An Improvement in Speech Signals Corrupted by Impulsive Noise using Wavelets Wiener Filter and ICA Technique

Neha Verma
M.Tech Scholar, Department of ECE, BTKIT, Dwarahat, UK, India.

Lalit Singh Garia
Assistant Professor, Department of ECE, BTKIT, Dwarahat, UK, India.

Abstract – The aim of this paper is to investigate the effect of noises on performance of speech signal de-noising using the method based on wavelets, wiener filtering and ICA. Determination of voiced and unvoiced speech, low and high pitch, and methods for selecting appropriate wavelets for speech compression are discussed. Discrete wavelet transform (DWT) has been applied for suppression of additive noise. Soft thresholding are used in the process to detect time occurrence of noise corrupting the signal. Based on the number of samples at a stretch that are corrupted, wiener filter with a variable size window compressed them by wavelet transform then create an Improved ICA (Independent component analysis) technique which will remove artifacts of speech signal. The results of stimulation show that proposed techniques provide enhancement in quality and intelligibility of speech signal.

Index Terms – Speech signal, noises, wavelet transform, wiener filter, short time energy, spectral centroid, ICA.

1. INTRODUCTION

Speech is a very common method for humans to convey information from one person to other. Noise act as disturbance in any form of communication which degrade quality of information signal. Speech signals are liable to suffer from variety of noise. Speckle noise, Gaussian noise, salt and pepper noise, poison noise and impulsive noise are some varieties. One of the most important branches of speech processing is speech enhancement which focuses on finding an optimal estimate of clean speech from noisy speech signal. Noise removal from speech signals is an area of interest of researchers since years. Various authors proposed many de-noising approaches, few are: soft decision and recursion, diffusion filtering, Signal Dependent Rank Order Mean Algorithm and methods based on wavelets and median filtering. Demerits suffered by these techniques involve improvement in the performance of highly corrupted samples whereas that of improvement in lightly corrupted samples decreases. On increasing the size of window of median filter the complexity in design increases. To alleviate the noise effects, technique certainly distorts other clean samples of signal too. Fourier domain is not suitable for non-linear and non-stationary signal like speech and it also provide only the frequency information of signal. Adaptive filter cannot detect noise from higher fluctuating signal. Artifacts are not efficiently removed from corrupted signals. So the proposed technique has made efforts to sort out above weakness. It uses wavelets along with wiener filter and ICA (independent component analysis) technique that would minimize effects of artifacts, de-noise linear as well as non-stationary speech signals and also detect noise from high as well as low fluctuating signals.

2. DETECTION OF IMPLUSIVE NOISE IN TIME DOMAIN

A. Discrete wavelet transform

Jean Morlet, introduced the concept of a `wavelet' in 1982. The wavelet are the small waves that enable us to obtain time and frequency information of a signal. Wavelet transform is an improved version of Fourier transform that has been used in various fields of signal processing.

Let a discrete time signal be f(x) with 'x' being sample of signal x=0,1…..M-1 and 'M' is number of sample. Discrete Wavelet Transformation (DWT) investigation in spatial domain provides good performance in detecting discontinuities or indirect changes.If the function being prolonged is a sequence of numbers, the resulting coefficients are called the DWT of f (x). The DWT transform pair is defined as following:

Scaling function term:

$$W_{\varphi}\ (j_0, k) = \frac{1}{\sqrt{M}} \sum_x f(x)\, \varphi_{j_0,k}(x)$$

Wavelet function term:

$$W_{\Psi}\ (j, k) = \frac{1}{\sqrt{M}} \sum_x f(x)\, \Psi_{j,k}(x)$$

Where f (x), $\varphi_{j_0,k}(x)$ and $\Psi_{j,k}(x)$ are functions of the discrete variable.

