Audio Noise Removal – Literature review

[1]Srashti Badkul[2]Brajlata Chourasiya

Department of Electronics & Communication, GGITS, Jabalpur

ABSTRACT: In this paper different audio denoising techniques are discussed. Most of the audio denoising techniques reduce Gaussian white noise from audio signals. Diagonal estimation techniques and non-diagonal estimation techniques are discussed. Different audio denoising techniques and noises are shown through the taxonomy.

KEYWORDS: audio denoising, diagonal estimation, Gaussian noise, musical noise, non-diagonal estimation, thresholding, White noise

_____

I.   INTRODUCTION

Audio is corrupted by different types of noise during acquisition of audio. The aim of noise removal from audio is to attenuate the noise without modifying the original signal. Various applications of audio denoising are music and speech restoration. Diagonal estimation techniques and non-diagonal estimation techniques are two types of audio denoising techniques. To attenuate the noise from audio signals diagonal timefrequency audio denoising algorithms process each spectrogram coefficient independently. The drawback of these algorithms are they have a limited performance, denoised signal contains musical noise, denoised sound is contaminated and the audio perception is degraded due to the superposition of musical noise. To overcome these drawbacks non-diagonal estimation techniques are required [6], [7], [11].

II.   AUDIO DENOISING - RELATED WORKS

The problem of extracting the desired sound signal, corrupted by "additive white Gaussian noise (AWGN) has been of interest to many researchers for practical as well as theoretical reasons" [1]. The removal of AWGN is difficult as it persists at all the frequencies in the signal. Two of the popular methods of denoising the musical instrument signals are;

(i)   Those based on adaptive filter algorithms [11][19];
(ii)   Those based on wavelet based algorithms [14].

Spectral audio denoising methods usually make use of the magnitudes of a time-frequency representation of the signalis discussed in paper [1]. However, ifthe time-frequency frame consists of quadrature pairs of atoms (asin the short-time Fourier transform), then the phases of the coefficients also follow a predictable pattern, for which simple models areviable. In this paper, we propose a scheme that takes into accountthe phase information of the signals for the audio denoising problem. The scheme requires to minimize a cost function composed ofa diagonally weighted quadrature data term and a fused-lasso typepenalty. We formulate the problem as a saddle point search problemand propose an algorithm that numerically finds the solution. Basedon the optimality conditions of the problem, we present a guidelineon how to select the parameters of the problem. We discuss the performance and the influence of the parameters through experiments

Wavelet based algorithm for audio denoising is discussed in paper [2]. The authors focused on audio signals corrupted with white noise. White noise is especially hard to remove because it is located in all frequencies. The authors used Discrete Wavelet Transform (DWT) to transform noisy audio signal in wavelet domain. It was assumed that signal is represented by high amplitude DWT coefficients and noise is represented by low amplitude coefficients. To get audio signal with less noise, thresholding of coefficients are used and they are transformed back to time domain. The authors proposed modified universal thresholding of coefficients which results with better audio signal. Objective Degree Grade (ODG) was main criterion for evaluation of experimental results. The authors have also compared ODG with Mean Square Error (MSE) which is widespread used for estimating signal quality. Results show that MSE shows little enhancement or even loss while ODG and also informal listening tests prove significant enhancement of signal quality. This denoising algorithm worked better for lower noise signals but for higher noise signals higher threshold must be set, but except noise part of original signal is also removed by it causing audible artifacts in denoised signal.

In paper [3], block attenuation methods that were initially applied in orthogonal wavelet signal representations [4] is investigated by authors. Block size as well as thresholding level in redundant timefrequency signal representations is studied by authors and they found that the remaining noise artifacts in restored signals is eliminated by block attenuation and provides a

good approximation of the attenuation with oracle. A connection between the block attenuation and the decision-directed a priori SNR estimator of Ephraim and Malah is studied by authors. An adaptive block technique based on the dyadic CART algorithm [4, 5] is introduced by authors. The experiments show that the remaining noise artifacts is eliminated and transients of signals are preserved by the proposed method better than the methods which use short-time Fourier do [3]. The experiments were performed on speech signals sampled at 11 kHz. These speech signals were corrupted by white Gaussian noise. The performance of block attenuation is good when compared with the performance of other methods such as Adaptive Block Attenuation with Complex Wavelets, Hard Thresholding with Complex Wavelets, Ephraim and Malah decision-directed a priori SNR estimator + Wiener with Complex Wavelets / Short-Time Fourier. A number of experiments were performed on various music signals also.

