iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 8 **Issue:** X **Month of publication:** October 2020

DOI: https://doi.org/10.22214/ijraset.2020.31806

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue X Oct 2020- Available at www.ijraset.com

Speech Signal Analysis using Wavelet Domain

Alisha Kamra[1], Kulwinder Singh[2], Sandeep Singh Dhaliwal[3]
[1, 2, 3]Bhai Maha Singh College of Engineering, Sri Muktsar Sahib

Abstract: Different research algorithms and systems are being generated day by day for a variety of DSP applications for detection and processing of image, speech, audio, acoustic signals etc. Traditionally Fourier transform is used for the analysis of signal but it does not provide analysis of non stationary signal like speech signal. Therefore Gabor introduces another transform for analyzing the non stationary signal in which window technique is considered. In Short Time Fourier Transform (STFT), the signal is divided into small enough segments which are assumed as stationary. For this purpose a finite length window function "w" is chosen. This window's width has to be equal to the signal segment where its stationary is valid. The limitation of STFT is that once the size of the window is fixed then it cannot be changed. Therefore, another transform required which is known as Wavelet transforms (WT). Wavelet transform is an efficient method for analyzing non stationary signals i.e. signals in which frequency components changes with time like speech signals. Discrete wavelet transform is used to denoise the speech signal in which White Gaussian noise is added. Different threshold rules (Heursure, Rigrsure, Minimaxi, Sqtwolog) are used to determine the threshold limit of speech signal with different wavelets (db13 and sym13).

Keywords: Short Time Fourier Transform (STFT), Speech Processing, Wavelet Transform, Threshold Rules, Denoising, Signal to Noise Ratio (SNR), Thresholding.

I. INTRODUCTION

Wavelet transform (WT) is the most successful and familiar technique for denoising of signal having non-stationary signals like electrocardiogram (ECG) and electroencephalogram (EEG). WT success depends on optimal arrangement of controlled parameters often set experimentally. Optimality of these parameters combination can be calculated using mean square error (MSE) method [1]. A new method based on hybrid filter fuzzy logic system and genetic optimization algorithm is proposed to remove impulse noise in standard test images. The proposed method HFGOA introduced the three steps. In first stage mean and median filter is used for obtain the uncorrupted image to noisy image. In second stage two filter outputs are used in difference vector value and it is used as input of fuzzy logic system. In third stage the optimal rule is selected by using genetic optimization algorithm [2]. The importance of wavelet transform tool in image denoising is vital. WT is a localized time/space analysis in terms of operating frequency. Operation of telescopic translation is used for gradual multi-scale refining of signal function. Soft as well as hard thresholding is utilized for thresholding of the wavelet coefficients. Wiener filtering can also be utilized as an alternate approach for denoising of an image in wavelet domain. Gaussian white noise is added with grey scale image by utilizing two diverse denoising methods. In comparison of system performance with two suggested methods, Wiener filtering proved as more powerful in wavelet domain [3]. A technique of double wavelet denoising (DWAD) is presented for preserving more features in an original image. Although WT is already being used for the removal of noise, yet it still shows poor performance for signals with low frequency overlap or SNR. Differentiating from single function methods of wavelet denoising, DWAD filters the noisy wavelet coefficients of signal by thresholding functions using two wavelet domains. The DWAD can be applied for 1D signals and it is established that a few wavelet coefficients that are smaller than threshold value can be preserved during the process of noise removal. Besides, DWAD can obtain superior performance while preserving features of original images [4]. Noise, temperature and light affect the quality of image in image processing. Noise implication may introduce during image coding, development process and transmission. Noise elimination has become important field and an eye catching technique in the image processing [5]. Presence of noise in image degrades the quality of image which causes loss of necessary information in image and unsatisfying visual effect. Quality of original image is commonly affected due to noise [6]. To overcome the noise problems different types of filtering technique are used. Improved decision based algorithm is developed based on decision based algorithm for efficient SPN removing. System performance is compared with different filters such as hybrid median filter, median filter and adaptive median filter [7]. An efficient adaptive algorithm is proposed to reduce noise from color and gray scale images. In this method first of a 3*3 window is selected and then the central pixel of the window is considered as processing pixels. If processing pixels identified as uncorrupted then these are not replaced. If it is corrupted pixel then window size is increased as given condition in proposed method [8]. An easy second order pre-filter orthogonal design method is adopted for applying multi-wavelet of higher multiplicities. Noise is reduced using the steins unbiased risk estimator (SURE) with corresponding threshold selection rule.

