Implementation Of Strategy Based on Auditory Model Based Wavelet Transform Speech Processing on DSP dedicated to Cochlear prosthesis

Amira Derbel, Mohamed Ghorbel, Mounir Samet and Ahmed Ben Hamida
Technologies de l'Information & Electronique Médicale 'TIEM'
TIEM-LETI, ENIS, University of Sfax – Tunisia

ABSTRACT- **The use of auditory models in cochlear implant sound processing strategies aims to improve cochlear implant user perception of speech. This paper presents the implementation of a cochlear implant signal processing strategy on DSP. As for concretization, TMS320C6416 DSP board was chosen. The aim of the new strategy is to o GUbtain more natural sound in CIs by better mimicking the human auditory system. Knowing that, wavelet transform, is characterized by its time-frequency resolution that is similar to the hearing perception procedure, we concern our work to develop a strategy based on auditory model based continuous wavelet transform processing DSP implementation. To gain software flexibility and interactivity, the Assembly programming environment is used. The paper chooses the processing algorithm which is taken to achieve a real-time throughput. This experimentation provides directives guidelines on our work to treat the speech signal sample per sample. Code energies of input samples can be obtained and send one per one in weft to the cochlear implant internal part in duration less than the arrival of the second sample for real time operations.**

Keywords:
Cochlear prosthesis, Auditory model, Wavelet Transform, DSP implementation

I. INTRODUCTION

Communication via speech is one of the essential functions of human beings. Humans possess varied ways to retrieve information from the outside world or to communicate with each other, and the three most important sources of information are speech, images and written text. For many purposes, speech stands out as the most efficient and convenient one.

Around 10% of the population in developed countries suffers from hearing impairment [1].

Cochlear prosthesis systems are devices designed to restore hearing deafness by electrical stimulation of the remaining auditory-nerve fibbers of the cochlea. Typically, such biomedical device includes an external sound processor [2] that transforms microphone picked up signal into command data that would be transmitted into the internal part via radio frequency link. This encoded and transmitted data to an implanted receiver/stimulator would command several current sources dedicated to deliver electrical stimuli for the different nervous sites of the cochlea [3].

Signal processing, in particular, played an important role in the development of different techniques for delivering electrical stimuli from the speech signal.

Several stimulation strategies were proposed during the past few years including simple filtering modules distributed along one considered sounds spectrum. These filters were in the analogical form in the beginning of this biomedical research, and then improved to the numerical form with the emergence of the digital sounds processor s [4]. Other strategies were also numerical and based spectrum analysis including FFT algorithms [4]. As an important time-frequency analysing method, the continuous wavelet transform (CWT) that is similar to that of the hearing perception procedure is proposed as a signal processing method [5.

Basing on human auditory system, our considered auditory spectrum would be divided according to ERB model. With 16-kHz as a sampling frequency, our bank filtering module was composed of 21 filters [6].

This paper presents in its second section, several details concerning the proposed stimulation strategy dedicated to cochlear prostheses apparatus. In that describes temporal approach of filtering stage based algorithm.

The third section achieves the objectives of this paper which are to implement on a TMS320C6416 DSP board, a new speech processing strategy based on auditory system based continuous wavelet filtering module for use in cochlear implant processors and to optimize the time-consuming operations that achieves the analyzing basis in real time.

Finally, the latter section summarizes and draws conclusions.

II. STIMULATION STRATEGY ARCHITECTURE

The cochlear implant prosthesis conception normally includes two main parts: an external unit, the sounds analyzer, and an internal chirurgically implanted unit, the under-skin implant, which are connected by an inductive transcutaneous link, as illustrated in Figure 1.

The external part includes a microphone that picks up the sound and a digital signal processing unit that convert the sound into a specific electronic code based on the adopted stimulation strategy. The internal part uses the demodulated code to generate electrical stimuli that drive the electrode array which is usually placed close to the residual nerve fibres inside the cochlea.

