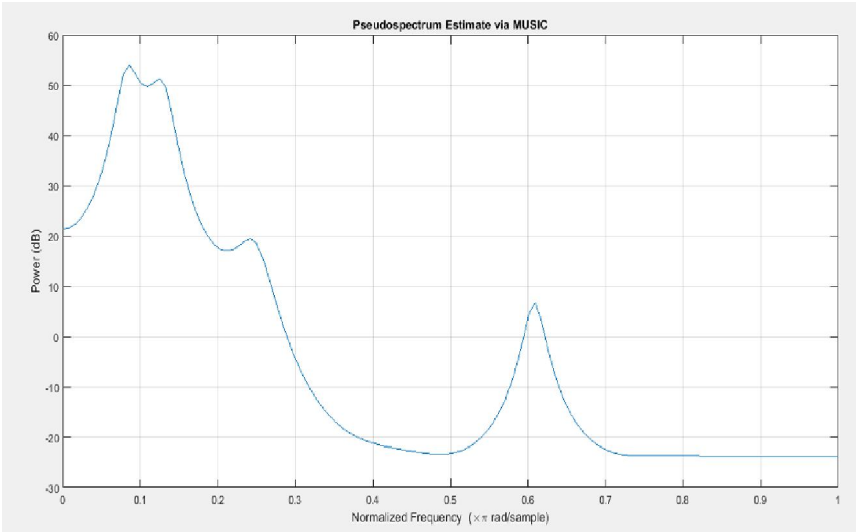


Fig.2 Pseudo Spectral Estimate of Output Speech

IV. EVALUATED PARAMETERS

The performance of the MV and AES filters are compared in terms of Speech Intelligibility Index (SII) and Mutual information (MI). The noise power is calculated for the signals applied in frequency domain.

A. Speech Intelligibility Index (SII)

Speech Intelligibility Index (SII) is the measures of recognizable and understandable amount of speech among the input and clean output. The SII should be within the range 0-1.
Extended Short-Time Objective Intelligibility (ESTOI) is used in measuring SII value. ESTOI predictor depends on the input signal, noisy signal and signal frequency. Short time discrete Fourier Transform (STDFT) is applied over each signal based on the sampling frequency. STDFT is used in process the clean signal in presence of noise to improve its intelligibility in each time frame. The basic principle for estimating SII is partitioning both speech and noise signal into small time frames. ESTOI predictor resample the input and noise, removes the silent frames, in terms of particular frame size. STDFT is applied over the signals to get the spectrum. Filters remove the extra noise present in the input signal for all frames. Normalized row and column values are obtained using this correlated values are computed for input and noise signals. Mean value of correlation among signals is represented as SII. Within each time frame the SII is determined, yielding the speech information available to the listener at that time frame. Next, the SII values of these time frames are averaged, resulting in the SII for that particular condition.

B. Mutual Information (MI)

Mutual Information (MI) is calculated based on entropy of the input and output speech, which returns the output of the SIMI (Speech Intelligibility prediction based on Mutual Information intelligibility predictor). Entropy function depends on amplitude

©IJRASET: All Rights are Reserved

Volume 4 Issue XII, December 2016
ISSN: 2321-9653

value of signal and noise, with SII value. Mean entropy value of clean processed signal value is calculated and compared with input speech.

V. SIMULATION RESULTS

Simulation is carried on observed signal and desired signal corrupted by noise. Spectrogram is used in representation of the recorded signal. Following are the results representing the pseudo spectral value of the desired signal using MUSIC algorithm. Noise power is represented for the input as well desired signals.

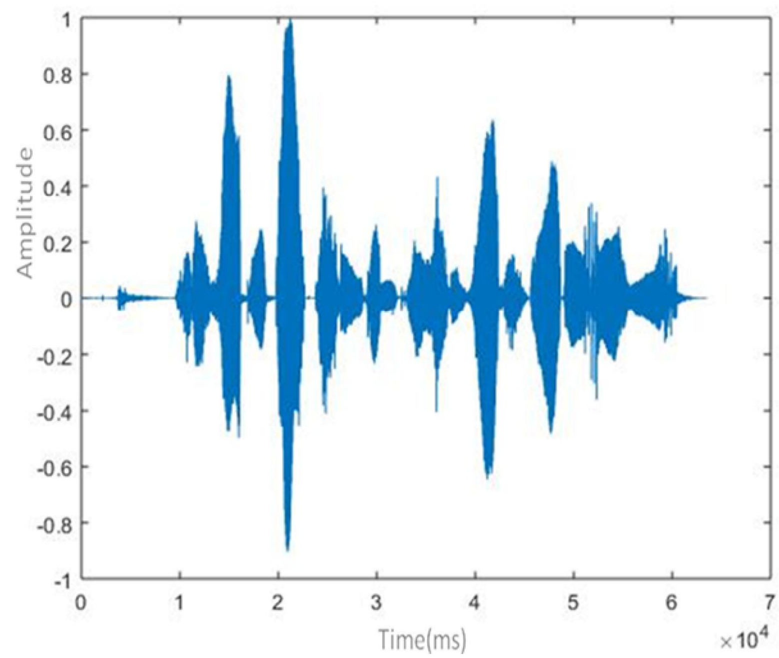


Fig.3 Recorded Input Speech

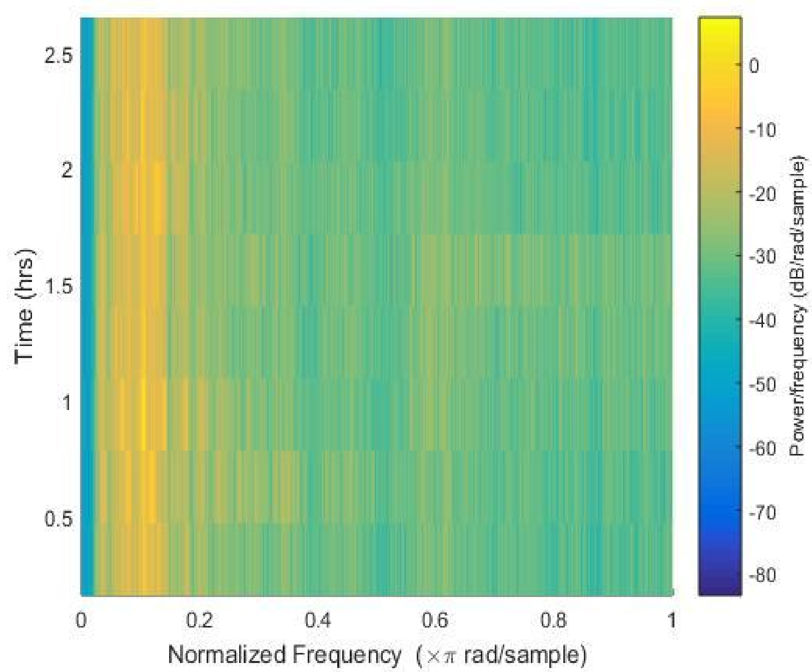


Fig.4 Spectrogram Representation of Noisy Image

©IJRASET: All Rights Reserved

Volume 4 Issue XII, December 2016
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

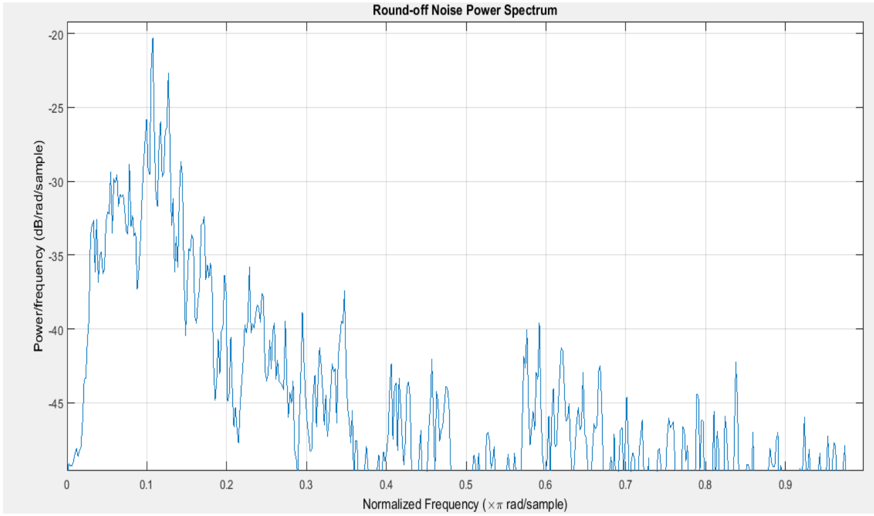


Fig.5 Noise Power Value of a Input Speech

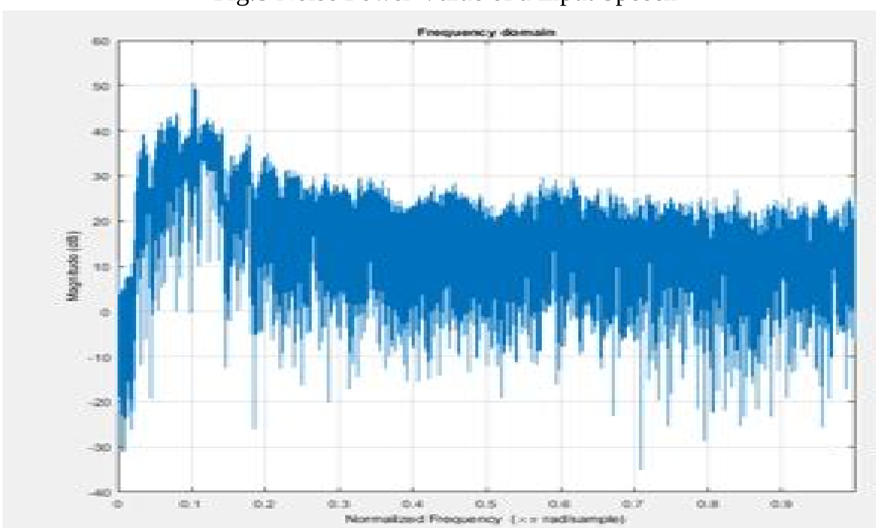


Fig.6 Representation of estimated clean image in Frequency domain

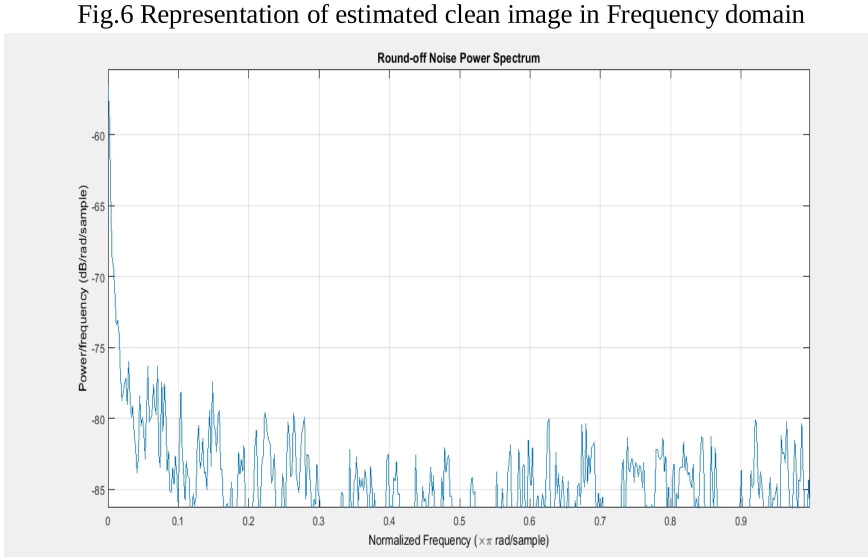


Fig.7 Noise Power Value of Clean Output Speech

©IJRASET: All Rights are Reserved

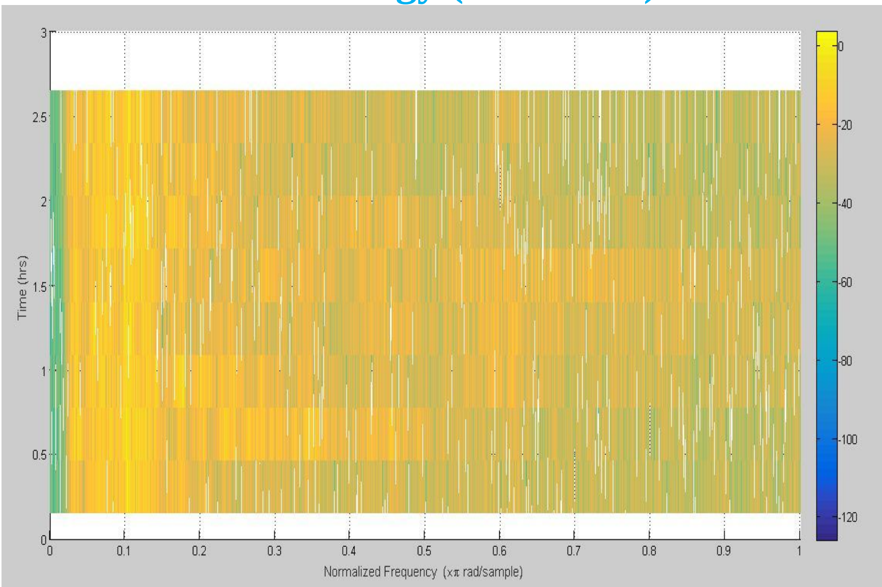


Fig.8 Spectrogram for a Clean Output Speech

VI. RESULT ANALYSIS

The clean speech, being corrupted by speech shaped noise is filtered by MV and AES filters, with the generation of covariance matrix, in order to estimate the signal and noise power of the speech signals before and after filtering operations in terms of Speech Intelligibility (SII) and Mutual Information (MI).The following table represents the comparison of evaluated SII and MI value for input speech and output clean speech signals.

TABLE I
COMPARISON OF SII AND MI VALUES

Signal Type	Speech Intelligibility Index (SII)	Mutual Information(MI)
Input Speech Signal	0.2966	0.1306
Output Clean Speech Signal	0.7927	0.1798

VII.CONCLUSION

The Non- Stationarity of speech is taken into account to increase the performance of enhancement filters. The voiced speech has been described with a harmonic chirp model, using which the frequency at any interval of time can be varied and two filters based on the Minimum Variance (MV) filter and Amplitude EStimation (AES) based filter were presented in order to reduce the noise present in speech signal. The filters are derived and their performance is evaluated in terms of Speech Intelligibility Index (SII) and Mutual Information (MI). MUSIC algorithm is used in estimating the pseudo spectral properties, measures the power value present in speech signal. This method of digital noise reduction can be employed in applications such as hearing aid techniques. The above method of deriving filters with respect to noise covariance matrix, and comparing the estimated parameters of clean speech with input speech signal values in terms of Speech intelligibility index is found to increase from 0.3 to 0.8, along with the Mutual Information which increase from 0.13 to 0.17 which indicates the correlation among input and output speech there by showing an increase in quality of clean speech signal.

REFERENCES

[1] S. Boll, “Suppression of acoustic noise in speech using spectral subtraction,” IEEE Trans. Acoust., Speech, Signal Process., ASSP-27, no. 2, pp. 113–120, Apr. 1979.
[2] Subha S, V.Murugan “Reverberation Suppression of Noisy Speech Signal” International Journal of Innovative Research in Computer and Communication Engineering, Vol.2, Special Issue 1, March 2014.
[3] Rupali V. Mane, Dr.M.T.Kolte “Implementation of Adaptive Filtering Algorithm for Speech Signal on FPGA” International Journal of Innovative Research in

©IJRASET: All Rights Reserved

Electrical, Electronics, Instrumentation and Control Engineering Vol. 2, Issue 3, March 2014.

[4] J. R. Jensen, J. Benesty, M. G. Christensen, and S. H. Jensen, "Enhancement of single-channel periodic signals in the time-domain," IEEE Trans. Audio, Speech, Lang. Process., vol. 20, no. 7, pp. 1948–1963, Sep. 2012.

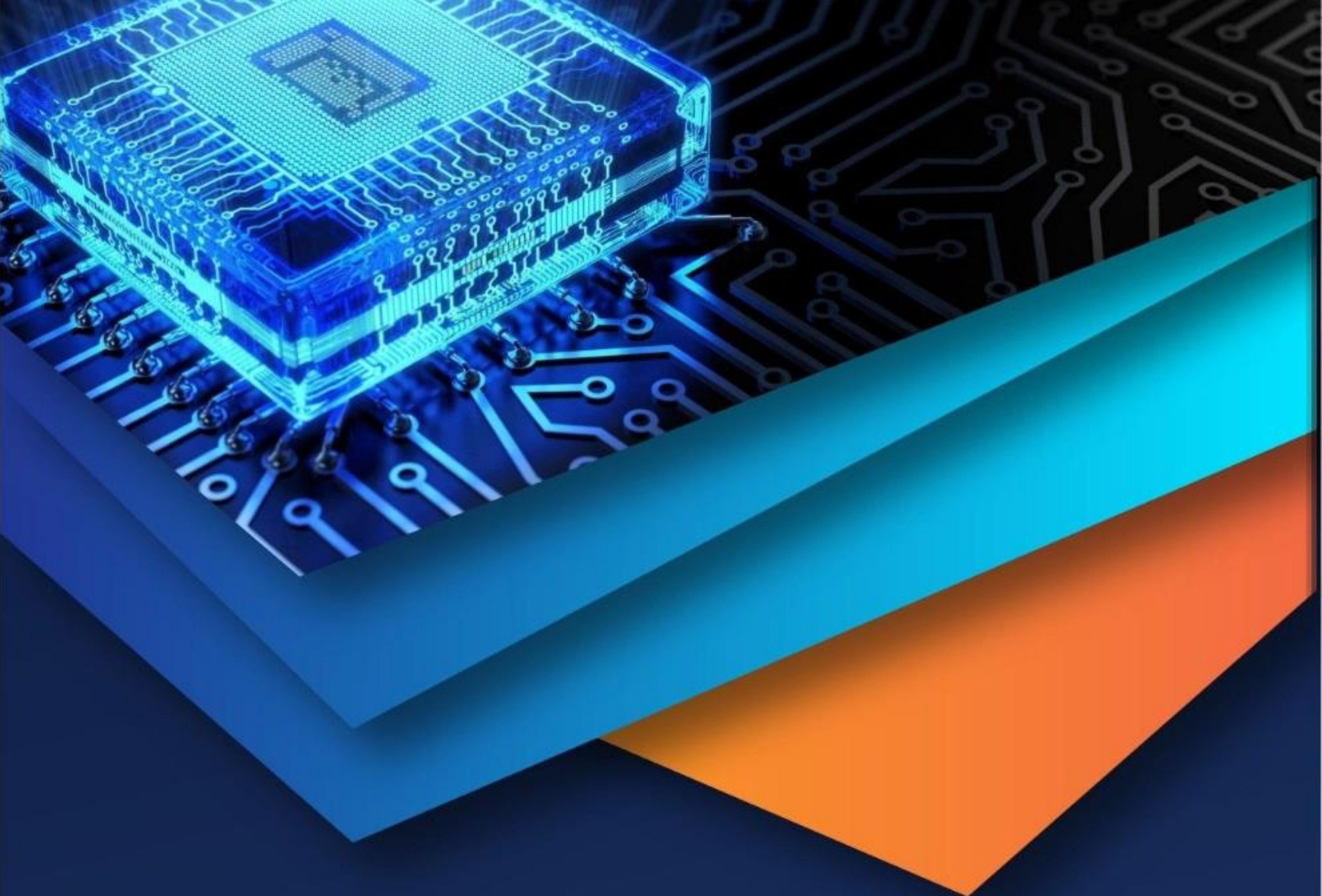
[5] M. G. Christensen and A. Jakobsson, "Optimal filter designs for separating and enhancing periodic signals," IEEE Trans. Signal Process., vol. 58, no. 12, pp. 5969–5983, Dec. 2010.

[6] Y. Pantazis, O. Rosec, and Y. Stylianou, "Chirp rate estimation of speech based on a time-varying quasi-harmonic model," Proc. IEEE Int. Conf. Acoust., Speech, Signal Process., pp. 3985–3988, Apr. 2009.

[7] Y. Doweck, A. Amar, and I. Cohen, "Joint model order selection and parameter estimation of chirps with harmonic components," IEEE Trans. Signal Process, vol. 63, no. 7, pp. 1765–1778, Apr. 2015.

[8] T. Gerkmann and R. C. Hendriks, "Unbiased MMSE-based noise power estimation with low complexity and low tracking delay," IEEE Trans. Audio, Speech, Lang. Process., vol. 20, no. 4, pp. 1383–1393, May 2012.

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)