Higher multiplicities wavelet gives better denoise results [9]. Speech plays an important role in communication system but if it is incorporated by noise then the quality of the information signal is degrades. Therefore it is necessary to remove noise from audio signal to improve the signal quality or to regenerate the signal. Algorithm is proposed for denoising speech signal using different wavelets. Noise is reduced from speech signal by inserted Gaussian noise in audio signal which evaluate the use of different wavelets [10].

In this paper, **Section 1** describes the comprehensive literature review of various papers published by different authors on denoising in speech processing. Various threshold rules are discussed in **Section 2**. Denoising algorithm is discussed in **Section 3.** Results and discussions are described in **Section 4**. Also, conclusions drawn from present research work are provided in **Section 5.**

II. VARIOUS THRESHOLD RULES

The rules of threshold selection are designed on the basis of an underlying model expressed as:

$$y = f(t) + N \tag{1}$$

Here e N is the white Gaussian noise and f (t) is a signal.

There are mainly four types of rules for threshold selection [8].

1) *Rigrsure:* In this rule, Stein's Unbiased Risk Estimate (SURE) principle is used to select the threshold. It provides a risk estimate at a particular value of threshold 't'. Threshold value is given by risk minimization in 't'.
2) *Sqtwolog:* It is a fixed form of threshold for attaining Minimax performance It is generally expressed as sqrt (2* log (length (s))).
3) *Heursure:* A mixture of Rigrsure and Sqtwolog methods is used for selecting the threshold in this method. Due to this SNR is very small and SURE estimate becomes noisy.
4) *Minimaxi:* Minimax principle is utilized for selecting this threshold rule. Minimax performance is yielded by choosing a fixed threshold. Approximately 3 % coefficients are saved using SURE estimate and Minimax threshold.

III. DENOISING ALGORITHM

Denoising algorithm consist of basic three steps

A. Decomposition
B. Threshold
C. Reconstruction.

These steps are explained as follows:

1) *Load Speech Signal:* First step of algorithm is that the Speech signal which is to be used for denoising can be loaded, from free data base of many signals i.e. ECG signal, speech signal, and audio signal.
2) *Add White Gaussian Noise:* When the speech signal is loaded then White Gaussian noise is added in speech signal for the further denoising process.
3) *Decomposition:* Speech signal is decomposed by DWT after passing through HPF and LPF, so that the detailed and approximate coefficients are obtained but before decomposition one should have to select wavelet i.e. in this algorithm different wavelets used are db13, db40, sym13, sym21, where db13, db40 are the daubechie wavelet which has orthogonal property and Sym13, sym21 are symlet wavelet that have similar properties to dbN.
4) *Thresholding:* After decomposition thresholding is chosen i.e. Soft and Hard threshold which is followed by four different threshold rules.
   a) Heursure
   b) Rigrsure
   c) Minimaxi
   d) Sqtwolog
5) *Reconstruction:* Denoised Speech signal can be reconstructed back by IDWT (Inverse Discrete wavelet transform) after passing threshold detailed and approximate coefficients through HPF and LPF by which noise free speech signal can be constructed.
6) *SNR:* After reconstruction, noise-free speech signal is constructed that is evaluated on the basis of SNR. Signal to noise ratio is the measurement of power ratio between a signal (meaningful information) and the background noise (unwanted signal) which determine the quality of signal.