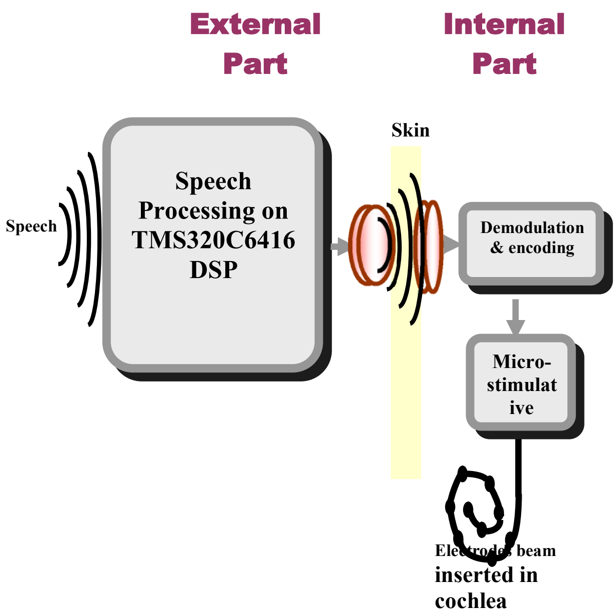


Fig.1. Basic Architecture of Cochlear Prosthesis

The following architecture (figure 2) gives several details concerning the proposed stimulation strategy dedicated to cochlear prostheses apparatus. We distinguish the main part including filtering stage algorithm, a second part for energy estimation to be compressed with logarithmic compression and forwarded as stimulation level to the inner part by a biphasic current pulses [6].

Our suggested filtering stage use Continuous Wavelet Transform (CWT) as a signal processing method.

We consider the case of a 21 stimulation electrodes prosthesis which requires twenty one filters bank distributed on the audible spectrum limited to 8000Hz.

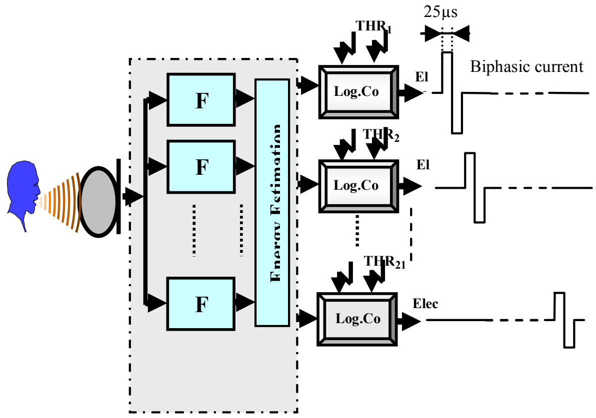


Fig.2. Bloc architecture of the proposed stimulation strategy

A. Continuous Wavelet Transform (CWT)

Basing on the non linear cutting of auditory system, continuous wavelet transform is chosen as a suitable form of wavelet transform.

The temporal approach stimulation was based on 'CWT' algorithm. The CWT of a temporal signal s(t) is given by this equation:

$$CWT_s(a,b) = \int_{-\infty}^{+\infty} \psi^*_{a,b}(t)\, s(t)\, dt \qquad (1)$$

This analysis used a family of functions $\psi_{a,b}(t)$ based on a mother wavelet ψ [11]. $\psi_{a,b}(t)$ is given by the following equation:

$$\psi_{a,b}(t)=\frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right)=\frac{1}{\sqrt{a}}\tilde{\psi}\left(\frac{t-b}{a}\right)e^{j\omega t} \quad a,b\in\Re \qquad (2)$$

Where 'a' is a dilatation parameter and 'b' is translation parameter.

In fact, we can use different mother wavelets such as Morlet and Marr wavelet [7]. Due to easy selection of its center frequency, and its exponential form, Morlet wavelet is chosen as the mother wavelet to decompose the audible spectrum for the cochlear bank filter. Morlet wavelet envelope is defined as:

$$\tilde{\psi}(t) = \frac{1}{\sqrt{2\pi}} e^{\frac{-t^2}{2}} \qquad (3)$$

Basically, the idea of Auditory Model Based Wavelet Transform is to make the temporal envelope of the mother time varying according to the characteristics of the target signal by introducing the elementary paving. This last will adjust the Morlet wavelet to emulate the ERB model of the auditory system in the following way:

$$\tilde{\psi}_{ERB}(t) = \frac{1}{\sigma}\tilde{\psi}\left(\frac{t}{\sigma}\right) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-t^2}{2\sigma^2}} \qquad (4)$$

The first σ is just the scaling factor to insure the energy is the same for every mother wavelet and the second σ adjusts the envelope of the mother wavelet ψ(t) without adjusting its center frequency.