The performance of adaptive block attenuation is good when compared with the performance of conventional thresholding operators. Sharper note transitions is obtained than the estimate with short-time Fourier. However, denoising using short-time Fourier performs better than the wavelet counterpart for the stationary parts when high pitch is involved because in high frequency bands short-time Fourier has higher frequency resolution than wavelet representation.In paper [8], denoising problem is considered from the viewpoint of sparse atomic representation. The authors proposed a general framework of time-frequency soft thresholding which encompasses and connects well known shrinkage operators as special cases. Convergence of the corresponding algorithms is numerically evaluated and their performance in denoising real life audio signals is compared to the results of similar existing approaches. The novel approach is competitive with respect to signal to noise ratio and improves the state of the art in terms of perceptual criteria. From the denoising point of view the neighborhood weighting could be considered as non diagonal estimation. Musical noise naturally arising in diagonal estimation is reduced by these approaches.

In paper [9], significant improvements in audio denoising is obtained by exploiting the persistence properties of signals. In this contribution, a novel denoising operator based on neighborhood smoothed, Wiener filter like shrinkage is derived. The purpose of the paper is concerning the operator design and derives a novel audio denoising operator, the persistent empirical Wiener estimate, which fuses recent developments in the field of structured sparsity with the properties of empirical Wiener filtering. According to a given performance criterion a rationale for adaptive threshold selection is proposed. Compared to the optimal thresholds a plain linear model depending on the level of the noise achieves minor performance differences. A simple method for estimating this noise level in case it is unknown is proposed. The proposed operators perform competitively compared to the state of the art, while being much more computationally efficient and robust to minor perturbation of the noise level. The method presented in [10] is based on the Singular Value Decomposition (SVD) of the frame matrix representing the signal in the Overlap Add decomposition. Both the singular values and the singular vectors of the representation are modified to perform denoising. For the former a tapering model is used and for the latter a nonlinear PDE method is used. The aim of the proposed technique is to reduce additive random noise which has corrupted the signal. To test this method the authors performed tests on a variety of sounds from speech and music after corrupting them with additive gaussian noise. The authors used the sampling rate 16 kHz for speech and 44.1 kHz for music. The authors compared their method with SavitzkyGolay filter in terms of MSE and SNR. Results show that performance of their method is good in reducing noise from signal.

In paper [11], the method used is non-diagonal in which block parameters are automatically adjusted to the nature of the audio signal. This is done by minimizing a Stein estimator of the risk which is calculated analytically from noisy signal values. Block thresholding method is used to eliminate musical noise. This block thresholding method performs attenuation of time-frequency coefficients after grouping the time frequency coefficients in blocks. In diagonal time-frequency audio denoising algorithms there is lack of time frequency regularity because of which it create isolated time frequency structures. This isolated time frequency structures are interpreted as musical noise. Block thresholding is used for audio time frequency denoising which regularizes the estimate and musical noise is reduced efficiently. In paper [12], Adaptive time-frequency Block Thresholding procedure using discrete wavelet transform is used to reduce the noise from the audio signal and to achieve better SNR of the audio signal. For audio signal denoising discrete-wavelet transforms based algorithms are used. For denoising both soft thresholding and hard thresholding are used. In the paper the authors compared the results of soft thresholding and hard thresholding. Results showed that performance of soft thresholding is better than performance of hard thresholding.

The optimality property discussed in [12] is "related to the ability of the wavelet basis to capture most of the signal energy in a small number of coefficients". The standard wavelet basis has been shown to be optimal in this regard for representing signals that have local singularities. In this wavelet based denoising methods for speech enhancements have been discussed. In the first method, traditional "spatially selective noise filtration technique" is proposed and second method is based on "Undecimated Discrete Wavelet Transform". These methods can be used for edge detection satisfactorily. Wavelet methods have been used for uni-dimensional (1D) and two dimensional (2D) signal analysis, producing and analyzing irregular signals [8]. "The fundamental idea behind wavelets is to analyze according to scale.

Matching Pursuit (MP) is a greedy algorithm that iteratively builds a sparse signal representation. An analysis of Matching Pursuit in the context of audio denoising is presented in the work [13]. The algorithm is interpreted as a simple shrinkage approach, the authors identified factors critical to its success and several approaches to improve its performance and robustness is proposed. The authors have presented experimental results on a wide range of audio signals and shown that the method is able to yield results that are competitive with other audio denoising approaches. The authors introduced a new audio denoising approach called Greedy Time-Frequency Shrinkage (GTFS) that is able to produce competitive denoising results in terms of standard performance metrics, Signal to Noise Ratio (SNR) and Perceptual Evaluation of Audio Quality (PEAQ). The

authors focused on the removal of uncorrelated Gaussian white noise from music and speech signals. The various audio denoising techniques are shown in the taxonomy of figure 1 where MMSE-LSA is Minimum Mean Square Error Log Spectral Amplitude Estimation algorithm. Different noises are shown in the taxonomy of figure 2.

Figure 1. Audio denoising taxonomy

Figure 2. Noise taxonomy
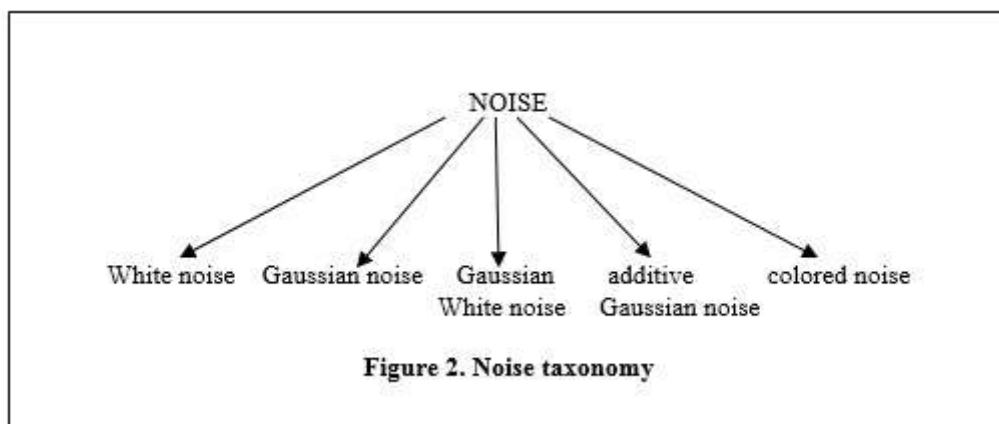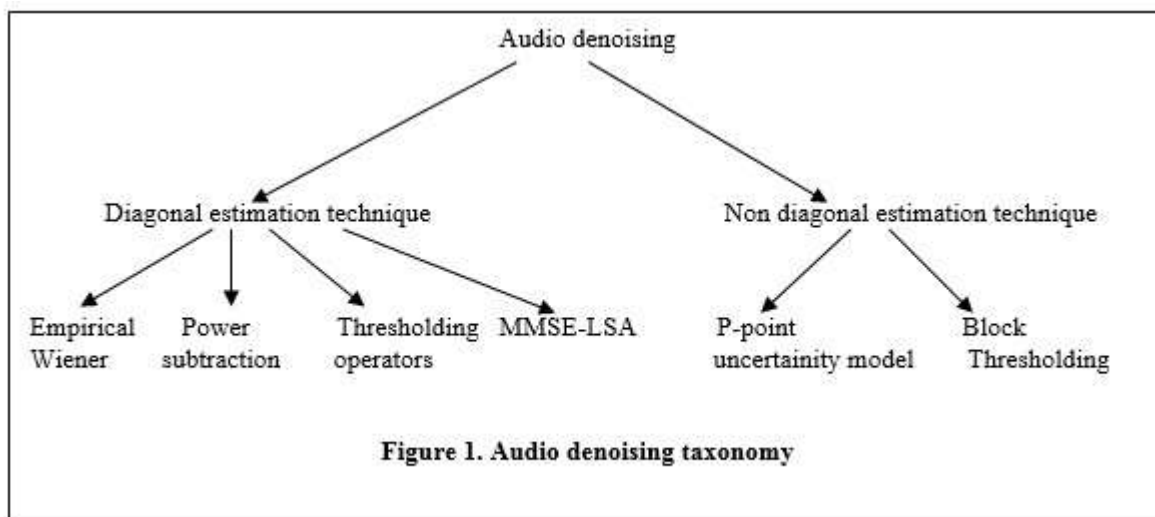
Performance comparison of Block Thresholding [7], Block Thresholding (BT) with soft thresholding wavelet [12], Block Thresholding (BT) with hard thresholding wavelet [12], Minimum Mean Square Error Log Spectral Amplitude Estimation algorithm (MMSE-LSA) [7], Minimum Mean Square Error Log Spectral Amplitude Estimation algorithm by using Decision Direct method (MMSE-LSA-DD) [7] of Mozart signal for different SNR values is shown in the below table 1.

Table 1. Performance comparison

Signal and SNR	Block Thresholding	BT with soft thresholding wavelet	BT with hard thresholding wavelet	MMSE-LSA	LSA-DD
Mozart 5dB	14.90	11.57	9.84	7.625	7.625
Mozart 10dB	18.31	15.02	12.76	12.625	-
Mozart 15dB	22.03	18.75	15.97	17.727	-
Mozart 20dB	25.14	22.64	19.48	22.825	-
Mozart 25dB	30.29	27.43	23.62	28.785	-

From the above performance comparison table 1, we are concluding that the block thresholding technique is more efficient than other listed techniques because signal to noise ratio value of block thresholding technique is high.

III.   CONCLUSIONS

Audio is corrupted by different types of noise during acquisition of audio. The process of removing such noise from audio signals is audio denoising. In this paper different audio denoising techniques are discussed. From the survey, we are concluding that the non-diagonal estimation techniques are efficient compared to diagonal estimation techniques as they avoid producing musical noise.

REFERENCES

[1]   Ilker Bayram, "Employing phase information for audio denoising", IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP), 2014.

[2]   Matko Saric, Luki Bilicic and Hrvoje Dujmic, "White Noise Reduction of Audio Signal using Wavelets Transform with Modified Universal Threshold", University of Split, R. Boskovica b. b HR, volume 21000, 2005.

[3]   Guoshen Yu, Emmanuel Bacry and Stephane Mallat, "Audio Signal Denoising with Complex Wavelets and Adaptive Block Attenuation", IEEE International Conference on Acoustics, Speech and Signal Processing, Volume 3, 2007.

[4]   T. Cai and B.W. Silverman, "Incorporation information on neighboring coefficients into wavelet estimation", Sankhya, 63, 127-148 , 2001

[5]   L. Breiman, J. Friedman, R. Olshen, and C.J Stone, Classification and Regression Trees, Belmont, CA: Wadsworth, 1983.

[6]   D. L. Donoho, "CART and best-ortho-basis: a connection", Ann. Statist. 25 1870–1911.

[7]   Guoshen Yu ,Stephane Mallat, Emmanuel Bacry, "Audio Denoising by Time-Frequency Block Thresholding", IEEE Transactions on Signal Processing, Vol. 56, No. 5, May 2008.

[8]   Lalitha kumari, Karunakar Reddy, Hari Krishna and Venkata Subash, "Time-Frequency Block Thresholding Approach for Audio Denoising", International Journal of Advances in Science and Technology, Vol. 2, No. 5, 2011.

[9]   Kai Siedenburg and Monika Dorfler, "Audio Denoising by Generalized Time-Frequency Thresholding", Audio Engineering Society Conference: 45th International Conference: Applications of Time-Frequency Processing in Audio, 2012.

[10]   Kai Siedenburg, "Persistent Empirical Wiener Estimation with Adaptive Threshold Selection for Audio Denoising", Proceedings of the 9th Sound and Music Computing Conference, pages 426 - 433, 2012.

[11]   George Baravdish, Gianpaolo Evangelista, Olof Svensson and Faten Sofya, "PDE-SVD Based Audio Denoising", Proceedings of 5th International Symposium on Communications, Control and Signal Processing, IEEE, 2012.

[12]   K.P. Obulesu and P. Uday Kumar, "Implementation of Time Frequency Block Thresholding Algorithm in Audio Noise Reduction", International Journal of Science, Engineering and Technology Research (IJSETR), Volume 2, Issue 7, July 2013.

[13]   S. S. Joshi and Dr. S. M. Mukane, "Comparative Analysis of Thresholding Techniques using Discrete Wavelet Transform", International Journal of Electronics Communication and Computer Engineering, Volume 5, Issue (4) July, 2014.

[14]   Gautam Bhattacharya and Philippe Depalle, "Sparse Denoising of Audio by Greedy Time-Frequency Shrinkage", 2014 IEEE International Conference on Acoutics, Speech and Signal Processing (ICASSP), 2014.

First Author:

Ms.Srashti Badkul, M. Tech Student (4th Sem.) from Gyan Ganga Institute of Technology and Sciences, Jabalpur (M.P.). She has published 1 research paper in international journal..

Second Author:

Mrs.Brajlata Chourasiya Asst.Professor in Gyan Ganga Institute of Technology and Sciences, Jabalpur (M.P.). I have published 10 research papers in various national and international journals and conferences. Out of these 10, four are from International Journals.