IV. RESULTS AND DISCUSSIONS

Thresholding is a technique by which signal is estimated. For denoising of the signal, it is used to exploit the capacities of signal transformation. In this paper analysis of different threshold rules i.e. Rigrsure, Sqtwolog, Heursure, and Minimaxi is done with three levels of decomposition. Decomposition process consists of two types of filters HPF and LPF whose detailed and approximate coefficients act as threshold for removing the noise present within the speech signal in order to construct noise free speech signal.

A. Analysis of Different thresholding Rules for Wavelet db13 and sym13

For analyzing different thresholding rules, denoised speech signal is decomposed to different levels. In this paper, Heursure, Rigrsure, Minimaxi, and Sqtwolog are analyzed with three levels of decomposition of denoised speech signal**.**

B. Decomposition Level 1

Decomposition values for a Level 1 denoised speech signal is indicated in **Table 1.** Both soft & hard thresholds are used in Level 1 decomposition. Decomposition values of Minimaxi and Sqtwolog threshold rules are same for wavelets db13 and sym13. However, decomposition values for Heursure & Rigrsure threshold rules are different.

Table 1: Decomposition at Level 1

Wavelet Type	Threshold	Minimaxi	Sqtwolog	Heursure	Rigrsure
db13	SOFT	3.1895	4.6034	0.3763	0.3602
	HARD	3.1895	4.6034	0.3310	0.3622
sym13	SOFT	3.1895	4.6034	0.3435	0.3235
	HARD	3.1895	4.6034	0.3232	0.3390

In table 1, results showed that the decomposition values of Minimaxi and Sqtwolog are same for both the wavelets (db13 & sym13). The decomposition value at level 1 for Minimaxi threshold is 3.1895 which are constant for both soft and hard threshold of two wavelets. Similarly the decomposition value of Sqtwolog threshold is 4.6034 that are constant for two wavelets. But the decomposition values of for Heursure and Rigrsure are different for both the wavelets.

