JOURNAL OF MECHANICS OF CONTINUA AND MATHEMATICAL SCIENCES
www.journalimcms.org

ISSN (Online) : 2454 -7190 Vol.-16, No.-7, July (2021) pp 67-78 ISSN (Print) 0973-8975

# SUPPRESSION OF WHITE NOISE FROM THE MIXTURE OF SPEECH AND IMAGE FOR QUALITY ENHANCEMENT

**Tabassum Feroz[1], Uzma Nawaz[2]**

[1]Department of Electrical Engineering, University of Engineering and Technology Peshawar, Pakistan

[2]Department of Electrical Engineering, University of Engineering and Technology Peshawar (Abbotabad Campus), Pakistan

[1]engineer.tabassumferoz00@gmail.com, [2]Uzma.nawaz@uetpeshawar.edu.pk

Corresponding Author: **Tabassum Feroz**

https://doi.org/10.26782/jmcms.2021.07.00006

(Received: May 6, 2021; Accepted: July 8, 2021)

## Abstract

*This study proposed a correlation analysis of two recent approaches. The FAST ICA technique is used for the separation of the multimodal data (i.e, mixture of audio, noise and image signal) and the minimum mean-square error (MMSE) is used for the removal of white noise from the audio signal. Initially, multimodal data will be formed by combining all three signals (i.e. a mixture of audio, noise and image signals). For creating an ideal situation and for SNR comparisons, separation of the signals will be performed using the Fast ICA technique. ICA, Independent element analysis is a recently developed technique in which the goal is to seek a linear interpretation of non-Gaussian knowledge for the elements to be as statistically free as possible. Such representations record the key structure of the data in several applications, including signal quality and signal separation. ICA learns a linear decay of the data. ICA can find the basic elements and sources included in the data found where traditional methods fail. After the separation of the mixed data, denoising will be performed using the MMSE technique. The main purpose of the MMSE technique is to remove White Noise from the unmixed audio signal which will be further used for overall and segmental SNR comparisons for quality enhancement. Based on the designed algorithms, both of these techniques are real-time data-driven programs. These techniques are explored with standard De-noising methods using several different estimation methods like signal-to-noise ratio (SNR). Experimental results prove that the proposed MMSE technique works well for both noise segmentation and overall consideration of noise distortion signals. These statistical techniques can be used in many applications, such as in different communication systems to eliminate background noise and in channels to reduce channel interference between different applications in speech communications.*

**Keywords**: Minimum Mean Square Error (MMSE), Filtering and Thresholding Techniques, Additive White Gaussian Noise (AWGN), Signal-to-Noise Ratio (SNR), Fast ICA, Whitening, Centering.

## I. Introduction

Quality enhancement includes noise reduction, image enhancement or speech enhancement and echo cancellation. It results in improving the quality and intangibility of the noisy signal. Quality enhancement is one of the prior parameters of the improved signal, particularly in the presence of the white noise which can be removed or suppressed by several proposed methods. The presence of the effect of noise or interference into the speech and image signals directly results in low and poor performance of the desired applications. In other words, quality improvement is defined as the mechanism for improving the quality of speech and image signals by applying the techniques of white noise reduction and removal from any communication systems. MATLAB simulation will be used to implement the De-Noising Mechanism. A real-time noisy speech/audio sample will be recorded into *.WMA format and image/fingerprint will be taken by the scanner in *.PNG. Both the signals will be combined to form multimodal data. For SNR comparison the mixed signals will be separated using the Fast ICA technique will be used. Once the signals are separated then apply the MMSE Algorithm to remove noise from the Unmixed Audio Signal. Since noise is removed, now mix the signals (Image and noise-free audio signal) to obtain again a noise-free multimodal mixed signal. At the end of the process image and speech signals will be separated for further processing. Plot all the signals. Compare the overall and segmental SNRs results before and after the MMSE technique and will be clearly observed that maximum noise has been removed from the signal. In signal processing, noise is a general term for an undesirable signal that may be generated due to interference with the transmission, ISI, processing, or conversion. These signals are random and useless and need to be recovered. Over the years, statisticians have worked over noise control.

In the early ’50s, a problem of point detection method for statistical process control arose and from that year several contributions have been made on noise control. In image processing, as we de-noised a signal using filtration, there is a chance to lose the information of the edge of the image. Edges of an image are a salient feature in a scan line therefore Piecewise constant signal with noise model is important for the digital image. Solving this problem Marr-Hildreth [I] (1980) invented an algorithm for correct detection of edges which involved convolution of the image with Laplacian of Gaussian function. This algorithm has limitations because at curved edges it gives localization error. Perona and Malik (1990) [III] proposed nonlinear diffusion filtration for image processing called anisotropic diffusion. In this method, image noise is removed without removing image contents. Generally, for noise removal from a signal, it removes high-frequency components from the low pass filter and is effective at removing noise when the recovered signal is uniform. Rudin [IV-X] further proposed the ROF model (in 1992) to remove the unwanted signal. Later Chang and Vetterli [XI, XII] gave the concept of wavelet soft thresholding for de-noising the image where the signal is analyzed in form of wavelets. The mathematical functions in which the analysis of data is based on scale or resolution are wavelets. Its threshold is derived in Bayesian framework. It is an adaptive method is widely used in signal processing. Later Bayesian thresholding method was used for working on Mean Square Error. MMSE filtration was invented by Ephraim and Malah and it stands for minimum mean square error filtration. Portilla [XII] presented a model, called BLS-GSM which was based on a Gaussian scale. He presented a Bayesian least squares estimator as an image denoising solution by demonstrating a corrupted image. This method is known as Portilla's method. White Gaussian noise is the thermal noise generated by the movement of electrons in electronic circuits. It has a uniform spectral power density at all frequencies. It is needed to remove such corruption. For such systems, an efficient algorithm was then proposed which worked for the removal of white Gaussian noise from a signal that is discretely sampled and has a broadband spectrum. Speech enhancement is the deletion of noisy speech to improve the quality of speech signals. Noise in a speech signal can be colored noise, impulse noise, transient noise, etc. Over the years many ways to perform this have been proposed. Romain Serizel, Marc Moonen, (2010) [XIII-XIV] introduced an effective noise control system and noise reduction algorithm. It works best if both the active noise control system and noise reduction algorithm are disclosed to provide the system remains causal. To remove noise LMS algorithms are very efficient. For removing noise of distance communication channel Miry, [XVI-XIX] proposed a noise cancellation algorithm that was based on fuzzy and neural networks. Pankaj and Anil, (2012) [XX] did further research on fuzzy and neural algorithms and they declared this as an efficient method. This paper presents an experimental study of the Fast ICA technique and the Minimum Mean-Square Error Technique. While there are several methods for removal of noise and signal filtration picking Fast ICA and MMSE have got significance.

## II. Methodology

Steps adopted to achieve the set objectives are:

### i. Formation of multimodal data

- Pre-processing for Multimodal Data
- Uploading an image file in MATLAB
- Conversion of 3-D to 1-D
- Size Mismatching Problem
- Multimodal Data Formation

### ii. Filtration using Fast ICA technique

- Data Pre-processing for FASTICA
- Centering
- Whitening

### iii. Removal of noise using MMSE technique

- Uploading a file in MATLAB
- Gaussian Noise Merging
- Defining Window
- Framing and noise-silence area detection

**iv. SNR comparison between input signal and enhanced signal**

- Writing the Audio Files in Mat lab
- Equalization of Ranges
- Computation of SNR

**v. Conclusion**

**vi. Suggestion for Future Work**

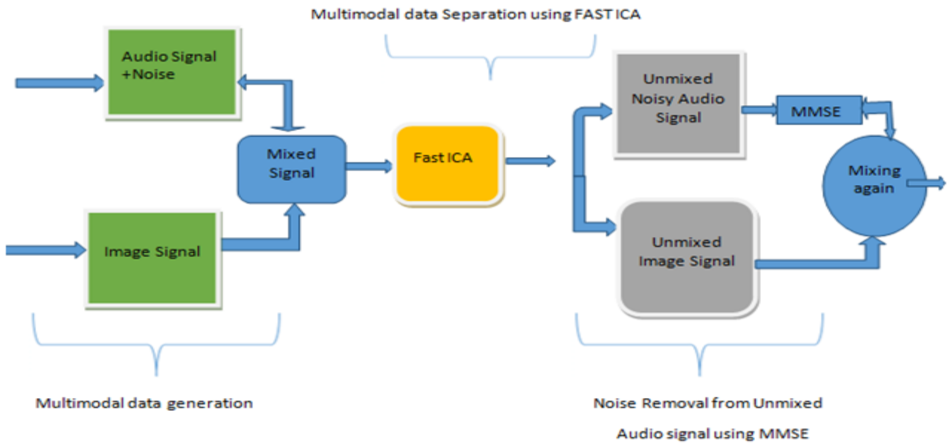


**Fig. 1.** Block diagram of the proposed method

### Multimodal Data Formation

The first stage of this portion is the making of multimodal data which can be done by multiplying the two matrices. The rule of matrix multiplication is followed here that the number of columns in the first matrix must be equal to the number of rows in the second matrix. The first matrix is given the name 'one' which is specifying our audio matrix and the 'two' will point out our image/audio matrix. The two matrices are multiplied under the multiplication rule and as a result, we will get a square matrix which is our multi-modal data. This matrix is then passed through Fast ICA for the separation of data [XXI].

### Filtration using Fast ICA technique

A major problem in signal processing research, and for many, is finding the right representation of multivariate data. The Representation is often desired as a direct modification of the original data. In other words, each representation object is a direct combination of the original variable. Popular methods of change include key component analysis, factor analysis and the pursuit of speculation. Independent component analysis (ICA) is a newly developed approach where the goal is to seek a more accurate picture of non-Gaussian information so that the elements are statistically independent, or as independent as possible [VIII]. Such a presentation appears to record the main data formation in many applications, including feature extraction and signal separation.ICA learns direct decay of data. The ICA tries to find original objects or sources by guessing something simple about their mathematical features. ICA is therefore a data-driven method.

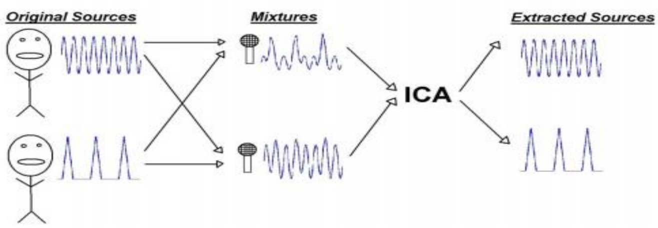


**Fig. 2.** Image showing the basic concept of Fast ICA

**Removal of noise using MMSE technique**

In signal processing, the most common problem is the change in the ratio of the signal to a given type. This problem is called signal recovery. A suitable linear filter known as a minimum mean square error filter will be designed and used for damaged signals. Filters are designed to minimize the mean square error between the desired signal and the available noisy or twisted signal. When the filter is linear, low-resolution (MMSE) filters can be created using matrix expression formats thus improving the quality of the emerging advanced signals [XXII-XXIV].

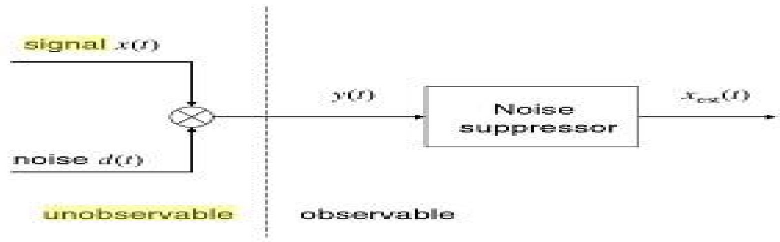


**Fig. 3:** Block Diagram of Noise Removal from Signal

## III. Experimental Results and Discussion

Multimodal Data Formation:

**Fig . 4:** Noise Free Image

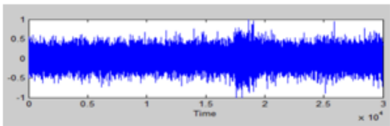


**Fig. 5:** Plot for Noisy Audio Sig

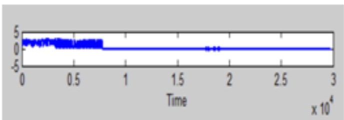


**Fig. 6:** Plot for Noisy Mixed Sig 1

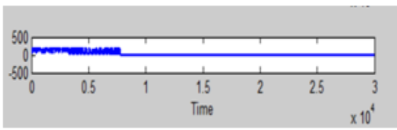


**Fig. 7:** Plot for Noisy Mixed Sig 2

Removal of Noise using MMSE:

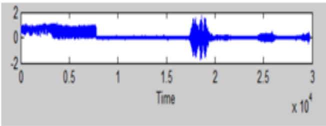


**Fig. 8:** Plot for Noise Free Mixed Sig 1

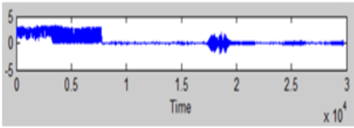


**Fig. 9:** Plot for Noise Free Mixed Sig 2

Filtration using Fast ICA:

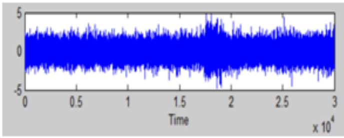


**Fig. 10:** Plot for ICA Unmixed Audio Sig

## Results:

Multimodal Data Formation:

**Fig.11:** Noise Free Image

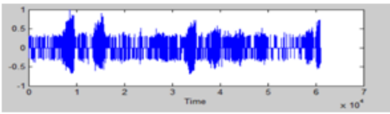


**Fig. 11**: Plot for Noisy Audio Sig

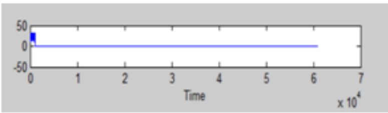


**Fig. 12:** Plot for Noisy Mixed Sig 1

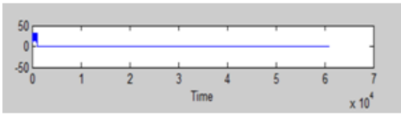


**Fig.13:** Plot for Noisy Mixed Sig 2

Removal of Noise using MMSE:

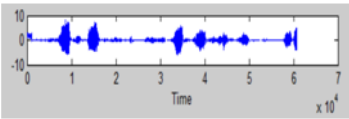


**Fig. 14:** Plot for Noise Free Mixed Sig 1

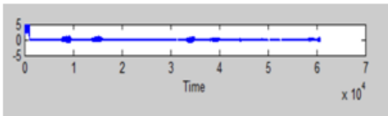


**Fig. 15:** Plot for Noise Free Mixed Sig 2

Filtration using Fast ICA:

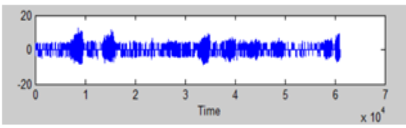


**Fig. 16:** Plot for ICA Unmixed Audio Sig

**SNR Results for Audio Signal**

**Table 1: SNR results for an audio signal before/after MMSE**

| Audio Signal | | | | |
|---|---|---|---|---|
| Samples | Before MMSE | | After MMSE | |
| | Overall SNR (dB) | Segmental SNR (dB) | Overall SNR (dB) | Segmental SNR (dB) |
| 1. | -9.13737 | -9.229253 | 1.187640 | 1.053248 |
| 2. | -1.067405 | -9.621992 | 1.12358 | 1.0 |

Table 1 shows the results of Overall SNRs and Segmental SNRs for the audio signal before/after applying the MMSE technique.

**SNR Results for Mixed Signal**

**Table 2: SNR results for mixed-signal before/after MMSE**

| Mixed Signal | | | | |
|---|---|---|---|---|
| Samples | Before MMSE | | After MMSE | |
| | Overall SNR (dB) | Segmental SNR (dB) | Overall SNR (dB) | Segmental SNR (dB) |
| 1. | 0.39861 | -8.271959 | 7.235209 | 2.857278 |
| 2. | -6.354761 | -3.437476 | 2.297458 | 1.03 |

Table 2 shows the results of Overall SNRs and Segmental SNRs for the mixed-signal before/after applying the MMSE technique.

**SNR comparison between audio signal and mixed-signal**

In signal processing, the quality of the signal greatly depends on the SNR value. The small the SNR, the weak the signal will be. The greater the SNR, the better the signal will be. SNR can have both positive and negative values.

Furthermore, SNR is a measurement parameter used in the scientific and engineering fields that compares the magnitude of the desired signal with the background noise. In other words, SNR is the ratio of signal power to the noise power, and its unit of output is typically decibels (dB). Segmental SNR is considered an excellent cognitive model because it evaluates measured noise with energy in each basic speech segment.

**Overall SNR comparison between audio signal and mixed-signal before MMSE**

**Table 3: Overall SNR comparison before MMSE**

| | **Overall SNR (dB) comparison before MMSE** | |
|---|---|---|
| | Sample 1 | Sample 2 |
| **Audio Signal** | -9.1374 | -1.06741 |
| **Mixed Signal** | 0.39861 | -6.35476 |

**Overall SNR Comparison between audio signal and mixed-signal after MMSE**

**Table 4: Overall SNR comparison after MMSE**

| | **Overall SNR (dB) comparison after MMSE** | |
|---|---|---|
| | Sample 1 | Sample 2 |
| **Audio Signal** | 1.38764 | 1.12358 |
| **Mixed Signal** | 7.235209 | 2.29746 |

TThe Histograms in the below-given Figures represent the SNR of an audio signal along with its mixed-signal in decibels. We have considered the SNR of mixed signals as a reference for their respective audio signals. Since mixed signals are a combination of 2 independent signals so they have a high level of noise present in them. We have obtained negative SNRs for mixed signals before the application of the MMSE technique. Similarly, we can also see that the SNRs of the mixed samples are positive. This is because maximum noise has been removed from them through the MMSE technique. The SNR's of mixed-signal are positive than those of audio signals due to the presence of noise. In the same way, the SNR values for both the audio and mixed signals before the MMSE techniques are negative which shows that lots of noise are present in both the signals. However, after the application of the MMSE technique, we have observed from the Figures that maximum noise has been removed from either signal as shown below.

Overall SNR Comparison:

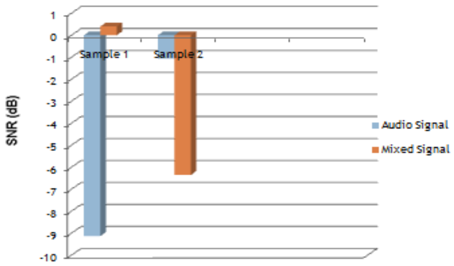


**Fig. 17:** Histogram showing Overall SNR Comparison between Audio Signal & Mixed Signal before MMSE

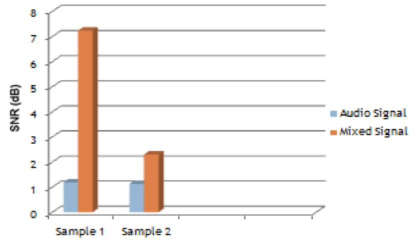


**Fig. 18**: Histogram showing Overall SNR Comparison between Audio Signal & Mixed Signal after MMSE

**Fig. 17** shows the comparison of Overall SNR between audio and mixed-signal before applying the MMSE technique. The Histogram shows negative values of SNR which elaborates that a lot of noise is present in both the signals. Similarly,

**Fig. 18** shows the comparison of Overall SNRs between the audio mixed signals after applying the MMSE technique. The Histogram shows positive values of SNR which elaborates that after the application of the MMSE technique, maximum noise has been removed from the mixed-signal which results in the improved and enhanced values of SNR as shown in the above Histogram. This also shows that the SNR for the mixed signals is high as compared to that of the audio signals.

**Segmental SNR comparison between audio signal and mixed-signal before MMSE**

**Table 5: Segmental SNR comparison before MMSE**

| | Segmental SNR (dB) comparison before MMSE | |
|---|---|---|
| | Sample 1 | Sample 2 |
| **Audio Signal** | -9.22925 | -9.62199 |
| **Mixed Signal** | -8.27196 | -3.437476 |

**Segmental SNR comparison between audio signal and mixed-signal after MMSE**

**Table 6: Segmental SNR comparison after MMSE**

| | Segmental SNR (dB) comparison before MMSE | |
|---|---|---|
| | Sample 1 | Sample 2 |
| **Audio Signal** | 1.053248 | 1 |
| **Mixed Signal** | 2.857278 | 1.03 |

Segmental SNR Comparison:

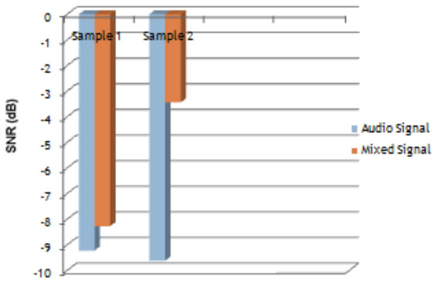


**Fig. 19:** Histogram showing Segmental SNR Comparison between Audio Signal & Mixed Signal before MMSE

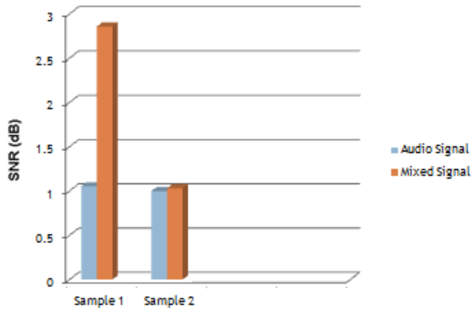


**Fig. 20**: Histogram showing Segmental SNR Comparison between Audio Signal & Mixed Signal after MMSE

**Fig. 19** shows the comparison of Overall SNR between audio and mixed-signal before applying the MMSE technique. The Histogram shows negative values of SNR which elaborates that a lot of noise is present in both the signals. Similarly,

**Fig. 20** shows the comparison of Overall SNRs between the audio mixed signals after applying the MMSE technique. The Histogram shows positive values of SNR which elaborates that after the application of the MMSE technique, maximum noise has been removed from the mixed-signal which results in the improved and enhanced values of SNR as shown in the above Histogram. This also shows that the SNR for the mixed signals is high as compared to that of the audio signals.

## IV. Conclusion

The following conclusion has been summarised from the current experimental study.

i. The values of SNR have improved for both the audio signal as well as for the mixed-signal after applying the MMSE technique.

ii. About 50-60% of the noise has been removed from the noisy samples. However, the results show that fast ICA does not work successfully for the mixture of speech and image in terms of separation.

iii. It has also been concluded from the results that fast ICA works properly in the separation of the only speech signal from the mixture data.

iv. IIn this whole project the application of the signal processing field has been explored through the use of Mat lab.

v. An example of this is the activities running within the human brain sometimes involve both image and audio which can be separated by this technique.

vi. The said technique has already been proven the best technique among all the proposed technique.

vii. These mathematical methods have many applications in automatic speech recognition (ASR) systems to remove background noise, in channels to reduce-channel interference between various applications in speech communications.

## V. Suggestion for Future Work

i. The readers are highly encouraged to further vast their experience in this particular workout. Certain expectations of multi-modal data separation are still needed to be deft and explored further.

ii. The artifacts which are present in the audio signal due to data compression are needed to be tackled.

iii. Fast ICA can be made more durable in the future.

iv. One can further enhance the accuracy of the Communication System.

v. Moreover, one can also improve and enhance the quality of dealing with multiple signals simultaneously.

**Conflict of Interest:**

There was no relevant conflict of interest regarding this paper.

## References

I. A Rapid Match Algorithm for Continuous Speech Recognition by Laurence S. Gillick and Robert Roth Dragon Systems, Inc. 90 Bridge St. Newton MA. 02158.

II. A. Ravi, Leela Satyanarayana. V., : GREY WOLF OPTIMIZATION WITH WAVELET SCHEME FOR SAR IMAGES DENOISING. *J. Mech. Cont.& Math. Sci., Vol.-14, No.-5, September-October (2019) pp 558-570*. DOI : 10.26782/jmcms.2019.10.00040.

III. Blind Separation of Sources: A Non-Linear Neural Algorithm Gilles BUREL (1992).

IV. D. Marr; E. Hildreth Proceedings of the Royal Society of London. Series B, Biological Sciences, Vol. 207, No. 1167. (Feb. 29, 1980), pp. 187-217

V. Elad m. 2002 ways to improve bilateral filter IEEE trans on image processing 11(10) pg 1141-1151.

VI. E. Abreu, M. Lightstone, S. K. Mitra, and K. Arakawa, A new efficient approach for the removal of white noise from highly corrupted images, IEEE Transactions on Image Processing, vol. 5, no. 6, pp. 10121025, 1996.

VII. Gonzalez, Woods, and Richard E. Woods. ”Eddins, Digital Image Processing Using MATLAB.” Third New Jersey: Prentice Hall (2004).

VIII. International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-2, Issue-5, June 2013.

IX. Independent component analysis, A new concept Pierre Comon (1994).

X. Image Analysis and Mathematical Morphology by Jean Serra, ISBN 0-12-637240-3 (1982).

XI. International Journal on Recent and Innovation Trends in Computing and Communication, ISSN: 2321-8169 Volume: 2 Issue: 6 1657 – 1661.

XII. J Portilla, V Strela, M Wainwright, and E P Simon celli, “Image denoising using scale mixtures of Gaussians in the wavelet domain,” IEEE Trans. Image Processing., In Press. 2003.

XIII. L. Rudin, S. Osher, E. Fatemi Nonlinear total variation based noise removal algorithms Physica D, 60 (1992), pp. 259–268.

XIV. Michael, Weeks. ”Digital Signal Processing Using MATLAB .”Pearsonpublications, ISBN81-297-0272-X 2.13 (2011): 15-16.

XV. Motwani M. C., Gadiya M. C., Motwani R. C. and Jr. Harris F. C. (2004), Survey of image denoising techniques, Proceedings of Global Signal Processing Expo andConference (GSPx 04),, Santa Clara, CA, USA.

XVI. Mallat, S., 1989. A theory for multiresolution signal decomposition: The wavelet representation. IEEE transactions on pattern analysis and machine intelligence 11, 674–693.

XVII. P. Mrazek, J. Weickert, and A. Bruhn. Geometric Properties from Incomplete Data, chapter On Robust Estimation and Smoothing with Spatial and Tonal Kernels. Springer, 2006.

XVIII. P. Prandoni, “Optimal Segmentation Techniques for Piecewise Stationary Signals,” Ph.D. thesis, Ecole Polytechnique Federale de Lausanne, Switzerland (1999).

XIX. PietroPerona and Jitendra Malik (July 1990). "Scale-space and edge detection using anisotropic diffusion" IEEE Transactions on Pattern Analysis and Machine Intelligence pg 629–639.

XX. Rafael C. Gonzalez, Richard E. Woods, Steven L. Eddins. Digital Image Processing Using MATLAB [M]. McGraw-Hill Press, 2011.

XXI. Scott E Umbaugh, Computer Vision and Image Processing, Prentice Hall PTR, New Jersey, 1998.

XXII. S. H. Chung and R. A. Kennedy, "Forward-Backward Nonlinear Filtering Signals from Noise, "Journal of Neuroscience Methods, vol. 40, pp. 71-86, 1991.

XXIII. Utpal Barman, Ridip Dev Choudhury. : ‘Prediction of Soil pH using Smartphone based Digital Image Processing and Prediction Algorithm’. J.Mech.Cont.& Math. Sci., Vol.-14, No.2, March-April (2019) pp 226-249. DOI : 10.26782/jmcms.2019.04.00019.

XXIV. Wayne Niblack, An Introduction to Image Processing, Prentice-Hall, NewJersey, 1986.filter, IEEE Xplore PDCAT pp.826-828.