Considering a constant surface of elementary paving in time frequency domain, we can prove that:

$$\sigma_t * \sigma_f = \frac{1}{2\pi} \qquad (5)$$

The result of a mathematical demonstration to find the relation enters the temporal elementary paving σ and the band width B in each corresponding frequency can be summarized as:

$$\sigma = \sigma_t = \frac{\sqrt{0.7}}{\pi B} \qquad (6)$$

In order to share the considered audible spectrum and fix the width of each band corresponding to ERB model [8], we can use two different function as Munich and Cambridge function. The first is expressed by this BM equation:

$$B_M = 25 + 75(1 + 1.4 f_0^2)^{0.69} \qquad (7)$$

and the second is expressed by this B_C equation:

$$B_C = 24.7(1 + 4.3 f_0) \qquad (8)$$

Where f_0 is the band frequency center in KHz.

A comparative study between these two approximate functions showed that the Munich critical bands loom with ERB critical bands than the Cambridge critical bands [6]. Thereafter, we are interested only in the development of an optimal wavelet base using Munich model for speech signal parameterization.

The final expression of the Munich based mother wavelet is given by:

$$\psi_{ERB}(t) = \sqrt{\frac{\pi}{1.4}}(25 + 75(1 + 1.4 f_0^2)^{0.69}) * e^{\frac{-[25+75(1+1.4 f_0^2)^{0.69}]^2 \pi^2 t^2}{1.4}} e^{-i2\pi f_0 t} \qquad (9)$$

Following the stimulation mode, the number of stimulation channel (or electrodes) used for cochlear prostheses and pathologic constraints of the patient, the division of the audible spectrum could be done in various forms, such as linear subdivision, logarithmic subdivision, anti-logarithmic subdivision…etc.

In our case, we took the ERB critical bands as a sharing spectrum model and a 16 KHz sampling frequency. We consider 21 bands bench filters, then 21 parameters to be extracted.

III. DSP IMPLEMENTATION

Still, DSPs are a very common and efficient method of processing real-time data [9]. All the Clarion devices are based on Texas Instruments TMS320 family DSP's. The latter presents a class of hardware devices that fall somewhere between an ASIC and a PC in terms of performance and design complexity [10]. They can be programmed with either assembly code or the C programming language, which is one of the platform's distinct advantages.

In this section, we describe the implementation of our proposed Strategy for cochlear prosthesis using a TMS320C6416 digital signal processor (DSP) board.

In this implementation, MATLAB is used as a procedural routine to operate needed tools represented as input data, these software algorithms were designed to resemble the hardware algorithms as closely as possible.

Concerning our continuous wavelet transform based strategy implementation; we need to operate as input data, in discrete time, the transfer function and the signal input that we have process, and in real time, we need only the transfer function as input data.

A. TMS320C6416 DSP Board Overview

For the scope of this project we used the TMS320C6416 digital signal processor (DSP) board. C6416 is a recent version of the TMS320 family of DSPs provided by Texas Instrument. It uses VelociTI, a high-performance very long-instruction-word (VLIW) architecture. The C6416 register file extends this by additionally supporting packed 8-bit types and 64-bit fixed-point data types. There are two general-purpose register files, A and B, which have 32-bit registers in total [10].

The TMS320C6000 family C compiler is a notable exception. The point here is that the memory consumption and the cycle time is not as important to the micro-processor programmer. The C language is usually a simpler way to develop programs even though the short assembly code does support this claim.

In order to create a high performance C6416 implementation design, a designer has three methods for implementation algorithm. We are programmed with C programming language then linear assembly code, next optimizer assembly code which is one of the platform's distinct advantages.

B. Discrete Time Implementation

One of the initial goals of this project was to implement our proposed strategy in discrete time.

The processing of the CWT, such realized by Matlab, consists at making the scalar product of the fixed input signal with the wavelet mother Psi during its moving on the right, for each value of translation parameter b; this is repeated for each value of dilatation parameter a. This phenomenon is presented in figure 3.

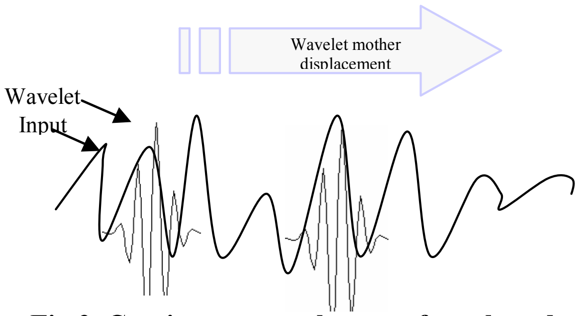


Fig.3. Continuous wavelet transform based processing for a fixed dilatation parameter "a"

This process is not appropriate for a processing in real time; since the signal is variable for each sample. In fact, this new processing method gives the same result with more of simplicity.

It consists at leaving the wavelet function fixed, while making move the input signal on the left for each sample; this is repeated for each value of dilatation parameter a. like illustrated in the figure 4.

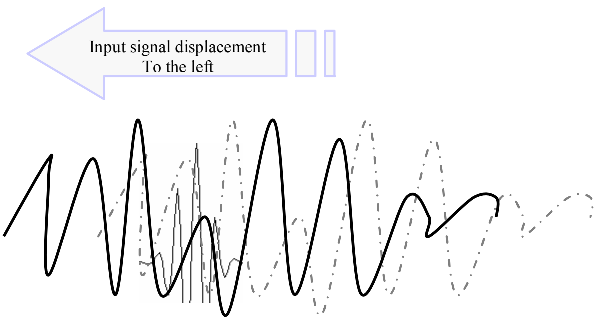


Fig.4. Continuous wavelet transform based processing dedicated to real time for a fixed dilatation parameter "a"

With two data form sends in memory: transfer function and sampling speech signal, we are implemented the continuous wavelet transform based algorithm described above, with language C then with Linear Assembly code using the TMS320C6416.

The C6416 continuous wavelet transform based algorithm implementation is presented in figure 14. In this implementation, for each sample input $x[n] = x_0(nT)$ placed at its specified level in an "N" sized buffer that initialized at zero, the output for each filter band will be computed for each dilatation parameter "a". It is given by:

$$Y_a[n] = x[n] * Psi_a[n] \qquad (10)$$

Where: a is the number of the filter band $1 \le a \le 21$; $Psi_a[n]$ is the discrete transfer function (mother wavelet) for each filter band.

In this way our filters bank based on CWT can be looked like an FIR filters bank. Therefore the complexity of the algorithm to be implemented on the DSP will be significantly reduced.

So, the result is resumed to have twenty one energies for input samples. Figure 5 shows CWT based algorithm implementation diagram.

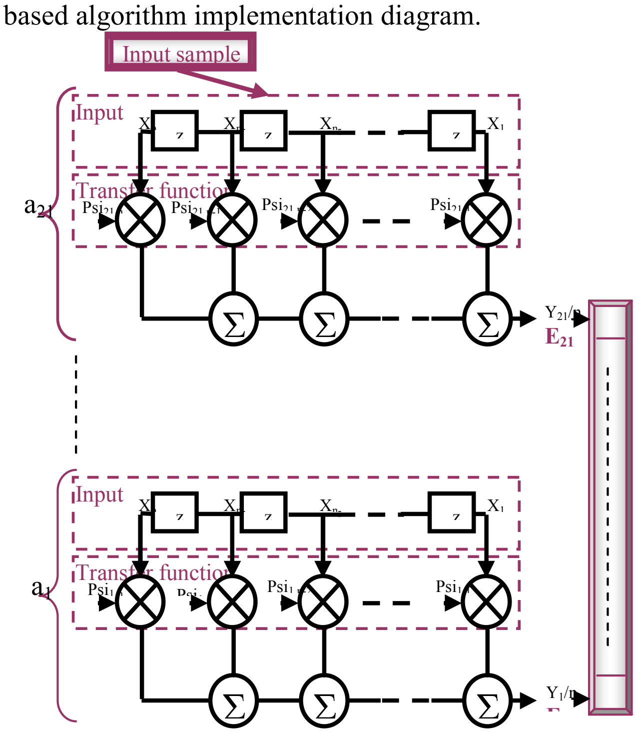
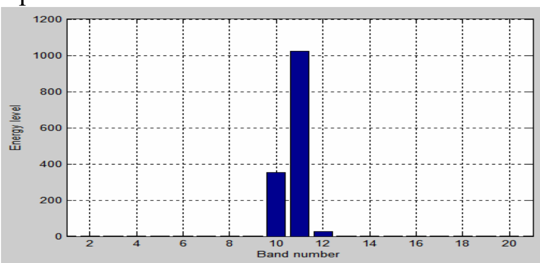


Fig.5. Continuous wavelet transform based algorithm implementation diagram

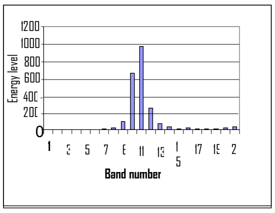
The two implementations are successfully tested using harmonic signal and then real speech signals extracted from TIMIT database. Testing results are similar to those based on Matlab implementation [6].

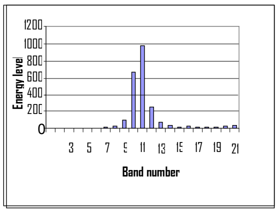
Some comparisons are acquired between parameterizations with Matlab implementation, C6416 implementation based on Langage C and Linear Assembly for some signals.

The following figures take the cases of a harmonic at 1.3 kHz, a composite harmonic 550 Hz + 4000 Hz, and a phoneme "aa" extracted from TIMIT database successively; and presents a comparison between parameterizations with (a) Matlab implementation, (b) Langage C and (c) Linear Assembly based TMSC6416 implementation.



(a)





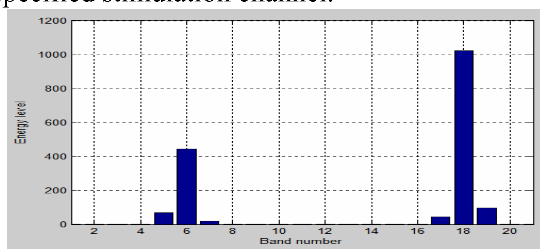
(b) (c)

Fig.6. A 1.3 kHz harmonic parameterizations with (a) Matlab implementation, (b) Langage C and (c) Linear Assembly based TMS C6416 implementation.

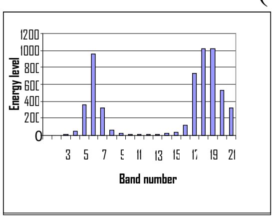
The visual measure proposed with this stimulation algorithm, the graphical spectrogram, was formed by a matrix having variable columns proportional to the energy levels calculated at each filter band. The displayed columns indicated affected frequency bands and gave an estimate of electrical stimulating pulses to be delivered to cochlea's nervous endings.

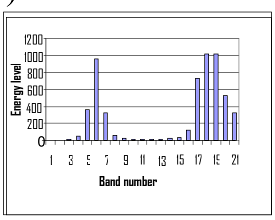
Stimulating pulse generation, in the assigned stimulation channels of the cochlear implants, was estimated thanks to the processed band energies E1, E2, …., E21.

Each band energy 'Ei' would be affected to one specified stimulation channel.



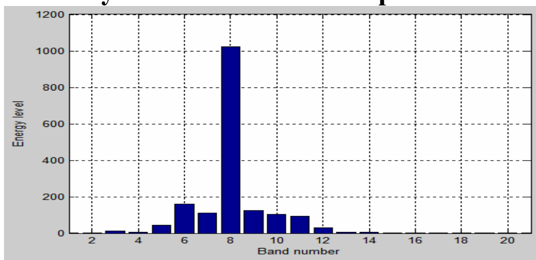
(a)



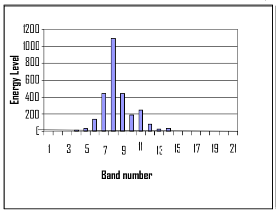


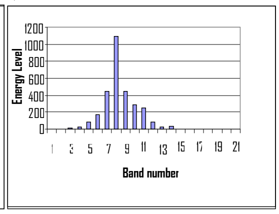
(b) (c)

Fig.7. A 550 Hz + 4000 Hz composite harmonic parameterizations with (a) Matlab implementation, (b) Langage C and (c) Linear Assembly based TMSC6416 implementation.



(a)





(b) (c)

Fig.8. A a phoneme "aa" extracted from TIMIT parameterizations with (a) Matlab implementation, (b) Langage C and (c) Linear Assembly based TMSC6416 implementation.