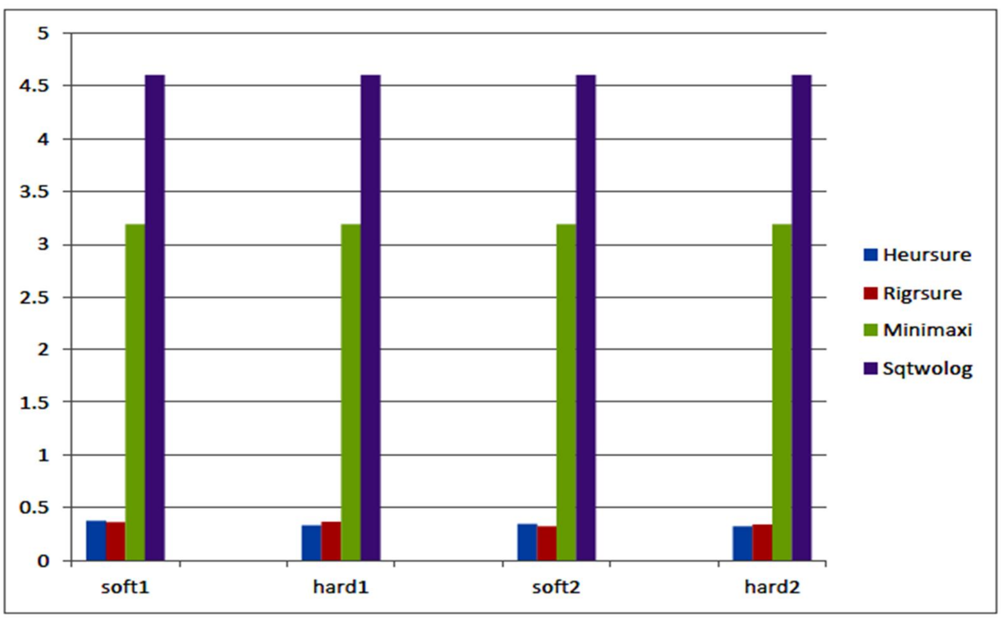


Fig. 1: Decomposition graph at Level 1

Fig. 1 indicates the level 1 decomposition values for a denoised speech signal. Soft1 and hard1 indicates the speech signal decomposition values for db13 wavelet. Whereas, soft2 and hard2 indicates the speech signal decomposition values for sym13 wavelet. Highest decomposition value is achieved using Heursure threshold. As shown in graph Heursure has the maximum value of decomposition (4.6036) which is throughout constant. Similarly Minimaxi threshold rule also has constant decomposition value 3.1897 that remains constant throughout. But the values of decomposition for Heursure and Rigrsure threshold rules are different.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue X Oct 2020- Available at www.ijraset.com

C. Decomposition at Level 2

Decomposition values for a Level 2 denoised speech signal is indicated in **Table 2.** Both soft & hard thresholds are used in Level 2 decomposition. Decomposition values of Minimaxi and Sqtwolog threshold rules are same for Wavelets db13 and sym13. However, for Heursure & Rigrsure rules decomposition values are increased as compared to Level 1.

Table 2: Decomposition at Level 2

Wavelet Type	Threshold	Minimaxi	Sqtwolog	Heursure	Rigrsure
db13	SOFT	3.1895	4.6034	0.5511	0.5125
	HARD	3.1895	4.6034	0.5400	0.5251
sym13	SOFT	3.1895	4.6034	0.4783	0.5114
	HARD	3.1895	4.6034	0.4931	0.5009

In this table the results showed that the decomposition values for Sqtwolog and Minimaxi threshold are same for both the wavelets (db13 & sym13). The decomposition value for Minimaxi threshold is 3.1897 which are throughout constant. Similarly the decomposition value of Sqwtolog threshold is 4.6036 that are also throughout constant. When Heursure threshold rule is used, then the decomposition value of wavelet db13 (soft) increases from 0.3761 to 0.5511. Similarly decomposition value of wavelet db13 (hard) also increases from 0.3310 to 0.5400. When Rigrsure threshold rule is used then the decomposition value of wavelet sym13 (soft) increases from 0.3235 to 0.5114. Similarly decomposition value of wavelet sym13 (hard) also increases from 0.3390 to 0.5009.

Fig. 2 indicates the level 2 decomposition values for a denoised speech signal. Soft1 and hard1 indicates the speech signal decomposition values for db13 wavelet. Whereas, soft2 and hard2 indicates the speech signal decomposition values for sym13 wavelet. Highest decomposition value is achieved using Heursure threshold. The graphs shows that the decomposition values of Rigrsure and Heursure threshold increases as compared to level 1 decomposition values but values of Minimaxi and Sqtwolog thresholds remains constant.

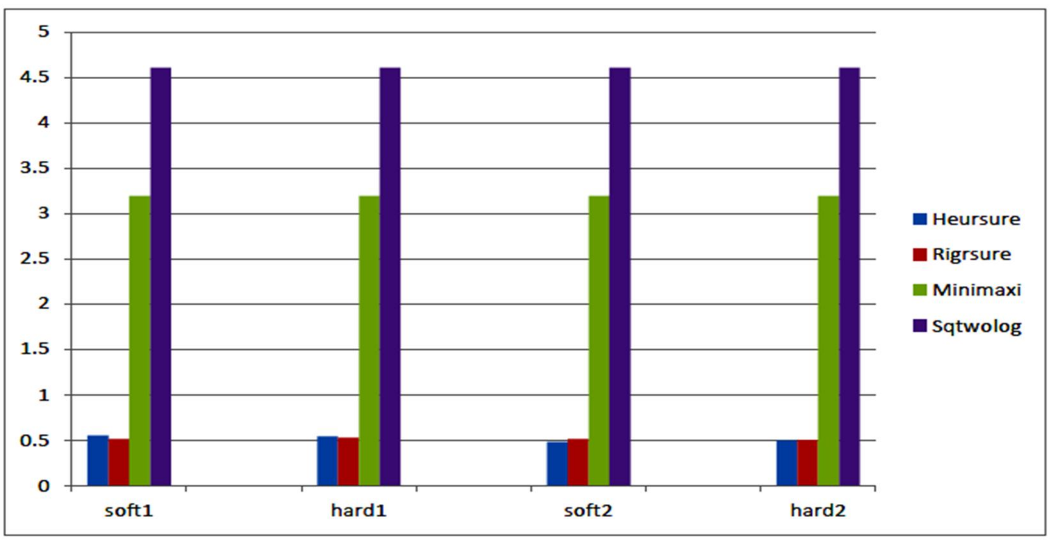


Fig. 2: Decomposition graph at Level 2

The graph shows that the decomposition values of speech signal when Rigrsure, Heursure Minimaxi and Sqwotolog threshold are used.

D. Decomposition at Level 3

Table 3: Decomposition at Level 3

Wavelet Type	Threshold	Minimaxi	Sqtwolog	Heursure	Rigrsure
db13	SOFT	3.1895	4.6034	0.7637	0.8230
	HARD	3.1895	4.6034	0.8061	0.7911
sym13	SOFT	3.1895	4.6034	0.7791	0.7893
	HARD	3.1895	4.6034	0.8071	0.8019

Table 3 indicates Level 3 decomposition values for a denoised speech signal. Heursure threshold provides highest decomposition values that remains constant as compared to Rigrsure and Sqtwolog whose decomposition values increases. The decomposition value of wavelet db13 (soft) is increased from 0.5511 to 0.7637, when Heursure threshold rule is used. Similarly decomposition value of wavelet db13 (hard) also increases from 0.5400 to 0.8061. Similarly decomposition values of Rigrsure threshold also increases but the decomposition values for Minimaxi & Sqtwolog rules remains the same (3.1895 & 4.6034).

Fig. 3 indicates the level 3 decomposition values for a denoised speech signal. Soft1 and hard1 indicates the speech signal decomposition values for db13 wavelet. Whereas, soft2 and hard2 indicates the speech signal decomposition values for sym13 wavelet.

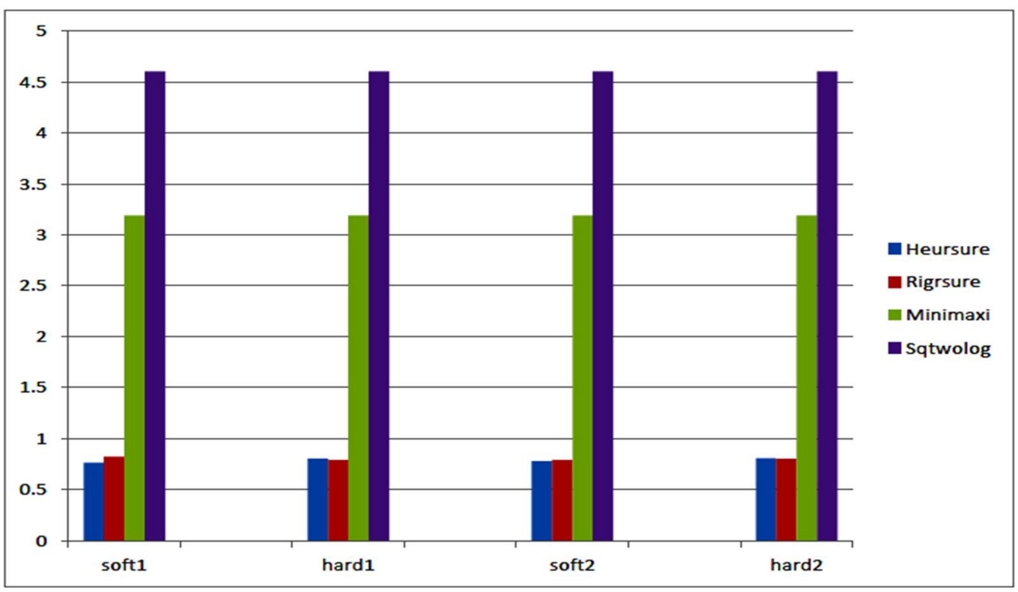


Fig. 3: Decomposition graph at Level 3

The graph shows the level 3 decomposition of denoised speech signal. Different threshold rules i.e. Heurure, Rigrsure, Minimaxi, and Sqwtolog are utilized for calculating level 3 decomposition values. Decomposition values of minimaxi and sqwotlog threshold remains constant as compared to Heursure and Rigrsure threshold whose decomposition values increases from level 1 to level 3.

V. CONCLUSIONS

Denoised speech signal is decomposed from level 1 to level 3. Two different types of wavelets are used for estimating the decomposition values using four threshold selection rules i.e. Heursure, Rigrsure, Minimaxi and Sqtwolog at different levels. Decomposition value of (db13 and sym13) wavelets remains constant when Minimaxi and Sqtwolog threshold rules are used as compared to Heursure and Rigrsure threshold rules whose decomposition values increases from level 1 to level 3. The decomposition values of Minimaxi and Sqtwolog threshold remains constant throughout as compared to Heursure and Rigrsure threshold rules. Therefore, Minimaxi and Sqtwolog threshold provides better results for denoising speech signal at different levels of decomposition.

REFERENCES

[1] Z. A. A. Alyasseri, A. T. Khader, M. A. Al-Betar, A. K. Abasi and S. N. Makhadmeh, "EEG Signals Denoising Using Optimal Wavelet Transform Hybridized With Efficient Metaheuristic Methods," in IEEE Access, vol. 8, pp. 0584-101605, 2020

[2] Selvi, A. Senthil, et al. "De-noising of images from salt and pepper noise using hybrid filter, fuzzy logic noise detector and genetic optimization algorithm (HFGOA)." Multimedia Tools and Applications, Vol. 79, No. 5 (2020), pp. 4115-4131.

[3] W. Fan, W. Xiao and W. Xiao, "Image denoising based on wavelet thresholding and Wiener filtering in the wavelet domain," in The Journal of Engineering, vol. 2019, no. 19, pp. 6012-6015, 2019.

[4] Yongjun Wu, Guangjun Gao and Can Cui, "Improved Wavelet Denoising by Non-Convex Sparse Regularization Under Double Wavelet Domains," IEEE Access, Vol. 7, pp. 30659-30671, 2019.

[5] Ahmad Kamran, Jawad Khan, and Muhammad Salah Ud Din Iqbal. "A comparative study of Different Denoising Techniques in Digital Image Processing." 2019 8th International Conference on Modeling Simulation and Applied Optimization (ICMSAO). IEEE, 2019.

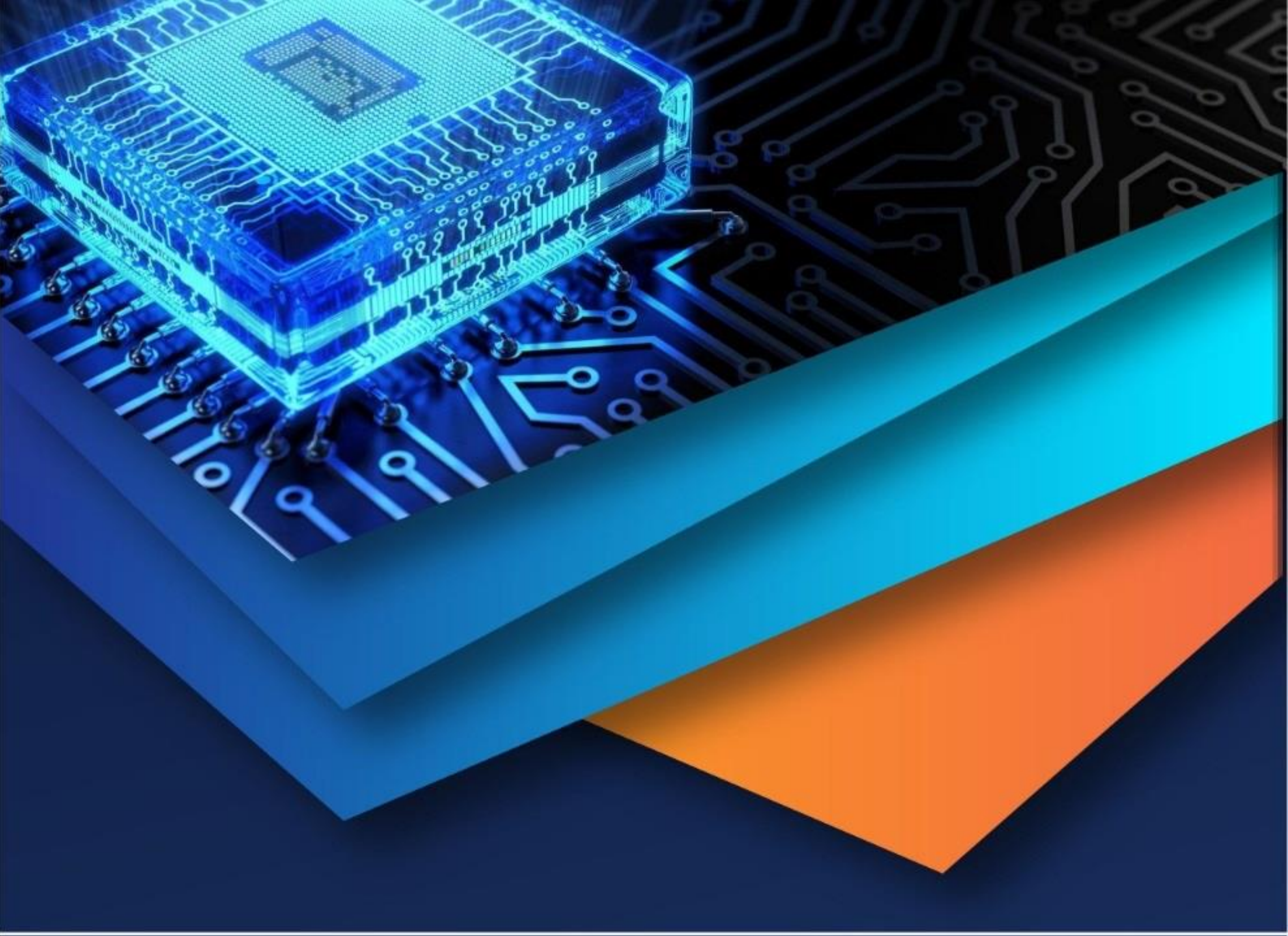
[6] Azzeh Jamil, Bilal Zahran, and Ziad Alqadi. "Salt and Pepper Noise: Effects and Removal." JOIV: International Journal on Informatics Visualization, Vol. 2, No. 4 (2018), pp. 252-256.

[7] Alias, Muhammad Syafiq Alza, Norazlin Ibrahim, and Zalhan Mohd Zin. "Salt and pepper noise removal by using improved decision based algorithm." 2017 IEEE 15th Student Conference on Research and Development (SCOReD). IEEE, 2017.

[8] Sathua Sujaya Kumar, Arabinda Dash, and Aishwarya Rani Behera, "Removal of salt and pepper noise from gray-scale and color images: an adaptive approach." arXiv preprint arXiv, Vol. 1703, No.02 (2017).

[9] Tai-Chiu Hsung, Deniel Pak–Kanglun, and K.C. Ho "Optimizing the multi-wavelet shrinkage Denoising", IEEE transaction on signal processing, Vol. 53, pp. 2, 2014.

[10] Roopali Goel, Ritesh Jain, "Speech Signal Noise Reduction by Wavelets", International Journal of Innovative Technology and Exploring Engineering (IJITEE), Vol. 2, No. 4, pp. 191-193, 2013.



10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY