THE APPLICATION AND COMPUTER SIMULATION OF MULTI-CHANNEL COCHLEAR IMPLANT BASED ON All PHASE DFT FILTER

Lan Tian[†,††] *Zhengxin Hou*[††] *Zhimin Pang*[†]

[†]School of Information Science and Engineering of Shandong University, Jinan 250100, China;
[††]School of Electronic and Information Engineering, Tianjin University, Tianjin 300072, China

ABSTRACT

Band-pass filter bank is an essential part in cochlear implant (CI). Its characteristics are important for synthesized sound quality because the outputs of filter bank contain a lot of fine structure cues which need to be encoded and transmitted to implanted electrode and then to stimulate auditory nerves. All phase DFT (APDFT) filter is a novel and high efficient digital filter, which possesses concurrent merits, such as zero-phase (or all-phase), abrupt cut-off characteristic. This paper presents the principle and the design method of APDFT filter bank and applies this filter into CI signal processing. Under the same acoustic simulation conditions, based on the continuous interleaved sampling (CIS) strategy, we analyze and evaluate the simulation results. Comparing with classical Butterworth filter bank adopted in CI processors, some objective waveform and spectra show that the output signal from APDFT filter bank is much closer to the original sound. For the normal-hearing (NH) listener, the synthetic speech and music by the APDFT filter bank have higher quality, and perceptual tests on six NH listeners primarily verify the improvement of intelligibility in noise. Thus, the APDFT filter could potentially improve the hearing quality for CI user.

Key words: digital filter, speech synthesis, electronic cochlea (cochlear implant), spectral analysis

1. INTRODUCTION

Cochlear implant is a therapeutic device to restore hearing ability for profoundly deaf people. Most deaf individuals have lost the ability to translate sound into electrical signals which are normally present on the fibers of the auditory nerve. Because these electrical signals are the inputs to the brain that result in sound sensation, cochlear implants deliver electric stimuli to these fibers that mimic the patterns present in a normal-hearing ear. Current cochlear implants are comprised of external unit and internal unit. The external unit includes microphone, worn speech processor and transmitter. The internal unit consists of receiver-transmitter and electrodes array embedded in cochlear. The sound is picked up by microphone and processed by processor that analyzes sound and encodes the results into instructions to be received by the implanted receiver. By RF signal, the transmission system transmits the instructions to the implanted receiver, and internal circuit decodes the data and then generates electrical stimulus and drives electrode array to stimulate auditory nerves. In the processor of multi-channel CI devices, band-pass filter bank is an essential analyzing unit and the outputs contain a lot of fine frequency and temporal cues which need to be encoded and transmitted to the embedded electrodes. So we suppose that the characteristic of the filter might be correlative with the quality of encoded instructions for the synthesized sound.

The all phase DFT (i.e. APDFT) is a novel FIR filter which possesses concurrent merits of the conventional windowing and the frequency sampling filter, and can be used to design zero-phase FIR filter and strictly complementary sub-band filtering [1]. Studies [3,7,8] suggest that phase information, or fine temporal structure, is important for perceiving music and tonal languages. For the sake of improving the synthetic sound quality and compensating phase cues, we introduce APDFT filter into the current or next generation CI processor. Because the sound heard by the real CI users might be subject to many factors, such as the limited channel numbers, existence of electric field and the differences of the remains of auditory nerve for each patient, the sound heard by CI user must be different from the normal sound [2]. However, in the interest of comparing effect on the signal processing level, we use computer simulation to synthesize the sound heard by CI users. The perception tests from NH listeners can initially evaluate the experimental effects. Our research results show that the synthetic sound quality improvement by the APDFT filter bank is obviously audible for NH listeners and is distinguishable from the corresponding spectra.

This paper is organized as follows. In the section two, we briefly introduce the CIS strategy and experiment procedures in computer simulation. The principle and design method of the APDFT filter are described in section three, and the features of the filters are compared and analyzed by the frequency response. In section four, we give the simulation results and related analysis and evaluation. Finally, the section five is the conclusion and prospect.

2. COCHLEAR IMPLANT AND CIS STRATEGY

The main function of CI signal processor is to decompose the input signal into its frequency components, similar to a healthy cochlea. The auditory nerve fibers are described as having a tonotopic organization, exploiting the place mechanism for coding frequencies. So, the CI signal processor is responsible for decomposing input signal into a limited number of frequency bands or channels and delivering the filtered signals to the appropriate electrodes.

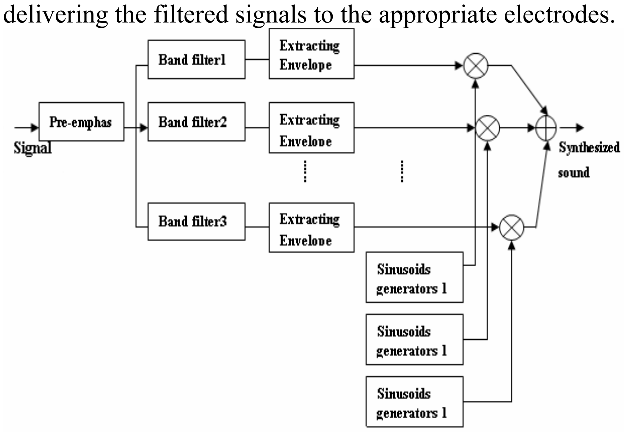


Fig.1 CIS Simulation Model

The Continuous Interleaved Sampler (CIS) strategy was proposed by Wilson in 1991[4]. The basic process is shown in Fig.1. First, the signal is pre-emphasized and then passed through a band-pass filter bank. The envelopes of the filtered waveforms are extracted by full-wave rectification and low-pass filtering (typically with 200 or 400 Hz cutoff frequency). The envelope outputs are finally compressed and then used to modulate biphasic pulses. The pulses are sequential, non-overlapping and interleaved-firing. That means only one electrode is stimulated at a time. So, CIS strategy meets minimizing electrodes interactions.

In our study, the same simulation model is used for the two designs of filter bank, i.e. six-order Butterworth and APDFT. In current CI processors, the six-order Butterworth band-pass filter bank is typically adopted. Logarithmic frequency spacing is arranged for every frequency band [7]. In order to compare the performances for the two kinds of filter bank, the same number of channel is selected, such as 8 or 16. 8 channel can reach the best asymptotic performance [5]. Other parameters are set the same for the two filters design. Signals were first processed through a 3-dB/octave pre-emphasis filter, and then respectively passed into the two kinds of filter bank. The envelope of the filter bank output was extracted by full-wave rectification, and low-pass filtering with a 200Hz (or 400 Hz) cutoff frequency. In the simulation of the synthesized sound, sinusoids wave was generated with amplitudes equal to the root mean-square (rms) energy of the envelopes computed every 4ms temporal interleaved, and carrier frequencies equal to the center frequencies of the band-pass filters. The phases of the sinusoids can be obtained from the Hilbert transform of the band pass signal, i.e. $\phi(t)$ components [7]. The sinusoids of each band were finally summed and the amplitude of the synthesized speech segment was adjusted to have the same rms value as the original speech segment [5].

3 THE PRINCIPLE OF APDFT FILTER

The APDFT concept is presented based on the theory below. For the point x(n) in a discrete sequence, there exist and only exist N-dimensional vectors that contains x(n) and have different intercept phases:

$$\left.\begin{aligned} X_0 &= [\, x(n), x(n+1), \ldots, x(n+N-2), x(n+N-1)\,]^T \\ X_1 &= Z^{-1}X_0 = [\, x(n-1), x(n), \ldots, x(n+N-3), x(n+N-2)\,]^T \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ X_{n-1} &= Z^{-N+1}X_0 = [\, x(n-N+1), x(n-N+2), \ldots, x(n-1), x(n)\,]^T \end{aligned}\right\} \quad (1)$$

Where Z^{-1} is the delay operator. Obviously, x(n) is the intersection of those vectors, i.e.

$$x(n) = X_0 \cap X_1 \cap \cdots \cap X_{N-1}$$

According to the regular representation of data matrices, the all phase data matrix of x(n) is defined as

$$X_{-N+1,0}(n) = [X_0, X_1, \ldots, X_{N-1}]$$

All the column vectors $\{X_l, l = 0,1,\ldots,N-1\}$ are composed of an all phase data space of x(n). Applying DFT/IDFT filtering to each column vector of $X_{-N+1,0}(n)$ yields

$$\begin{aligned} Y_l(i) &= \sum_{k=0}^{n-1} W_N^{-1}(i,k)\left[F_N(k)\sum_{j=0}^{N-1} W_N(k,j)X_l(j)\right] \\ &= \sum_{j=0}^{N-1} H_N(i,j)X_l(j) \end{aligned} \quad (2)$$

where $i, l = 0,1,\ldots,N-1$,

and $$H_N(i,j) = \sum_{i=0}^{N-1} W_N^{-1}(i,k)W_N(k,j)F_N(k) \quad (3)$$

$$i, j = 0,1,\ldots,N-1$$

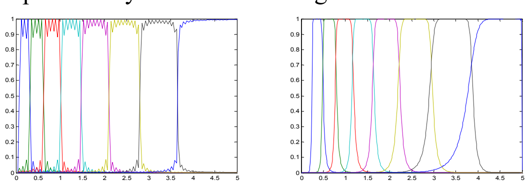
Here W_N and W_N^{-1} are the N×N DFT and IDFT transform matrix respectively, and F_N is the N-dimension expected response vector for DFT filtering. Obviously, it would yield N different phase values corresponding to the same point x(n) : $Y_i(l), l = 0,1,\ldots,N-1$. To eliminate the multiple filtering values caused by different phases and thereby the block effect, we take the mean of the values as the filtering output:

$$y(n) = \frac{1}{N}\sum_{i=0}^{N-1} Y_i(l) \tag{4}$$

In [6], it has been proved that, when the DFT response vector F_N is real symmetric, the all phase DFT filtering can be implemented by a zero-phase digital filter, or called the APDFT filter, whose impulse response $h_N(n)$ can be calculated by the symmetric extension IDFT of F_N respecting the origin and then weighted with a triangle window on it, i.e.

$$\begin{cases} h_N(n) = \dfrac{N-n}{N} IDFT\left[F_N(k)\right] \\ \qquad n, k = 0, 1, \ldots, N-1 \\ \qquad h_N(-n) = h_N(n) \end{cases} \tag{5}$$

The definition of the all phase DFT filter and the related equations are provided above. Computer simulation [6] shows that it has the advantage of the frequency sampling in overall filter performance, and can be used to realize strictly complementary sub-band filtering.



(a) (b)

Fig.2 (a) The frequency response of APDFT band pass filter bank
(b) The frequency response of Butterworth band pass filter bank
(Frequency dividing: 250 500/500 875/875 1150/1150 1450/1450 2000/2000 2600/2600 3800/3800 5512)

Fig.2 shows the filter bank frequency-responses for (a) APDFT and (b) traditional Butterworth. The APDFT filter is linear phase, the zero-phase characteristic makes filter output almost without time delay (i.e. phase distortion), and keeps high fidelity within every frequency band and keeps the phase cues consistency between neighboring channels that would be present in normal sound wave architecture. After full-wave rectification and LP filtering, the extracting envelop still contains more fine temporal cues. While the Butterworth filter is non-linear phase and lack of consistency of phase cues within one channel and between channels, phase distortions certainly exist. The case within one channel is shown in Fig.3.