To get back f(x), IDWT transform pair is defined as following:

For $j=j_0$

$$f(x) = W_{\varphi}\ (j_0, k) + W_{\Psi}\ (j, k)$$

$$= \frac{1}{\sqrt{M}} \sum_k W_{\varphi}\ (j_0, k)\varphi_{j_0,k}(x) + \sum_{j=j_0}^{\infty} \sum_k W_{\Psi}\ (j, k)\Psi_{j,k}(x)$$

B. Thresholding Function

The selection of thresholding function is the main issue of wavelet threshold de-noising.

1) Global threshold function:

One is the hard threshold function:

$$W\hat{}_{j.k} = \begin{cases} W_{j,k}, & |W_{j,k}| \geq \lambda \\ 0, & |W_{j,k}| < \lambda \end{cases}$$

One is the soft-thresholding function:

$$W\hat{}_{j.k} = \begin{cases} sgn(W_{j,k}).(|W_{j,k}| - \lambda), & |W_{j,k}| \geq \lambda \\ 0, & |W_{j,k}| < \lambda \end{cases}$$

Where sgn(*) is a sign function, $W_{j,k}$ stands for wavelet coefficients, $W\hat{}_{j.k}$ stands for wavelet coefficients after treatment, λ stands for threshold value and it can be expressed as:

$$\lambda = \sigma\sqrt{2\ln(N)} \qquad \sigma = median(|c|)/0.6745$$

Where N is the image size, σ is the standard deviation of the additive noise and c is the detail coefficient of wavelet transform.

2) Penalized Threshold Function

MATLAB code for Penalized Threshold

$$THR = wbmpen(C, L, \sigma, \alpha)$$

Where [C, L] is the wavelet decomposition of the signal to be de-noised, σ is the standard deviation of the zero mean Gaussian white noise in de-noising model,α is a tuning parameter for the penalty term. It must be a real number greater than 1.

3) Proposed Threshold Method

If there is small changes threshold value it also destroys some important signal details that may cause error and artifacts in speech. So, optimum threshold value should be found out. The proposed threshold is given as:

$$\lambda_{\text{proposed}} = \lambda + 2\beta$$

Where, λ = global threshold value and β = penalized threshold value.

3. WAVELET DENOISING USING WIENER FILTERING

To avoid few limitations of thresholding the wavelet coefficients, another method was investigated, which is based on "Weiner filtering" in the wavelet domain.

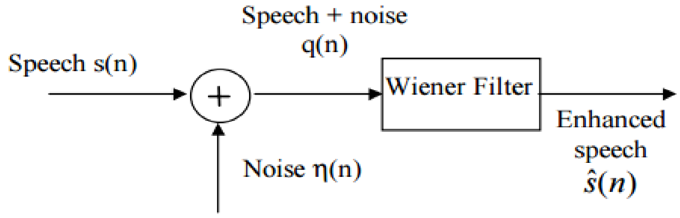


Fig1: Block diagram of Wiener filtering concept

From Figure 1, we can write:

q(n) = s(n) +η(n)

s^(n) = q(n) * h(n).

We can define an error signal, e(n), as follows:

e(n) = s(n) − s^(n)

Where s(n) is the speech signal, h(n) is the Wiener filter impulse response and η(n) is the additive noise component. Wiener filter is a filter processed by linear time-invariant (LTI) filtering to observe noisy process, assuming known stationary signal and noise spectra, and additive noise. The goal of the Wiener filter is to compute a statistical estimate of an unknown signal using a related signal as an input and filtering that known signal to produce the estimate as an output.

4.ICA(INDEPENDENT COMPONENT ANALYSIS)

ICA is a linguistics technique which is used to find hidden factors that underlie sets of random variables, measurements, or signals. ICA defines as a model which is generated to multivariate data, that is given as a large database of samples. In the model contain the data variables that are assumed to be linear mixtures of some unknown latent variables. These latent variables are assumed non gaussian and mutually independent, and they are called the independent components of the observed data.

Principal relating ICA is component analysis and factor analysis. ICA is a much more powerful technique, however, capable of finding the underlying factors or sources when these classic methods fail completely. It analyses the data originate from many different kinds of application fields, including digital images, document databases and economic indicators.

Applications of ICA are in separation of artifacts in MEG data, finding hidden factors in financial data and reducing noise in natural images.

5. IMPLEMENTATION AND RESULTS

A.Short time energy

In speech signal the amplitude varies with time. Speech have few voiced and few unvoiced segments and the amplitude of unvoiced is generally much lower than the amplitude of voiced.

$$E_n = \sum_m [s(m)\, w(n-m)]^2$$

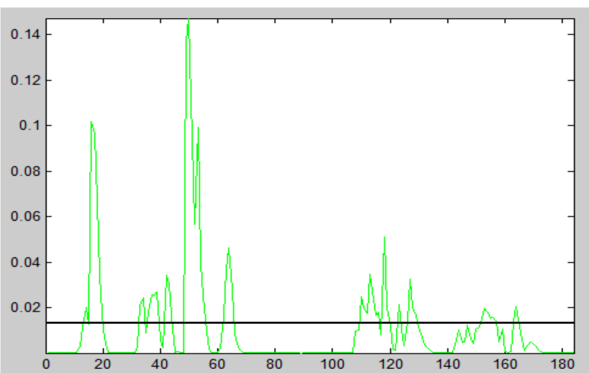
Where s(m) is a short time speech segment obtained by passing the speech signal x(n) through window w(n).

B.Spectral centroid

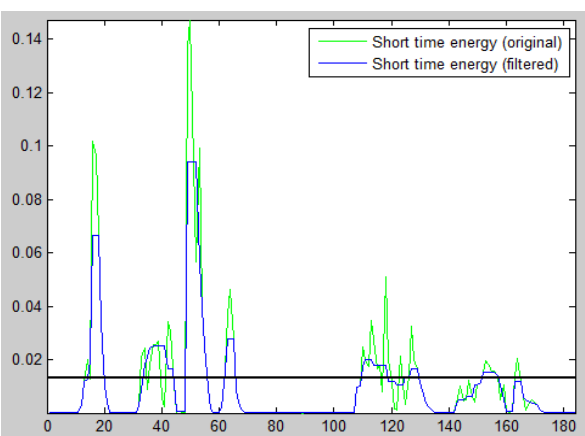
The spectral centroid is a tool used in digital signal processing to characterize a spectrum. Spectrum's center of mass is calculated by it. Centroids have connection with the impression of "brightness" of a sound by which low and high pitch of sound is determine.

$$\text{Centroid} = \frac{\sum_{N=0}^{N-1} f(n)x(n)}{\sum_{n=0}^{N-1} x(n)}$$

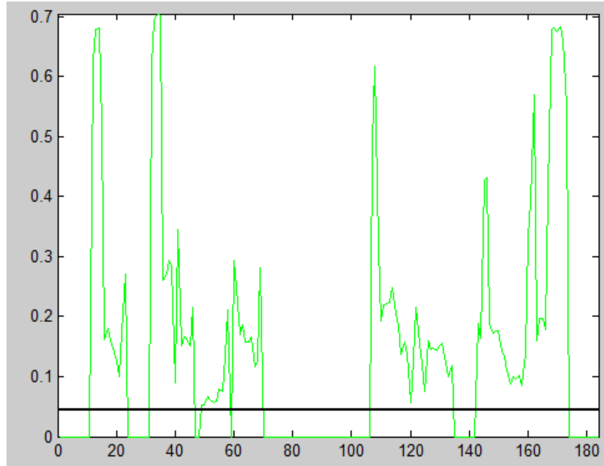
where x(n) represents the weighted frequency value, or magnitude, of bin number n, and f(n) represents the center frequency of that bin.



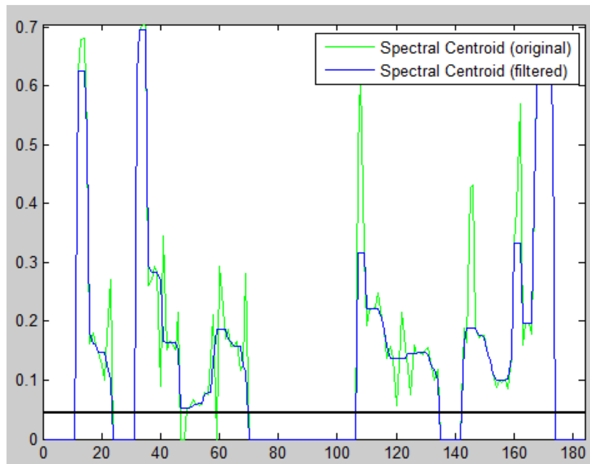
(1)



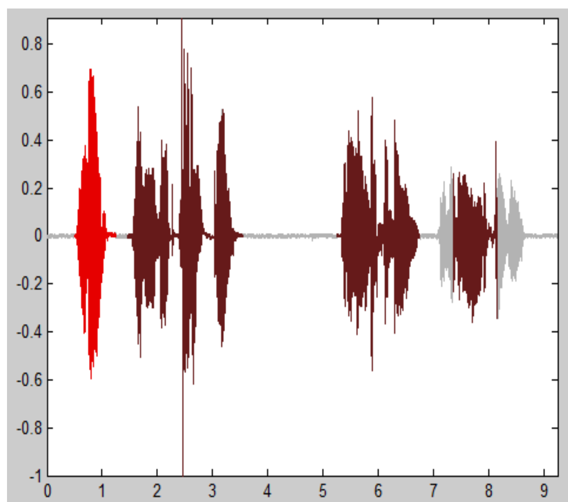
(2)



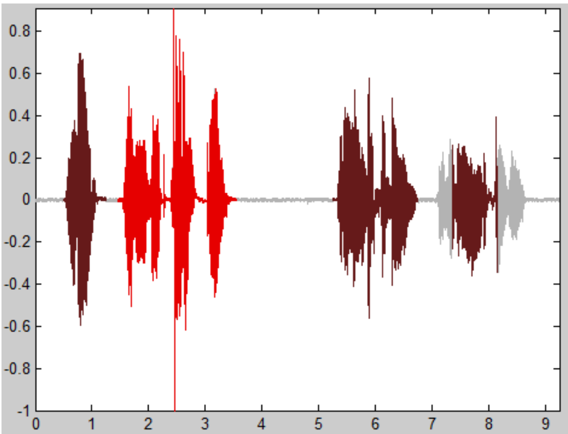
(3)



(4)



(5)



(6)

Result (1)(2) shows the de-noising of corrupted signal by determining voiced and unvoiced signal

Result (3)(4) shows de-noising according to high and low pitch.

Result (5)(6) noise is suppressed by windowing and clean signal is detected.

C.Detector efficiency

Efficiency γ is given as:

$$\gamma = \frac{100(N-\text{errors})}{N}$$

N= number of sample, errors= total number of errors detected.

The efficiency obtained when various wavelet families are used in above equation are as presented in Table1.

TABLE I

Detection efficiency using different wavelets

wavelet	**Efficiency(γ)**
Haar	46.41%
Daubechies	93.90%
Symlet	84.20%
Coifet	83.01%
Dmey (base)	97.9%
Dmey(proposed)	99.8%

D.Objective Measures

The signal-to-noise ratio is the most widely used measure to detect quality and intellectuality of speech signal. The SNR measure in frequency domain for k^{th} frame is defined by:

$$SNR_k = 10 \log_{10} \frac{\sum_n |X_k(n)|^2}{\sum_n [X_k(n) - \tilde{X}_k(n)]^2} \text{ [dB]}$$

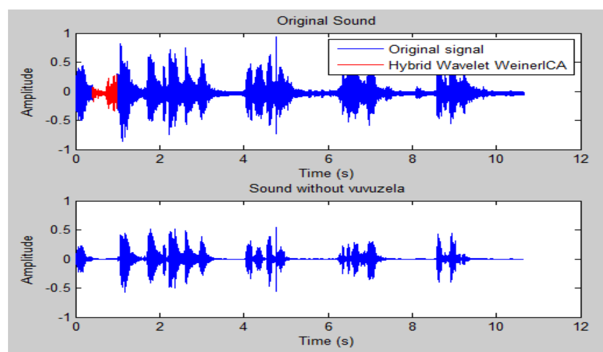
Where: $\tilde{X}_k(n)$ is the DFT of the de-noised speech signal and $X_k(n)$ is the DFT of the k^{th} frame of the clean speech. To get overall SNR these SNR_k for different frames are averaged.

2 pure speech signals corrupted by different kinds of additive noise (of varying time durations) are taken. By applying DWT we get approximate and detail coefficient, reconstruction in done in detail coefficient than thresholding is applied to get the time duration of noise added and once the noise is detected wiener filter of variable size window is applied on the corrupted speech signal to remove the noise, further an additional ICA techniques is applied to detect and remove remaining artifacts from the signal. The SNR of the corrupted speech signals (x1 and x2) comparing with previous technique is mentioned in Table 2.

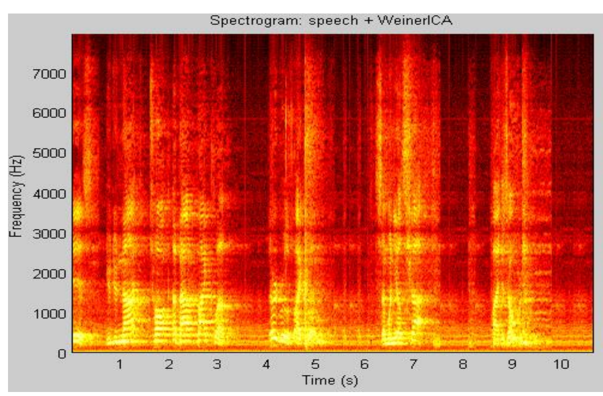
TABLE II

The improvement in SNR indicated for the proposed technique and wavelet adaptive median filtering approach.

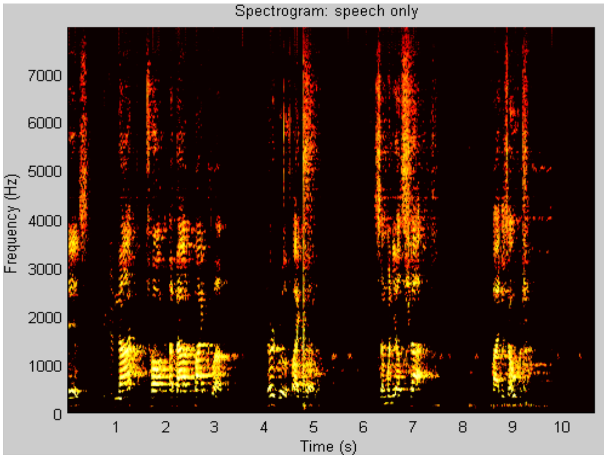
	Improvement in SNR	
	Proposed Technique	Wavelet adaptive filtering
x_1	8.330dB	8.93dB
x_2	23.330dB	18.7dB



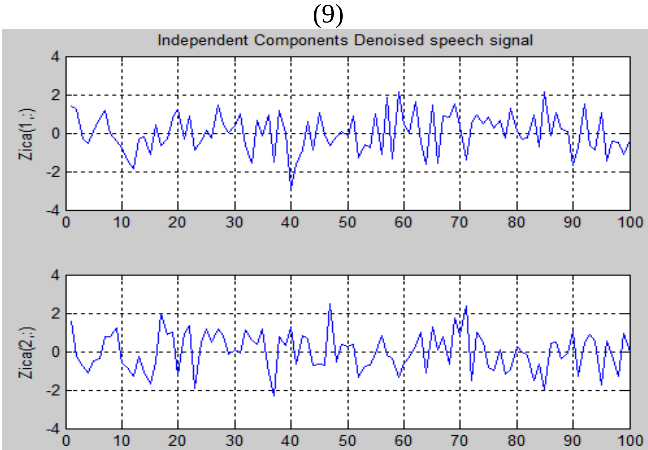
(7)



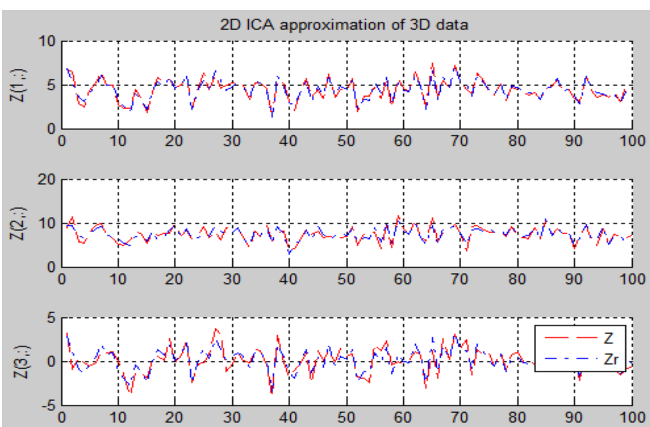
(8)



(9)



(10)



(11)

Result(7) shows amplitude of original sound and Hybrid wavelet Weiner ICA.

Result(8)(9)shows 'Frequency of Spectrogram' of speech + Weiner ICA on time and speech only on time.

Result(10) shows frequency and time graph, $Z_{ica}(1)$ is the de-noise speech signal only through wavelets, $Z_{ica}(2)$ is filtered de-noise speech signal through wavelet + wiener filter + ICA.

Result(11)shows the comparison between Z_r de-noise speech signal only through wavelets and Z filtered de-noised signal through wavelet + wiener filter + ICA at different frequencies positive and negative to show high level of performance.

6. CONCLUSION

The approach introduced is based on wavelet, wiener filtering and ICA techniques whose results indicate improvement in speech quality, has better performance and lower complexity than wavelet adaptive median filtering. It efficiently removes additive noise of varying time duration. Future work may include real time implementation of system.

REFERENCES

[1] S. V. Vaseghi, "Advanced digital signal processing and noise reduction." Wiley, 2008, pp. 29–43.

[2] R. C. Nongpiur, "Impulse noise removal in speech using wavelets," in Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on. IEEE, 2008, pp. 1593–1596.

[3] C. Chandra, M. S. Moore, and S. K. Mitra, "An efficient method for the removal of impulse noise from speech and audio signals," in Circuits and Systems, 1998. ISCAS'98. Proceedings of the 1998 IEEE International Symposium on, vol. 4. IEEE, 1998, pp. 206–208.

[4] Z. He, X. Guo, and M. Zhang, "Detection and removal of impulsive colored noise for speech enhancement," in Information and Automation (ICIA), 2010 IEEE International Conference on. IEEE, 2010, pp. 2320–2324.

[5] S. Zahedpour, S. Feizi, A. Amini, M. Ferdosizadeh, and F. Marvasti, "Impulsive noise cancellation based on soft decision and recursion," Instrumentation and Measurement, IEEE Transactions on, vol. 58, no. 8, pp. 2780–2790, 2009.

[6] A. Graps, "An introduction to wavelets," Computational Science & Engineering, IEEE, vol.2, no. 2, pp. 50–61, 1995.

[7] J. I. Agbinya, "Discrete wavelet transform techniques in speech processing," in TENCON'96. Proceedings. 1996 IEEE TENCON. Digital Signal Processing Applications, vol.2. IEEE, 1996, pp. 514–519.

[8] D. L. Donoho, "De-noising by soft-thresholding," Information Theory, IEEE Transactions on, vol. 41, no. 3, pp. 613–627, 1995.