Application of the Wavelet transformation to denoising and analyzing the speech

Phan Duy Hung, Nguyen Thi Lan Huong
Pham Thi Ngoc Yen, Eric Castelli
International Research Center **MICA**
Multimedia **I**nformation, **C**ommunication & **A**pplication
C10, Hanoi University of Technology
Tel: 084.8683087 Fax: 084.8683551

ABSTRACT: Wavelet transform (WT) has attracted most engineers and scientists because of its excellent properties. The coherence of practical approach and a theoretical basis not only solves currently important problems, but also gives the potential of formulating and solving completely new problems. It has been show that multi-resolution analysis of Wavelet transforms is good solution in speech analysis and threshold of wavelet coefficients has near optimal noise reduction property for many classes of signals. This paper proposed applications of wavelet in speech processing: pitch detection, voice–unvoice (V-UV) decision, denoising with the detailed algorithms and results.

I. INTRODUCTION

In speech compression, coding, analysis, synthesis and recognition applications, denoising the speech, calculate the pitch period and determinate voiced or unvoice are very important parameters.

When the transformed signal based on Fourier transformation, signal is transformed to frequency domain, a linear filter will separate the signal into original signal and noise approximatively. We can't do that at the frequency ranges piled up. So with the linear methods in time-frequency domain, one tries to select the combinative systems, in which the part of spectrum piled up is least.

The non - linear method is different. The spectrum of both the signal and the noise can piled up arbitrarily. Using the amplitude is better than spectral locating. This method permits to trim, threshold, contract the amplitude of signal transformed to denoise or separate. Wavelet has many special characters apply to digital processing. In the non - linear processing, they are the good location capability in time - frequency domain and the concentrated character of the WT coefficients.

The denoising method based on discrete wavelet transformation, the result is compared to a threshold. These values are set to zero if it is smaller than a threshold, then invert wavelet transformation. The denoising capability is very good based on the concentrated character of the WT coefficients. If a signal has energy concentrated at a few scales, the thresholding of wavelet coefficients will remove noise that has small amplitude or the signal unexpected.

Some classical algorithms to calculate the pitch period are: parallel processing in time domain, convolution function, mean differential function, analysis Cepstrum, Scaled Inverse Fourier Transformation (SIFT) algorithm; Determine V – UV: based on the pitch values, rate of zero passing, spectral flatness.

Each method has different strong and weak points, but the degree of accuracy isn't good. Wavelet transformation has degree of accuracy better.

II. METHOD

In dyadic form, the wavelet transform of the signal $f(t)$ is defined by the relation:

$$DWT_f\left(b,2^j\right)=\frac{1}{\sqrt{2^j}}\int_R \psi^*\left(\frac{t-b}{2^j}\right).f(t).dt$$

Where:

b is the time delay

2^j is the scale parameter

$\psi^*(t)$ is the complex conjugate of a wavelet function for which: $\int_{-\infty}^{+\infty}\psi(t)dt = 0$.

So that we can define the corresponding Fourier transform $\psi(\omega)$. The function $\psi(t)$ is a wavelet if and only if its Fourier transform $\psi(\omega)$ satisfies:

$$\int_0^{\infty}\frac{|\psi(\omega)|^2}{\omega}d\omega = \int_{-\infty}^{0}\frac{|\psi(\omega)|^2}{\omega}d\omega = C_\psi < \infty$$

A. Denoising with the DWT

Assume a finite length signal with additive noise of the form:

$$y_i = x_i + \varepsilon\, n_i, \qquad i = 1,\dots, N \qquad (1)$$

as a finite signal of observations of the signal x_i that is corrupted by white Gaussian noise n_i with standard deviation ε. $n_i \sim N(0,1)$. The goal is to recover the signal x from the noise observations y. Let W be a left invertible wavelet transformation matrix of the discrete wavelet transform. Equation (1) can be written in the transformation domain:

$$Y = X + N \quad \text{or} \quad Y_i = X_i + N_i \qquad (2)$$

Where capital letters denote variables in the transform domain, i.e, $Y = Wy$. Then the inverse transform matrix W^{-1} exists, and we have: $W^{-1}W = I$

Follow Dohono's approach, assume an orthogonal wavelet transform with a square W, i.e $W^{-1} = W^T$.

Let $\hat{X}$ denote an estimate of X, based on the observations Y. We consider diagonal linear projections $\Delta = \text{diag}(\delta_1, \ldots, \delta_N)$, $\delta_i \in \{0,1\}$, $i = 1, \ldots, N$. Which give rise to the estimate:

$$\hat{x} = W^{-1}\hat{X} = W^{-1}\Delta Y = W^{-1}\Delta W y$$

The estimate $\hat{X}$ is obtained by simply keeping or zeroing the individual wavelet coefficients. Since we are interested in the l_2 error, we define the risk measure:

$$R(\hat{X}, X) = E\left[\|\hat{x} - x\|_2^2\right] = E\left[\left\|W^{-1}(\hat{X} - X)\right\|_2^2\right] = E\left[\|\hat{X} - X\|_2^2\right] \quad (3)$$

Notice that the last equality in the equation (3) is a consequence of the orthogonality of W. The optimal coefficients in the diagonal projection scheme are $\delta_i = 1_{X_i > \varepsilon}$, i.e only those values of Y where the corresponding elements of X are larger than ε are kept, all others are set to zero. This leads to the ideal risk:

$$R_{id}(\hat{X}, X) = \sum_{n=1}^{N} \min(X^2, \varepsilon^2)$$

The ideal risk cannot be attained in practice, since it requires knowledge of X, the wavelet transform of the unknown vector x. However, it does give us a lower limit for the l_2 error.

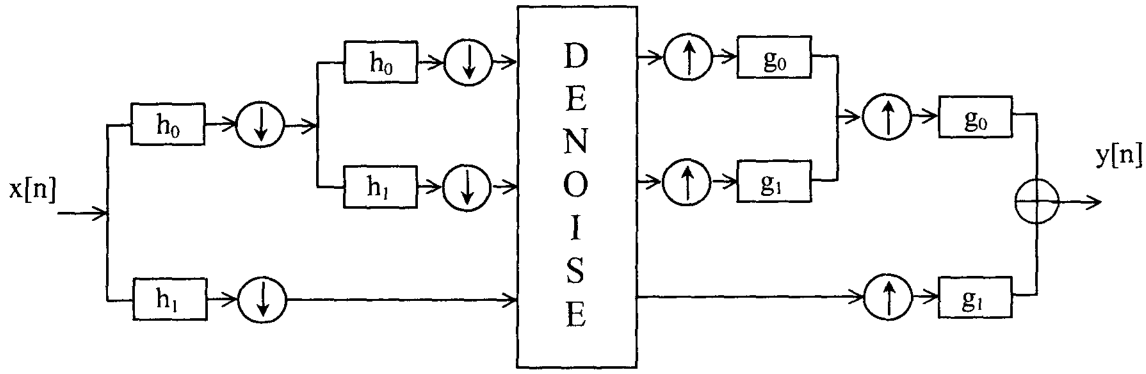

Figure 1: Diagram denoising with the DWT

h_0, h_1, g_0, g_1 are the impulse responses of the finite impulse response (FIR) filter in tree analysis and synthesis [3].

The scheme for denoising is follow [6]:

1. Compute the DWT: $Y = Wy$

2.Perform thresholding in the wavelet domain, according to so – called hard thresholding:

$$\hat{X} = T_h(Y, t) = \begin{cases} Y, & |Y| \geq t \\ 0, & |Y| < t \end{cases}$$

or according to so – called soft thresholding:

$$\hat{X} = T_S(Y, t) = \begin{cases} \text{sgn}(Y)(|Y| - t), & |Y| \geq t \\ 0, & |Y| < t \end{cases}$$

3. Compute the inverse DWT: $\hat{x} = W^{-1}\hat{X}$

Threshold $t(N, \varepsilon)$ based on:

+ The principle Stein's Unbiased Risk Estimate (SURE).

+ Fixed threshold: sqrt(2* log(length(y))).

+ Combine two above methods.

+ The principle minimax.

B. Calculate the pitch period and determine V - UV

- Calculate the pitch period

The signal is analyzed by Mallat algorithm: the signal is filtered by two low pass and high pass filters, then down - sampling. The parts low pass filtered are processed continuously with two above filters. The results are the detailed parts and approximation parts of the signal. The approximation parts show the information, which varies slowly. Through experiment, we said that at the 'scale 5', the approximation parts can present the global periodic form of the speech i.e we can calculate the pitch period in the time domain. Algorithm diagram as follows (Fig 2).

- Determine V – UV

We based on the results F_0 directly (Fig 3)

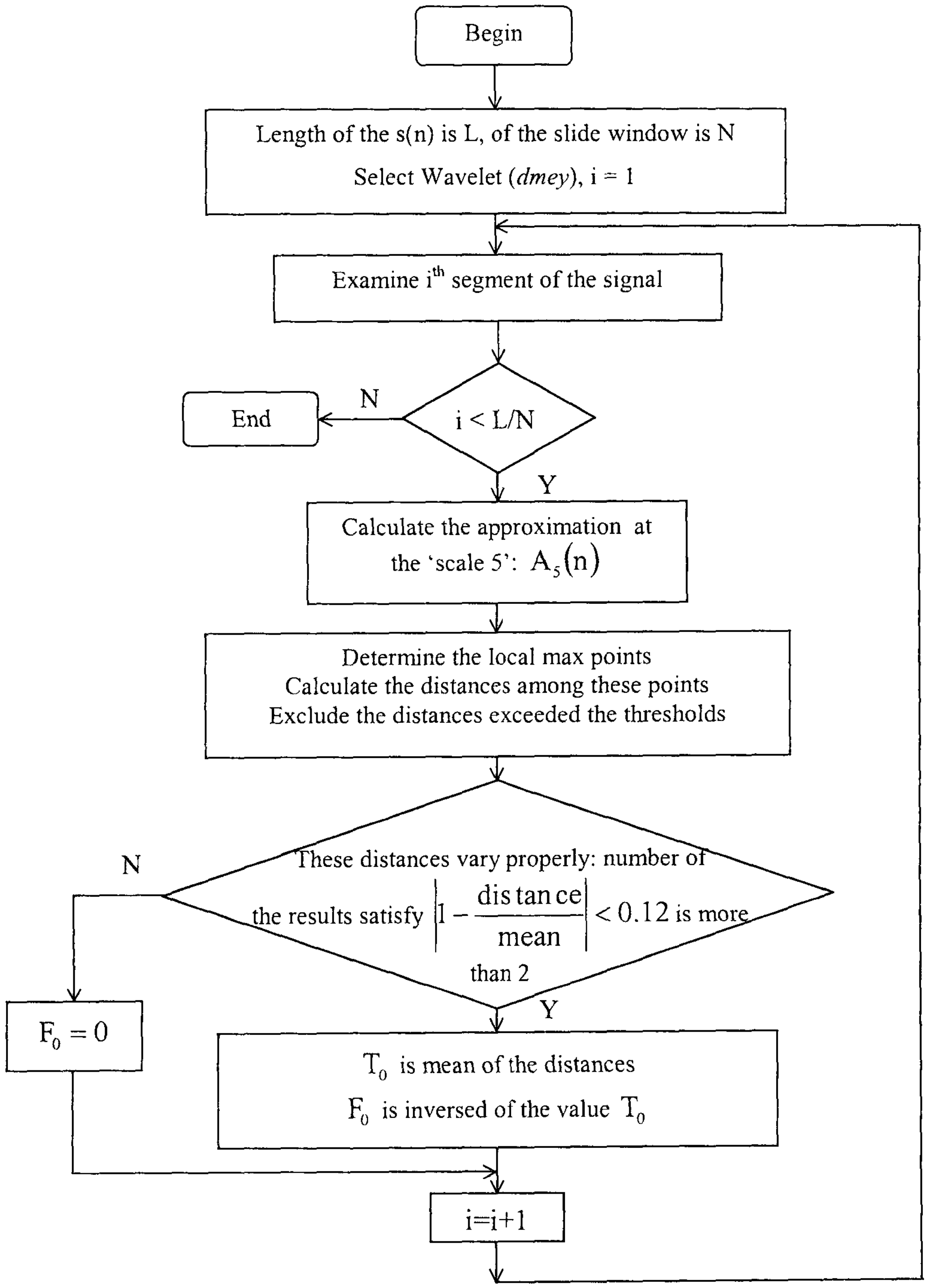

Figure 2: Diagram calculate the pitch period

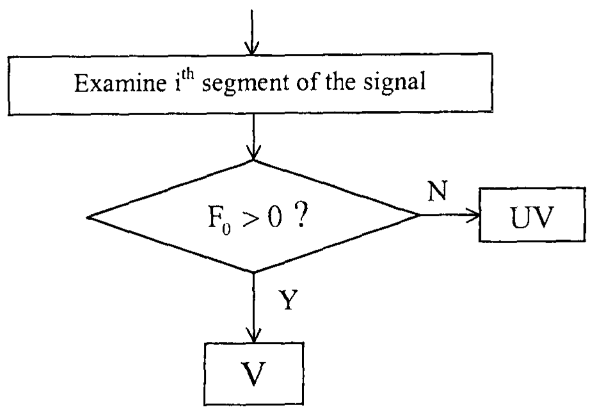

Figure 3: Determine V - UV based on F_0

IV. RESULTS AND EVALUATIONS

A program on Matlab language is realized, it has some functions, for example: record, display, edit, analyze, statistics.

Denoising algorithm here can apply to denoise arbitrary signal, where noise is white Gaussian noise. However, the algorithm denoising by the DWT has good result if the number of the non-zero coefficients is small. We can observe clearly with the short signal, the result is relative. With the long signals, we hear the denoised sounds, and the result is good.

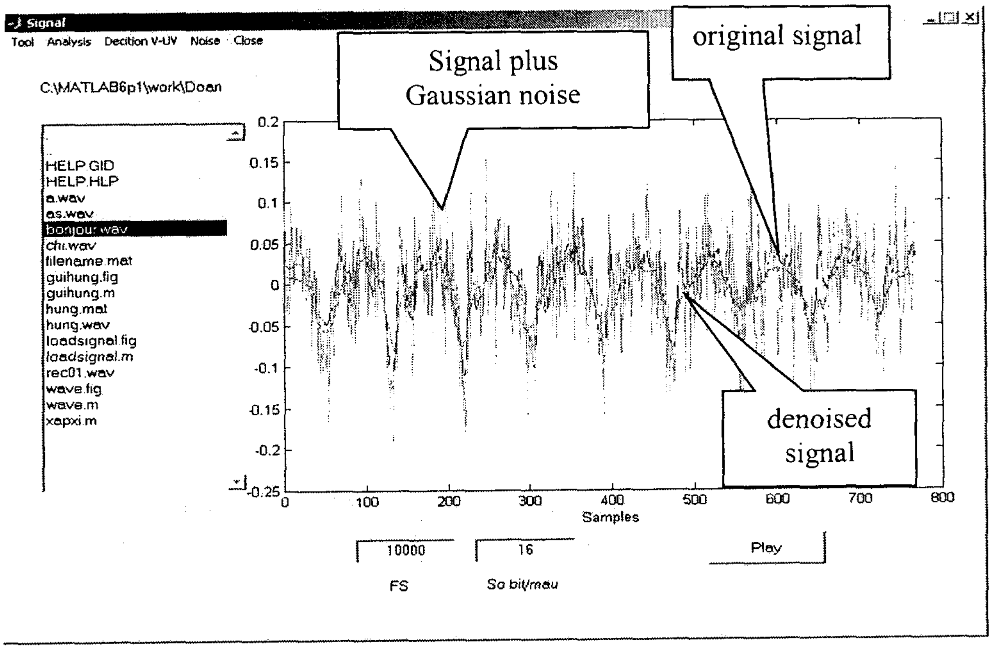

Figure 4: Illustration the addition noise and denoising

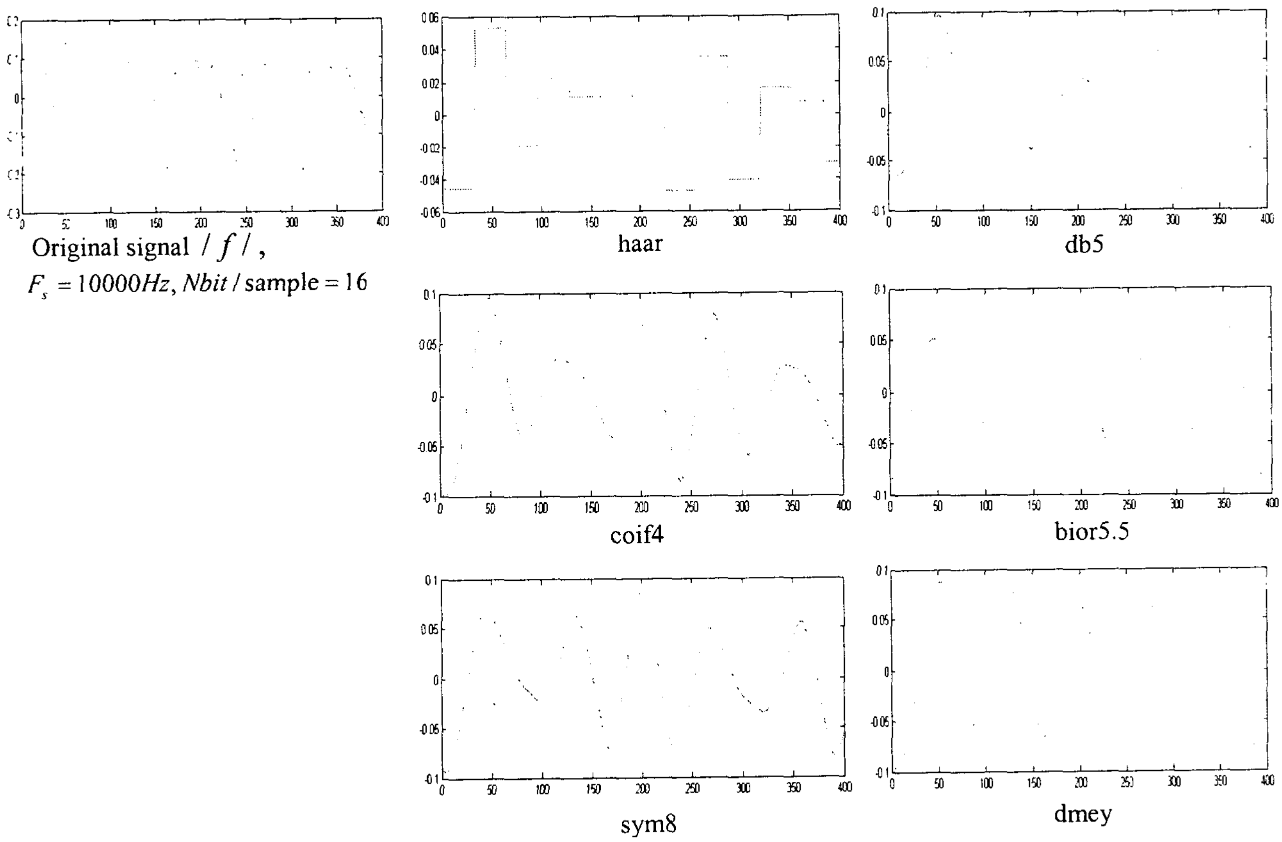

Figure 5: The approximation parts at the 'scale 5' correspond to the different wavelets

Survey the calculation the pitch period, the accuracy is 98% with some wavelets: dmey, bior – spline, some of the wavelet coif, sym, db. Wavelet Haar give the least accuracy, 15 ÷ 20%. This accuracy is better than the classical methods. With the continuous speech, wavelet transformation to calculate the pitch period and determine V – UV give the result, it's very good. Example here is the signal /the song / with sample frequency is 10000Hz and Nbits/sample is 16, selected wavelet is dmey.

/the song/

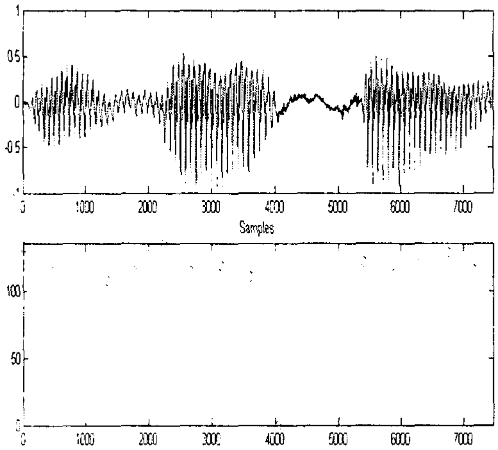

Figure 6: Determine V - UV of the speech

V. CONCLUSION

Wavelet transformation opened many tendencies for the digital signal processing applications. The algorithms and their degree of the complication were contracted more and more the researcher. This paper studied the wavelet transformation and applied to analyze, denoise the speech. Some main objects in future:

- Realizing the program in real time with the DSP C6701 TI, the results will be the kernel for the speech coding by multi-band excitation (MBE) - Wavelet method.
- Studying profoundly the Wavelet theory, the methods to design, select the optimal solution for some classes of the problems, the methods of the realization in real time. The classes of the problems are resolution, compression, denoising the signal; determine the points of the signal, which change suddenly in time domain, determine the global form of the signal, segment the signal.

References

[1] ***Bernard Mulgrew, Peter Grant and John Thompson***; *Digital Signal Processing, Concepts and Applications*; Macmillan press Ltd, ISBN 0 – 333 – 74531 – 0.

[2] ***René Boite, Hervé Bourlard, Thierry Dutoit, Joël Hancq et Henri Leich***; *Traitement de la parole*; Presses polytechniques et universitaires romandes.

[3] ***Martin Vetterli, Jelena Kovačević***; *Wavelets and Subband Coding*; Prentice Hall PTR, Englewood Cliffs, New Jersey 07632.

[4] ***Stéphane Mallat***; *A Wavelet Tour of Signal Processing*; Academic Press.

[5] *Scanning the special issue on Wavelets*;

Proceedings of the IEEE, April 1996 Vol 4, No.4.

[6] ***Y. MEYER***; *Wavelets and applications*; RMA20, Springer – Verlag.

[7] ***Tom H.Koornwinder***; *Wavelets: An Elementary Treatment of Theory and Applications*; World Scientific, Singapore – New Jersey – London – Hong Kong.

[8] ***Y. MEYER***; *Wavelets and applications*; RMA20, Springer – Verlag.

[9] http://www.mathworks.com

[10] http://www.wavelet.org

[11] http://www.acoustics.org/index.html