The experimental result of this analysis algorithm were satisfactory, we remark in this section, nearly similar results for each case with C language and Linear Assembly based C6416 implementation compared to results obtained with Matlab based implementation [6]. Usually harmonic energies are located in their corresponding frequency. The minor difference is caused to the fixed edge of input and output data necessary in C6416 based implementation (fixed point).

In summary, the design conversion resulted in the following performance: The Linear Assembly Decoder based C6416 implementation required 32.50 µsec of total processing time (this includes all of the outputs) compared to the 47.28 µsec required for the extraction output data by the C language based implementation. This enhancement equates to 31.26% of the original DSP processing time. The performance results are shown in figure 9.

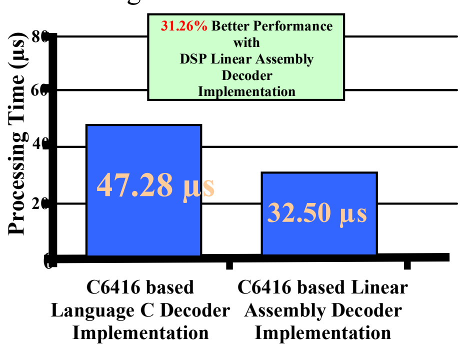


Fig.9. Performance C6416 based Assembly Decoder implementation

IV. CONCLUSION

In this study we were interested in cochlear implant research wish is actually a developing field regrouping researchers in different disciplines. All researchers are nowadays looking for new flexible techniques of stimulation and rehabilitation.

We purpose an implementation in DSP of auditory model based algorithm using continuous wavelet transform speech processing already validated by Matlab for our new stimulation strategy.

We implemented the proposed strategy algorithm on a TMS320C6416 DSP board for acoustic communications. Based on the program profiling, we identified the time-consuming operations of the strategy algorithms in discrete time with language C and then Linear Assembly. The duration of an implementation based on C language and Linear Assembly are successively 47.28 µs and 32.50 µs. nearly similar results were obtained for each case with C language and Linear Assembly based implementation to results obtained with Matlab based implementation.

The C language and Linear Assembly language based implementation in discrete time are successfully tested using harmonic signal and then real speech signals extracted from TIMIT database.

This current implementation does not meet the real-time operation requirements yet, but applicable to construct a real-time processing.

REFERENCES

[1] S.U.Ay,F.G.Zeng, and B.J.Sheu, 1997, “Hearing with bionic ears: Speech processing strategies for cochlear implants devices,” Circuits Devices, pp. 18-23.

[2] Fathi KALLEL, Ahmed Ben HAMIDA and Nouri MASMOUDI, ‘FFT Based Speech Processing Dedicated to a Fully Programmable Stimulation Strategy for Cochlear Prostheses’ 10th WMSCI conference, July 16-19, 2006 . Orlando, Florida, USA.

[3] CHAOU. M : Etude d'un amplificateur commandé en gain et enbande passante. Contribution à la conception d'une prothèse auditive, Mémoire de D.E.A à l'E.N.I.S Dec 1998

[4] A. BEN HAMIDA, "Etude et implantation d'algorithmes de traitement de la parole dédié à la stimulation électrique par la prothèse cochléaire", Thèse de doctorat E.N.I.S, 1998.

[5] F. S Yang, Wavelet Analysis and Its Application in Engeneering. Beijing, China: Tsinghua Univ., 1995, ch. 1pp.50-53.

[6] Amira, D., Fathi, K., Ahmed B., and Noureddine, E., 2007. Wavelet-Based Parametrisation of Speech Signal Dedicated to Cochlear Prosthesis, Fourth International Multi-Conference on Systems, Signals and Devices.

[7] Daubechies, I., 1992. Ten lectures on Wavelets, SIAM, Philadelphia, PA.

[8] Eddington, D., Dobelle, W., Brachman, D., Mladevosky, M., Parkin, J., 1978. Auditory prosthesis research using multiple intracochlear stimulation in man, Annals of Otology,Rhinology and Laryngology 87(53) 1-39.

[9] Moore, M., 1995. A DSP-Based Real-Time Image Processing System, Proceeding of the 6 th International Conference on Signal Processing Applications and Technology, Boston, MA.

[10] Texas Instruments, 1999.Inc, TMS320C4X User’s Guide, Houston, TX.

[11] Yang, F.S., 1995.Wavelet Analysis and Its Application in Engineering. Beijing, China, Tsinghua Univ, chapter1, pp.50-53.