Depending on the precise of cut-off frequencies and the computational complexity in CI simulation, the parameter N for APDFT filter may be set as 128. For all testing data, the sound signal is sampled at 1,1025Hz and 16bit. The central frequencies of 8-channel Butterworth filter bank are as follow: [375;687;1012;1300;1726;2300;3200;4660]. The corresponding APDFT filter bank response vector F matrix is as below.

```
Sampling rate Fs=11025Hz; APDFT Filter's N=128; channel=8
F1=[zeros(1,2) ones(1,3) zeros(1,119) ones(1,3) zeros(1,1)];
F2=[zeros(1,5) ones(1,3) zeros(1,113) ones(1,3) zeros(1,4)];
F3=[zeros(1,8) ones(1,3) zeros(1,107) ones(1,3) zeros(1,7)];
F4=[zeros(1,11) ones(1,4) zeros(1,99) ones(1,4) zeros(1,10)];
F5=[zeros(1,15) ones(1,6) zeros(1,87) ones(1,6) zeros(1,14)];
F6=[zeros(1,21) ones(1,6) zeros(1,75) ones(1,6) zeros(1,20)];
F7=[zeros(1,27) ones(1,12) zeros(1,51) ones(1,12) zeros(1,26)];
F8=[zeros(1,39) ones(1,51) zeros(1,38)]
```

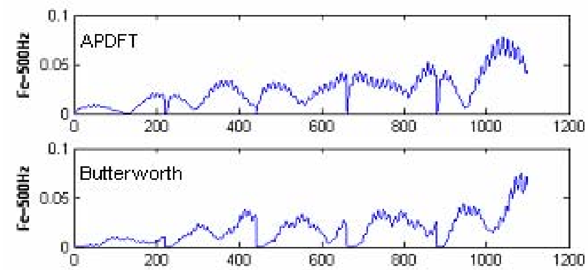


Fig.3 Extracting envelop on a low frequency band (Fc=500Hz) for a music patch after band pass filter + full-wave rectification + LP filtering

4. SIMULATION RESULTS

Based on the same CIS simulation method, 20 English sentences and 20 Chinese syllables are synthesized and analyzed, and then the corresponding acoustic spectra are evaluated. Because the wide-band spectra are difficult to display the fine frequency information, we adopt the narrow-band spectra. Some results for an English sentence sample are provided from Fig.4.1 to Fig.4.3. The spectra are for the original, APDFT filter bank outputs with pre-emphasis, Butterworth filter bank outputs with pre-emphasis respectively. Comparing Fig4.2 and Fig.4.3 with Fig4.1, It was found that the waveform of Fig.4.2 is much closer to the original signal of Fig.4.1 and NH listeners have much higher quality perception for Fig4.2. It was also found that the output of APDFT filter bank without pre-emphasis is almost the same quality as the original sound. For the output of Butterworth filter bank, the decreased spectra quality can be discerned on the spectrum because some horizontal "white bar" (i.e. losing frequency ingredient) are presented, and the decreased perception can be perceived easily. Based on the CIS process strategy, the synthetic sound spectra by the two filters are shown in Fig.4.4. It shows samples of time-domain waveforms and corresponding spectra for a piece of music using 16 channels. It can be found that (b) is much closer to (a) than (c), and its better perception is distinguishable. At the low and high frequency area, the APDFT method reconstructs more precise frequency structure. In addition, we also found that the APDFT method has better intelligibility in noise for Chinese syllables recognition. Six NH listeners are tested for the Mandarin ten monosyllable words when adding 10 dB

Gauss white noise. With the stimulation rate varying from 2ms to15ms, the average recognition rate for the APDFT method is 15% higher than the traditional method.

In summary, the simulation results show that the APDFT filter design algorithm can extract much more fine frequency and phase cues from each channel than the classical filtering method and the synthetic sound performance improvement could be potentially helpful for CI user in noise.

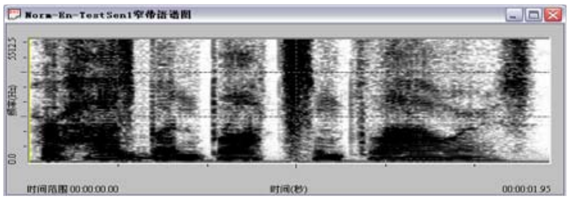


Fig.4.1 Spectra of the original speech

Fig.4.2 Spectra of the pre-emphasized APDFT filter bank output (ch=8)

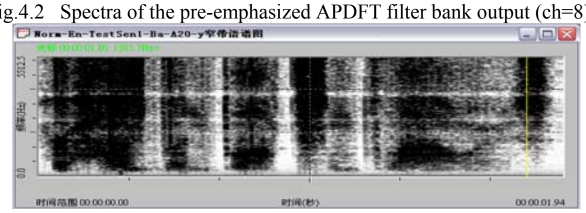


Fig.4.3 Spectra of the pre-emphasized Butterworth filter output (ch=8)

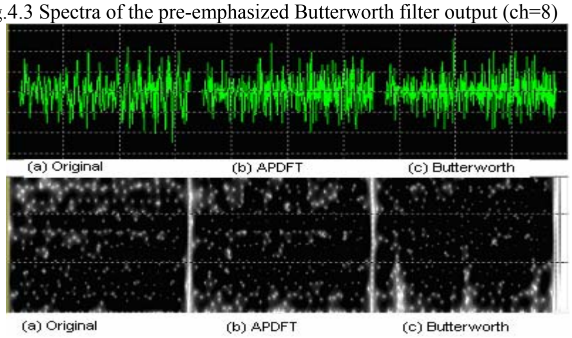


Fig 4.4 The waveforms and spectra of a music patch (ch=16)

5. CONCLUSION AND DISCUSSION

In this paper, we proposed the APDFT filter designing method and applied it to the signal preprocessing of CI. Based on the same experimental conditions and CIS processing strategy, the improvement contribution of the APDFT filter bank to CI vocoder is initially verified by NH listener and objective spectra. The high fidelity output signal from the APDFT filter bank and the perceptible improvement results on the corresponding simulated synthetic sound are obtained. But we note that the perception results are far away from the obvious improvement at the filter output level. Compared to the traditional Butterworth filter design method, the reception sound effect is better assuredly and we can conclude that APDFT filter is able to extract and deliver more accurate and fine spectral and temporal features because this kind of filter possesses the overall outstanding filtering characteristics. With the help of the APDFT fast algorithms [1], computation complexity and memory space will not be a practical problem. However, how to encode the fine structure cues, especially for phase information, to the hearing nerve fibers to obtain abundant hearing perception (such as music melody), will be more challenging.

From this study, we can conclude that more fine acoustic structure for improving the perception performance of cochlear implant user will mainly depend on the novel encoding and processing strategy that involve in the electrodes simulation. More complex processing strategies and new type electrode array [8], such as ‘virtual channel’ strategy and laser electrodes and experiments with CI users should be combined in our future studies to further verify APDFT filter improving effect.

6. ACKNOWLEDGEMENT

This paper was partially supported by the grant from China CSC committee and Shandong University. We are grateful to Prof. P. Loizou, Yang Liu, and Dr. Yi Hu, Ning Li at the University of Texas at Dallas (UTD) for constructive help when the first author visited at UTD.

7. REFERENCES

[1] Z-X. Hou, Z-H Wang, Y. Yang. Design and implementation of all phase DFT digital filter. ACTA Electronic SINICA, 31(4), pp.539-542, 2003.
[2] P. Loizou. Mimicking the human ear. IEEE Signal Processing Magazine, 15(5), pp.101-130, 1998.
[3] K-B Nie, A. Barco, and F-G Zeng. Spectral and temporal cues in cochlear implant speech perception. Ear & