Speech Enhancement Using Wavelet Transform

Manal Hade

Received on: 4 / 3 /2007
Accepted on: 29 / 6 /2008

Abstract

Speech enhancement is an important proceeding which it can be achieved by estimate the clean speech from noisy speech. The clean speech is corrupted by white Gaussian noise.

In this paper, thresholding the wavelet coefficients, that can be done by standard deviation method for each frame by level dependent thresholding using semisoft threshold and soft threshold. The results in this work indicate that using wavelet transform in speech enhancement application provides a good quality(measured in object SNR test) as well as provides other advantages which make it more suitable for some applications such as a split of frequencies and enhancement the speech in real time applications.

Key Words

Speech Enhancement, Wavelet Transform, Level Dependent Threshold.

الخلاصة

تكمن أهمية معالجات التحسين الصوتي بأنها تستطيع أن نتوصل وبواسطة التخمين إلى صوت اقرب ما يمكن للصوت الأصلي من صوت ضوضائي، حيث أن الصوت الأصلي يشوه بضوضاء من نوع Gaussian .
خلال هذه المعالجة تم استخدام طريقة تقليم عوامل المويجة Thresholding والتي انجزت بواسطة طريقة الانحراف المعياري لكل مقطع Frame باستخدام تقليم لكل مرحلة Level Dependent thresholding حيث تم استخدام نوعين من Threshold هما Soft threshold و semi soft threshold تشير النتائج المحصلة في هذا البحث إلى أن استخدام تحويل المويجة في تطبيقات إزالة الضوضاء من الصوت قد أعطى نتائج جيدة تم قياسها بالفحوصات الشيئية وكذلك تعطينا بعض الميزات التي تجعلها ملائمة لبعض التطبيقات مثل فصل ترددات الإشارة وكذلك يمكن استخدام تحويل المويجة في إزالة الضوضاء في الوقت الحقيقي .

Electromechanical Engineering Department University of technology

1- Introduction

In many speech processing applications, speech has to be processed in the presence of undesirable background noise like white Gaussian noise. 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 [1].

Speech noise reduction or speech enhancement has always been a non –trivial problem for communication engineers.The total removal of background noise is practically impossible and distortion of the speech content is inevitable [1] .

In this paper speech enhancement algorithms are designed relying on the wavelet based denoising techniaque with level dependent threshold.

The wavelet transform distinguishes itself in the analysis of non-stationary signals such as speech.The wavelet shrinkageis a powerful tool in denoising signal corrupted by different types of noise .Recent efforts made for speech enhancement using wavelet shrinkage include different types of threshold [1].

2- Discrete Wavelet Transform (DWT)

A set function $\Phi(t)\ and \psi_{j,k}(t)$ is constructed that could span all of $L^2(R)$,so the general statement of the expansion function $f(t) \in L^2\ (R)$ can be given by [2] :

$$f(t)=\sum_k c_j(k)2^{j/2}\Phi(2^j t-k)+\sum_k d_j(k)2^{j/2}\psi(2^j t-k)...(1)$$

Let the function $f(t)$ be a discretely sampled function .The decompostion of $f(t)$ in the wavelet basis is done by recursive filtering with H and G with down –sampling of factor of two in each set .A lower resolution signal is delivered by low pass filtering with half-band low pass filter H followed by down-sampling by two.The higher resolution (or detail) is computed by a high pass filtered followed by down-sampling by two [3].

The coefficients h (n) and g (n) ,used to construct the set of scaling and wavelet basis, are low pass (H) and high pass (G) FIR filter coefficients respectively. Where H= {h (n)} and G = {g (n)}.According to the equation (2), G is the reverse of H [3, 4].

$$g(n)=(-1)^n\ h\ (N-n)...(2)$$

Figure (1) shows filter band of discrete wavelet transform .The symbol ↓2 is down-sampler (decimator) that it takes a signal x(n) as input and produces an output of y(n) = x(2n) ,which mean that is discarded [5].

The inverse discrete wavelet transform is illustrated in figuer (2) , Wavlete reconstruction process of up-sampling and filtering up-sampling process of lengthening signal component by inserting zeros between samples. In other words ,the input signal is streched twice its original length and zeros are inserted in the even numbered samples [6,7].

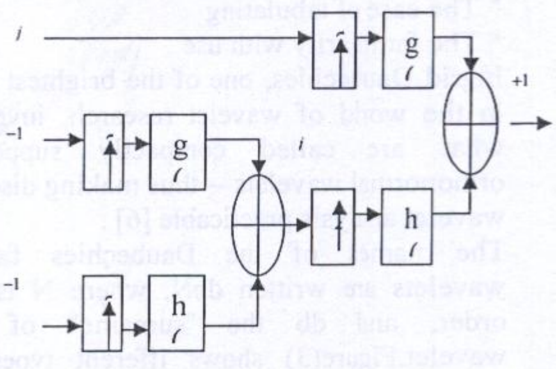


Figure (2) Two stage two -band tree signal synthesis tree of IDWT

The j scale coefficients segeunce c_j is up-sampled by doubling its length, then convoluting it with the scaling coefficients h(n),the same is done to the j level wavelet coefficient d_j sequence and the results are added to give the $(j+1)$level scaling function coefficients [5].

3-Wavelet Families

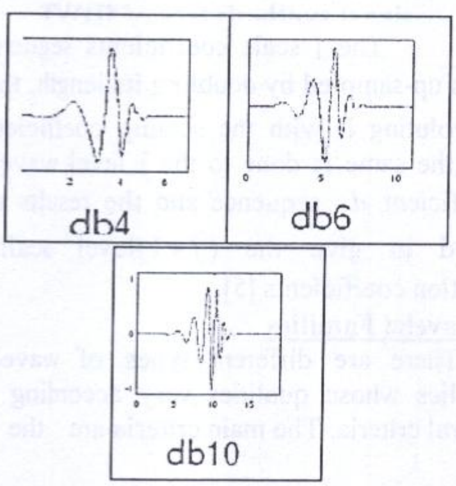
There are different types of wavelet families whose qualities vary according to several criteria. The main criteria are the supported of $\psi, \widetilde{\psi}(and \phi, \widetilde{\phi})$. The speed of convergence to 0 of these functions $\varphi(t) or \widetilde{\psi}(\omega)$ when the time t or the frequency goes to infinity, which quantifies both time and frequency localizations

The symmetry, which is useful in avoiding dephasing in image processing

* The number of vanishing moments for ψ or for ϕ (if it exists), which is useful for compression purposes
* The regularity, which is useful for getting nice features, like smoothness of the reconstructed signal or image, and for the estimated function in nonlinear regression analysis These are associated with two properties that allow fast algorithm and space-saving coding:
* The existence of a scaling function ϕ
* The orthogonality or the biorthogonality of the resulting analysis They may also be associated with these less important properties
* The existence of an explicit expression
* The ease of tabulating
* The familiarity with use

Ingrid Daubechies, one of the brightest stars in the world of wavelet research, invented what are called compactly supported orthonormal wavelets -- thus making discrete wavelet analysis practicable [6].

The names of the Daubechies family wavelets are written dbN, where N is the order, and db the "surname" of the wavelet.Figure(3) shows ifferent types of wavelet family.



Figuer(3) show the wavelet families db4 –db6

4- Speech Enhancement Techniques

Proposed a powerful approach for noise reduction.

It is based on the thresholding of the wavelet coefficients. Let y be a finite length observation sequence of the signal x that is corrupted by zero-mean white Gaussian noise m with variance σ^2 [10].

y = x+m .....(3)

In the wavelet domain, this gives :

$W_y = W_x + W_m$.....(4)

The clean signal can be estimated in the following way :

$x = W^{-1} X_{estimation} = W^{-1} Y_{thresh}$....(5)

where: Y_{thresh} represents the wavelet coeficients after thresholding.

W^{-1} denotes the inverse wavelet transform .

The wavelet transform projects the signal onto the transformed domain where the signal energy is concentrated in small number of coefficients, while the noise is evenly distributed across the transformed domain [1].

5- Wavelet Thresholding

Generally ,thresholding process has many types as will be shown below. Thresholding operations are applied to the coefficients of the wavelet .

5-1 Soft Thresholding

Soft thresholding can be stated mathematically as .[9,10]

$$THR(y) = \begin{cases} \text{sign}(y) \times (|y| - \lambda) & |y| > \lambda \\ 0 & |y| \leq \lambda \end{cases} .(6)$$

$$\lambda = \sigma_j \sqrt{2 \log N_j} ......(7)$$

with $\sigma_j = \frac{MAD(D_j)}{0.6745}$ ....(8)

where $MAD(D_j)$ is median absolute deviation of coefficients for each level, and N_j is the data length for each level .

$$\text{sign}(y) = \begin{cases} -1 & \text{if } \; y<0 \\ 0 & \text{if } \; y=0 \\ 1 & \text{if } \; y>0 \end{cases} \quad \text{...(9)}$$

Figure (4) shows input /output characteristics of the ramp signal for soft thresholding with $\lambda_1 = 0.2$

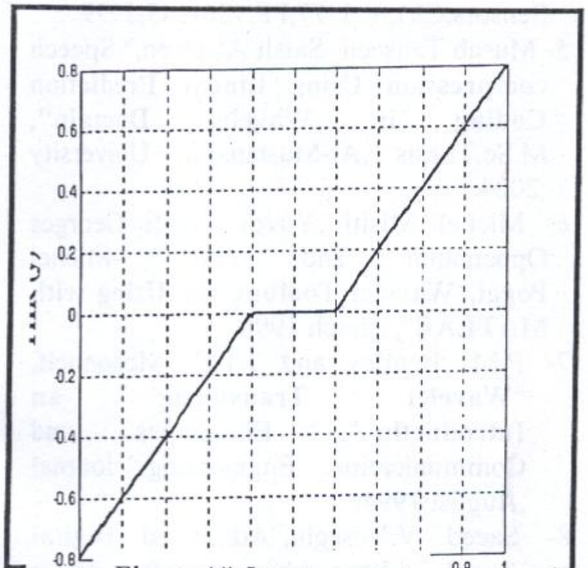


Figure (4) Input - output soft of thresholding characteristics λ =0.2.

5-2 Semi soft –Thresholding

The semisoft threshold function is given by [11, 12]

$$THR(y) = \begin{cases} 0 & |y| \le \lambda_1 \\ \text{sgn}(y)\dfrac{\lambda_2(|y|-\lambda_1)}{\lambda_2-\lambda_1} & \lambda_1 < |y| \le \lambda_2 \\ y & |y| > \lambda_2 \end{cases}$$

Where THR (y) represents the output valu after thresholding the wavelet coefficient λ_1 *and* λ_2 denote lower and upper hreshold respectively. Thresholding value λ_1 is determined by equation(7) and λ_2 is given by :

$$\lambda_2 = \sqrt{2}\lambda_1 \text{.......(11)}$$

Figure (5) shows input /output characteristics of the ramp signal for semisoft thresholding with $\lambda_1 = 0.2$ and λ_2 =2.8

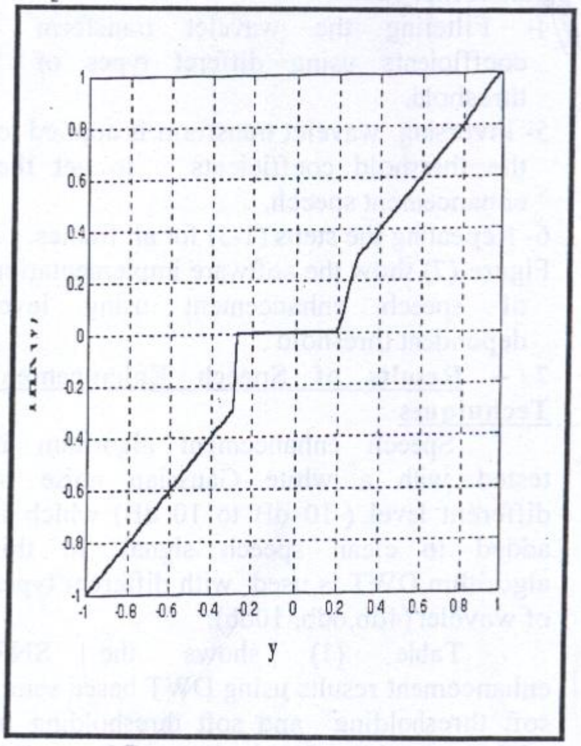
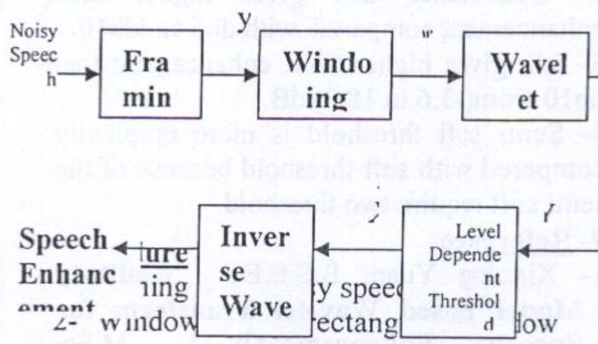


Figure (5) Input - output characteristics of semisoft -thresholding with $\lambda_1 = 0.2$ and λ_2 =2.8

6- W

Level dependent threshold is based on thresholding detail coefficients for each level. Figure (6) shows the block digram of speech denoising using level dependent threshold.



2- window (n).

$y_w(n) = y(n) \cdot w(n)$.........(12)

3- O speec band.

Figuer (6) Block diagram of speech enhancement using level dependent threshold

$Y = \omega^L y_w$..........(13)

Where ω^L denotes L- stage wavelet transform .

4- Filtering the wavelet transform coefficients using differet types of threshold.

5- Inversing wavelet transform is applied to the threshold coefficients to get the enhancement speech.

6- Repeating the steps (1-5) for all frames.

Figure (7) show the software implementation of speech enhancement using level dependent threshold .

7 – Results of Speech Enhancement Techniques

Speech enhancement algorithm is tested with a white Gaussian noise of different level (-10 dB to 10 dB) which is added to clean speech signal. In this algorithm DWT is used with different types of wavelet (4db,6db, 10db).

Table (1) shows the SNR enhancement results using DWT based semi-soft thresholding and soft thresholding in additive white Gaussian noise. All duabechies types have the same results and minimum SNR enhancement exceed 6 db.

8 - Conclusions

The following points are concluded from the simulation result.

1- Level dependent threshold using semisoft threshold gives higher SNR enhancement compared with soft threshold .

2- Duabechies db4 gives higher SNR enhancement compared with db6 anddb10.

3- 6db gives higher SNR enhancement than db10 from(-3.6 to 18.3) dB.

4- Semi soft threshold is more omplexity compared with soft threshold because of the semi soft require two threshold.

9- Reference

1- Xiaolog Yuan, B.S.E.E., **"Auditory Model Based Wavelet Transform for Speech Enhancement** ", M.Sc. Thesis,Marquette University,May 2003.

2- Oliver Rioul and Martin Vetterli,"**Wavelet and Signal Processing** ",IEEESP Magazine,October 1991

3- C. Sidney Burrus, Ramesh A.Copinath, and Haitao Guo,"**Introduction to Wavelet and wavelet Transform**", prentice Hall,1998

4- James F.Scholl,Jonathan R .Agre ,Loren P.Clare ,and Martin C.Gill,"**A low Power Impulse Classifier Using The Haar Wavlet Transform** ", proc.SPIE, Sensors,C31,V.3577,PP.136-145,1999 .

5- Musab Tahseen Salah Al-Deen," **Speech compression Using Linear Prediction Coding in Wavelet Domain**", M.Sc.Thesis ,Al-Mustansiria University ,2004.

6- Michel Misiti ,Yuves Misiti Georges Oppenheim and Jean -Michel Poggi,"**Wavelet Toolbox for Using with MATLAB**", March 1996.

7- P.M. Bentley and J.T.E. Mcdonnelf, "**Wavelet Transform: an Introduction**", Electronics and Communication Engineering Journal ,August 1994.

8- Saeed V.Vaseghi,"**Advanced Digital Signal Processing and Noise Reduction**", by John Wilry and Sons Ltd,2000.

9- Iain M.Johnstoe, and Bernard W. Silverman, "**Wavelet Threshold Estimation for Data with Corrected Noise** ",Stanford University, Bristol University, August 20,1996.

10- David L.Donoho,"**Denoising by Soft – Thresholding** ",Stanford Statistics Department ,1992.

11- Jong Won Seok and Keun Sung Bar, "**Speech Enhancement with Reduction of Nois Components in the Wavelet Domain** ", School of Electronic and Electrical Engineering, Kyunpook National University ,Taegu ,Korea,1997.

12- Hong –Ye Gao,"**Wavelet Shrinkage Denoising Using The Non-Negative Garrote**", American Statistical Association Institute of Mathematical Statistic and Interface Foundation of North America,1998.

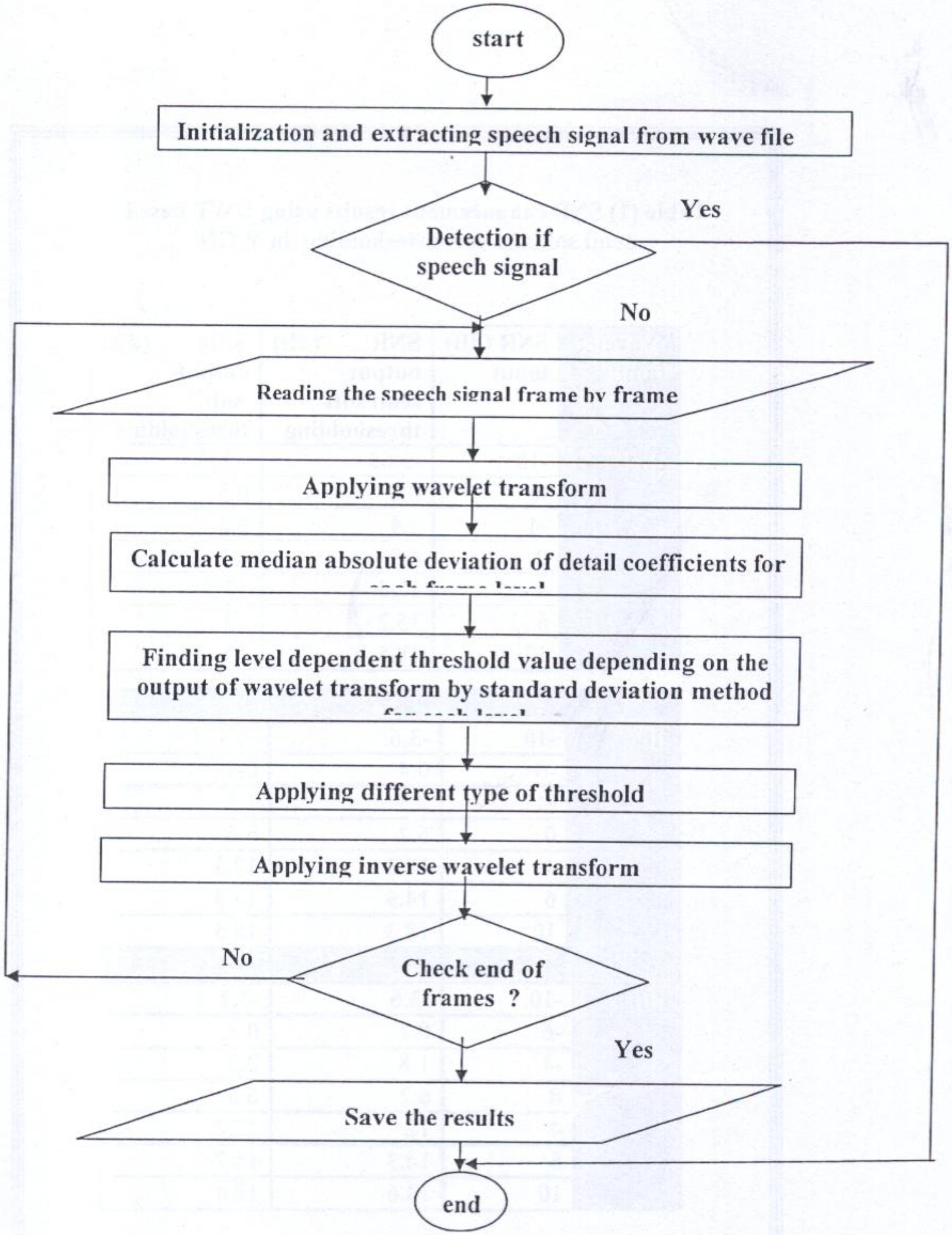


Figure (7) Flowchart of speech enhancement using level dependent threshold

Table (1) SNR enhancement results using DWT based semi soft and soft thresholding in WGN

Wavelet family	SNR (dB) input	SNR (dB) output semi soft thresholding	SNR (dB) output soft thresholding
db4	-10	-3.65	- 3. 5
	-6	0.6	0.5
	-4	2.4	2.3
	0	6.9	6.7
	4	12.5	12.5
	6	15.2	15
	10	18.5	18
db6	-10	-3.6	-3.4
	-6	0.4	0.4
	-4	2.3	2.2
	0	6.7	6.6
	4	12.3	12.3
	6	14.9	14.8
	10	18.3	18.5
db10	-10	-3.6	-3.7
	-6	0.2	0.3
	-4	1.8	2.2
	0	6.2	6.6
	4	12	12.2
	6	14.3	14.7
	10	13.6	18.4