ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

An Improved Speech Denoising and Enhancement Based on Radon Transform and Local Ridgelet Transform

P. P. S. Subhashini
Associate Professor,
Department of ECE,
RVR & JC College of Engineering,
Chandramoulipuram, Chowdavaram,
Guntur, Andrapradesh, India.

J. Ravindranadh
Associate Professor,
Department of ECE,
RVR & JC College of Engineering,
Chandramoulipuram, Chowdavaram,
Guntur, Andrapradesh, India.

ABSTRACT:

Recent years, De-noising or enhancing the quality of speech signals will be a challenging task in speech processing applications. Previously there are so many algorithms have been proposed to de-noise the speech signals. Here in this thesis we had done an experiment on denoising of various speech signals by introducing a new transformation technique called “Ridgelet”, which an extension for wavelet and curvelet. Ridgelet analysis can also be done in a similar way to wavelet analysis in the Radon domain as it is translates singularities along lines into point singularities. Simulation results had shown that the proposed technique performed better than the conventional techniques, and the quality of de-noised speech is measured in terms of Signal to Noise Ratio and Mean Square Error.

I.INTRODUCTION:

Speech is simple and a possible way for the human to convey the information from one place to another with emotion or for one person to others. In many applications these speech processing systems plays a vital role such as speech recognition, voice communication. The performance of speech/voice processing systems will degrade as the increment in background noise which in results reduces the intelligibility of speech. Therefore, someone needs to reduce or eliminate the unwanted information from the speech, and it is necessary to develop such enhancement algorithms in order to eliminate and improve the quality of speech, so that even in noisy environments, these systems can work efficiently. In early stage of research, many speech processing algorithms were developed to suppress background noise and improve the perceptual quality and intelligibility of speech. These methods were developed with some perceptual or statistical constraints placed on the noise and speech signals.

They involve simple signal processing and reduce only some extent of noise in the signal. And also, some of the speech signal is distorted during the enhancement process. Hence, these techniques have a tradeoff between the amount of noise removal and speech distortions which was introduced because of processing of the speech signal. A detailed review on these techniques is presented and the references there in. In the last two decades, a substantial progress has been made in the field of speech processing.

However, spatial filtering approaches like mean filtering or average filtering, Savitzky filtering, Median filtering, bilateral filter and Wiener filters had been suffered with loosing edges information. All the filters that have been mentioned above were good at denoise of speech but they will provide only low frequency content it doesn’t preserve the high frequency information. In order to overcome this issue Non Local mean approach has been introduced. More recently, noise reduction techniques based on the “NON-LOCAL MEANS (NLM). It is a data-driven diffusion mechanism that was introduced by Buades et al. in [1]. It has been proved that it’s a simple and powerful method for speech denoising. In this, a given coefficient or sample is denoised using a weighted average of other coefficients or samples in the (noisy) speech signal. In particular, given a noisy signaln_i, and the denoised signal d ̂=(d_i) ̂ at coefficient i is computed by using the formula.

$$\hat{d_i} = \frac{\Sigma_j w_{ij} n_j}{\Sigma_j w_{ij}} \qquad (1)$$

Where w_ij is some weight assigned to coefficienti and j. The sum in (1) is ideally performed to whole signal to denoise the noisy signal. NLM at large noise levels will not give accurate results because the computation of weights of coefficients will be different for some neibourhood samples which looks like same.

Most of the standard algorithm used to denoise the noisy signal and perform the individual filtering process. Denoise generally reduce the noise level but the signal is over smoothed due to losses like edges or lines. In the recent years there has been a fair amount of research on wavelet thresholding and threshold section for signal denoising [3], because wavelet provides an appropriate basis for separating noisy signal from the signal. Wavelet transform is good at energy compaction, the small coefficient are more likely due to noise and large coefficient due to important signal feature [8]. These small coefficients can be thresholded without affecting the significant features of the signal.

The wavelet transform (WT) is a powerful tool of signal processing for its multi resolutional possibilities [11]. Unlike the Fourier transform, the wavelet transform is suitable for application to non-stationary signals with transitory phenomena, where frequency response varies in time [2]. The wavelet coefficient represents a measure of similarity in the frequency content between a signal and a chosen wavelet function [2]. These coefficient are computed as a convolution of the signal and the scaled wavelet function, which can be interpreted as a dilated band pass filter because of its band pass like spectrum [5]. By wavelet analysis from a signal at high scales, extracted global information called approximations, and at two scales, extracted fine information called details.

The discrete wavelet transform (DWT) requires less space utilizing the space saving coding based on the fact that wavelet families are orthogonal or biorthogonal bases, and thus do not produce redundant analysis. The discrete wavelet transform corresponds to its continuous version sampled usually on a dyadic grid, which means that the scales and translations are power of two [5]. Thresholding is a simple non-linear technique, which operates on one wavelet coefficient at a time. In its most basic form, each coefficient is thresholded by comparing against threshold. If the coefficient is smaller than threshold then it set to be zero; otherwise it is kept or modified. We replace the small noisy coefficient by zero and inverse wavelet transform on the result may lead to reconstruction with the essential signal characteristics and with less noise. Since the work of Donoho and Johnstone [4], [9], [10], there has been much research on finding thresholds, however few are specifically designed for signals.

Experiments show the effectiveness of the new technique both in terms of signal-to-noise ratio (on simulated speech signals). In the proposed approach we had introduced a ridgelet transform based noise reduction and enhancement of speech signals, which is an extension for wavelet transforms. The ridgelet transform will give the fine and smooth curve instead of discrete curve. The experimental results will be compared with the existing techniques and will be shown that the proposed scheme will perform superior to the conventional methods.

II.RELATED WORK:

In this section we discussed various spatial filters and their performance when a noisy input will be given to them. Here in this section we had explained about each filter in detail.

a.Median Filter:

The median filter is a nonlinear digital spatial filtering technique, often used to noise removal from images. Median filtering is very widely used in speech processing applications. The main idea of the median filter is to run through the speech entry by sample, replacing each sample with the median of neighboring samples. The pattern of neighbours is called the “window”, which slides, sample by sample, over the entire speech.

b.Bilateral Filter:

The bilateral filter is a nonlinear filter which does the spatial averaging without smoothing edges information. Because of this feature it has been shown that it’s an effective speech denoising algorithm. Bilateral filter is presented by Tomasi and Manduchi in 1998. The concept of the bilateral filter was also presented in [2] as the SUSAN filter and in [3] as the neighbourhood filter. It is mentionable that the Beltrami flow algorithm is considered as the theoretical origin of the bilateral filter [4] [5] [6], which produce a spectrum of speech enhancing algorithms ranging from the linear diffusion to the non-linear flows. The bilateral filter takes a weighted sum of the speech samples in a local neighbourhood; the weights depend on both the spatial distance and the intensity length. In this way, edges are preserved well while noise is eliminated out.

c.Classic Non- Local Means:

It is a data-driven diffusion mechanism that was introduced by Buades et al. in [1]. It has been proved that it's a simple and powerful method for digital image denoising. In this, a given pixel is denoised using a weighted average of other pixels in the (noisy) image. In particular, given a noisy imagen_i, and the denoised image d ̂=(d_i ̂) at pixel i is computed by using the formula.

$$\tilde{d_i} = \frac{\sum_j w_{ij} n_j}{\sum_j w_{ij}} \qquad (2)$$

Where w_ij is some weight assigned to pixeli and j. The sum in (1) is ideally performed to whole image to denoise the noisy image. NLM at large noise levels will not give accurate results because the computation of weights of pixels will be different for some neibourhood pixels which look like same.

d.WAVELET THRESHOLDING:

Signal denoising using the DWT consists of the three successive procedures, namely, signal decomposition, thresholding of the DWT coefficients, and signal reconstruction. Firstly, we carry out the wavelet analysis of a noisy signal up to a chosen level N. Secondly, we perform thresholding [11] of the detail coefficients from level 1 to N. Lastly, and we synthesize the signal using the altered detail coefficients from level 1 to N and approximation coefficients of level N. However, it is generally impossible to remove all the noise without corrupting the signal. As for thresholding, we can settle either a level-dependent threshold vector of length N or a global threshold of a constant value for all levels.
According to D. Donoho's method, the threshold estimate d for denoising with an orthonormal basis is given by where the noise is Gaussian with standard deviation s of the DWT coefficients and L is the number of samples or pixels of the processed signal or image. This estimation concept is used by MATLAB. From another point of view, thresholding can be either soft or hard. Hard thresholding zeroes out all the signal values smaller than d. Soft thresholding does the same thing, and apart from that, subtracts d from the values larger than d. . In contrast to hard thresholding, soft thresholding causes no discontinuities in the resulting signal. In MATLAB, by default, soft thresholding is used for denoising and hard thresholding for compression.

II.PROPOSED RIDGELET TRANSFORM:

A two dimensional continuous ridgelet transform in R^2 is given in [11]. If a Smooth univariate function ψ: R R is taken with sufficient decay and also satisfy the admissibility condition given by Eq. (1) [12].

$$\int \frac{|\psi|^2}{|\xi|^2} d\xi < \infty \qquad (3)$$

Then, ψ has a vanishing mean $\int \psi(t)dt = 0$ and a special normalization about $\psi$is chosen so that

$$\int_0^\infty |\psi(\xi)|^2 \xi^{-2} d\xi = 1$$

For each scale (a) > 0, each position (b) ∈ $R$and each orientation(θ) ∈ [0,2π), the bivariate Ridgelet $\psi_{a,b,\theta}$ is defined as [11, 12]

$$\psi_{a,b,\theta}(x) = a^{-1/2}\psi(x_1 cos\theta + x_2 \sin\theta - b)/a \; (4)$$

A Ridgelet is constant along lines $x_1 cos\theta + x_2 \sin\theta = const.$ If these ridges are transverse, it becomes a wavelet. Given an integrable bivariate function f(x), Ridgelet coefficients are defined as [12]

$$R_f(a,b,\theta) = \int \overline{\psi_{a,b,\theta}(x)} f(x) dx \qquad (5)$$

And the extract reconstruction is obtained by using the above equation is

$$f(x) = \int_o^{2\pi}\int_{-\infty}^{\infty}\int_0^{\infty} R_f(a,b,\theta)\, \psi_{a,b,\theta}(x) \frac{da}{a^3} db \frac{d\theta}{4\pi} \; (6)$$

Ridgelet analysis can also be done similar to wavelet analysis in the Radon domain as it is translates singularities along lines into point singularities, for which the wavelet transform is known to provide a sparse representation [36]. Radon transform of an object f is defined as the collection of lines integral indexed by(θ,t) ε[0, 2π) ×R :

$$Rf(\theta,t) = \int f(x_1,x_2)\partial(x_1 \cos\theta + x_2 \sin\theta - t) dx_1 dx_2 \qquad (7)$$

In eq.7 '∂' is the Dirac distribution. It means that the Ridgelet transform is the application of 1D-DWT to the slices of the Radon transform where the angular variable θ is remained constant and t is having varying nature [12]. Therefore, for computing continuous Ridgelet transform, initially Radon transform is determined, then 1-D wavelet transform is applied to the slices. A detailed discussion on Ridgelet transforms is given in [11, 12] and the reference therein.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

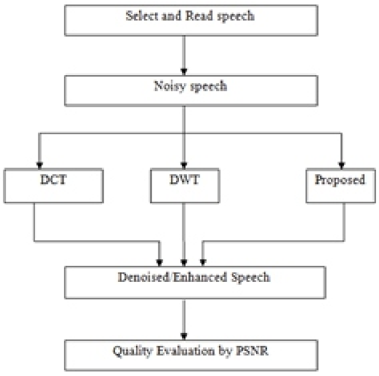


Fig.1 Block diagram of speech denoising/enhancement.

a.Local Ridgelet Transforms:

From literature review, it is found that Ridgelet transform is optimal to find only lines of the size of the image and for detecting line segments; the concept of partitioning was introduced [13]. From this, a two dimensional signal is decomposed into overlapping blocks of side length b pixels in such a way that the overlap between two vertically adjacent blocks is a rectangular array of size b by b/2 which avoid blocking artifacts [12]. If a 2D signal having n by n size, then 2n/b such blocks in each direction is counted. The authors in [12] have presented two competing strategies to perform the analysis and synthesis {12]. In first method, the blocks values are analyzed in such that the original pixel is constructed from the co-addition of all blocks. While in second method, only those block values are analyzed when the signal is reconstructed (synthesis) [11, 12]. From different experiment, the authors in [12] have concluded that the second approach yields improved performance. From second method, a pixel value, f(i,j)from its four corresponding block values of half size $\ell = b/2$, namely, $B_1 = (i_1, j_1)$, $B_2 = (i_1, j_1)$, $B_3 = (i_2, j_2)$ and $B_4 = (i_2, j_2)$ with $i_1, j_1 > b/2$ and $i_2 = i_1 - \ell, j_2 = j_1 - \ell$ is computed as [12]:

$$f_1 = \omega\left(\frac{i_2}{\ell}\right) B_1(i_1, j_1) + \omega(1 - i_1/\ell) B_2(i_1, j_1) \quad (8)$$

$$f_2 = \omega\left(\frac{i_2}{\ell}\right) B_3(i_1, j_2) + \omega(1 - i_2/\ell) B_4(i_2, j_2) \quad (9)$$

$$f(i,j) = \omega(j_2/\ell) f_1 + \omega(1 - j_2/\ell) f_2 \quad (10)$$

In above equations, $\omega(x) = \cos(\pi x/2)^2$. Detailed discussion on Ridgelet transform is given in [11-13] and the reference therein.

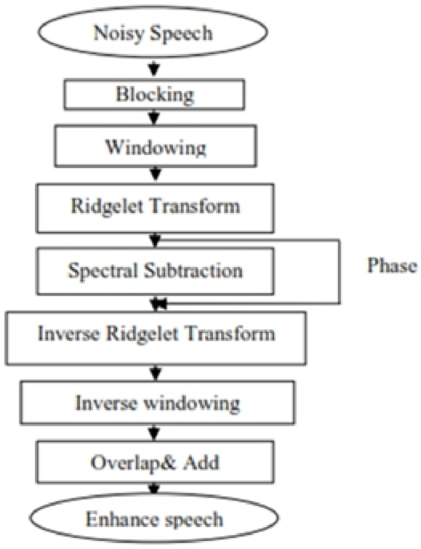


Fig.2 Proposed Ridgelet flow chart

IV.SIMULATION RESULTS:

To demonstrate the performance of proposed algorithm, here in this section we had presented the experimental results, which has been done in MATLAB coding in 2014a version with 4GB RAM and i3 processor. The experiments have been conducted on various speech samples, which is corrupted by gaussian noise of standard deviation sigma = 0.1 to 10. This corrupted speech signal has been denoised by using proposed technique. The results which have been shown in fig.2, show that the proposed ridgelet technique is significantly effective than the other techniques in terms of visual perceptual quality.

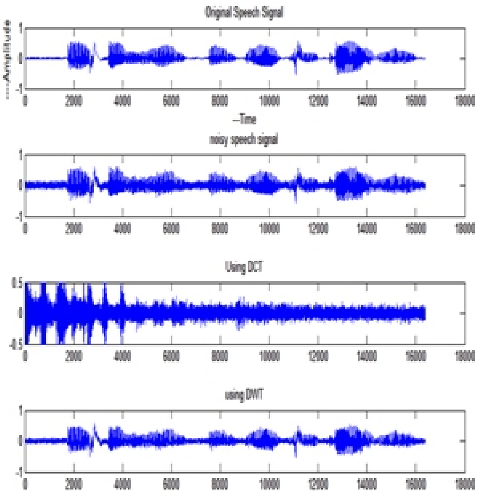
a.Statistical Parameters:

Here we are used Signal to Noise ratio (SNR), Mean Square error (MSE) and Maximum error (ME) to measure the quality of image, where PSNR will be used to measure the quality of image using a mathematical expression which as follows:

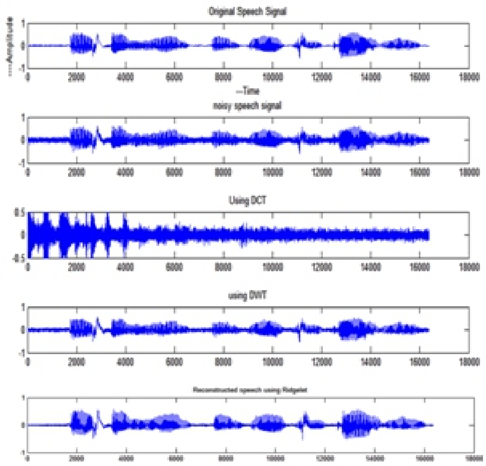
$$SNR = 10 * \log_{10}\left\{\frac{\sum x^2(n)}{\sum |x(n) - y(n)|^2}\right\}$$

$$MSE = \frac{1}{2}\sum_n |x(n) - y(n)|^2$$

$$ME = max|x(n) - y(n)|$$

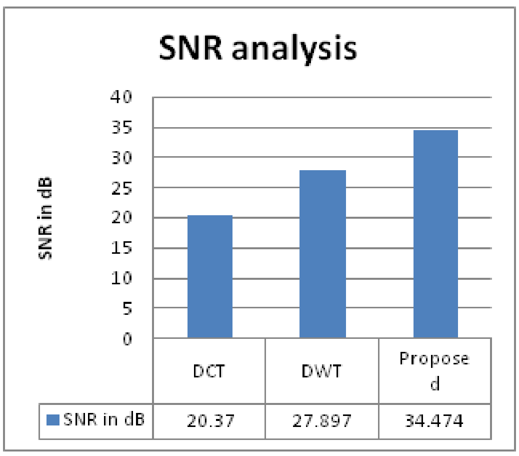


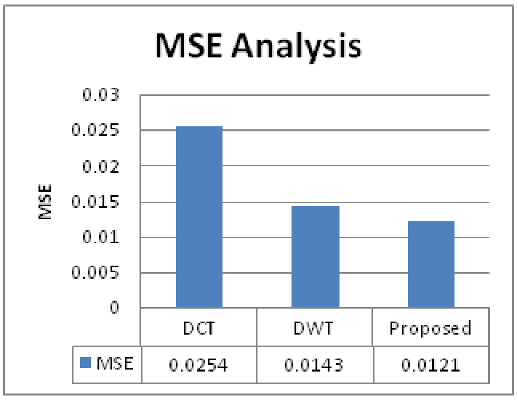
(a)



(b)

Fig2. Comparison of proposed and conventional techniques (a) Original, Noisy, DCT and DWT (b) Proposed algorithm





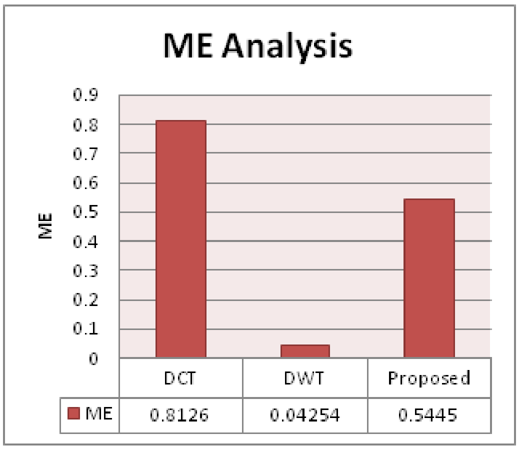


Fig3. Quality analysis of proposed and conventional techniques in terms of PSNR and MSE.

V.CONCLUSION:

Here we introduced a new speech denoising and enhancement algorithm in noisy environments. The proposed algorithm is based on transformations, which will give far better results and good performance than the spatial techniques. Here, we derived the ridgelet transformation for 1D and 2D signals also studied the performance analysis of proposed scheme over conventional techniques in terms of SNR, MSE and ME. Simulation results showed that the proposed scheme has performed well and superior to the existing denoising methods.

REFERENCES:

1.Buades, A, Coll, B, Morel J.M, “A non-local algorithm for image denoising”, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 20-26 June 2005, San Diego, CA, USA.

2.G. Oppenheim J. M. Poggi M. Misiti, Y. Misiti. Wavelet Toolbox. The MathWorks, Inc.,Natick, Massachusetts 01760, April 2001.

3.Yanna Ma* and Akinori Nishihara, "A modified Wiener filtering method combined with wavelet thresholding multi taper spectrum for speech enhancement", EURASIP Journal on Audio, Speech, and Music Processing 2014.

4.D.L. Donoho and I.M. Johnstone, Adapting to unknown smoothness via wavelet shrinkage, Journal of American Statistical Assoc., Vol. 90, no. 432, pp 1200-1224, Dec. 1995.

5.C. Valens. A really friendly guide to wavelets. eBook, 2013. http://perso.wanadoo.fr.

6.D.L.Donono, I.M.Johnstone. Adaptive to unknown smoothness via wavelet shrinkage [J]. Journal of the American Statistical Association. Vol.90, No.432, pp.1200-1224, 1995.

7.Chen juan, Chen qian-hui, Shi lu-huan. Edge detection technology in image tracking[J]. Chinese Journal of Optics and Applied Optics. Vol.2. No.1,pp.46-53.february 2009.

. Maarten Jansen, Noise Reduction by Wavelet Thresholding, Springer–Verlag New York Inc.- 2001.

9.D.L. Donoho and I.M. Johnstone, Ideal spatial adaptation via wavelet shrinkage, Biometrica, Vol. 81, pp. 425-455,1994.

10. D.L. Donoho and I.M. Johnstone, Wavelet shrinkage: Asymptopia?, J.R. Stat. Soc. B, ser. B, Vol. 57, no. 2, pp. 111-369, 1995.

11.Moore, R. Ezekiel, S. ; Blasch, E. "Denoising one-dimensional signals with curvelets and contourlets" IEEE National Aerospace and Electronics Conference,pp. 189-194, 2014.

12.J. L. Starck, D.L Donoho, E.J Candes, "Astronomical image representation by the curvelet transform" Aastronomy and astrophysics (berlin print), Vol. 398, pp.785-800, 2003.

13. M.J. Fadili , and J.-L. Starck, "Curvelets and Ridgelets", Technical report, 2007.

14.E. Candes, and D. Donoho, "Curvelets", Technical report, Statistics, Stanford University, 1998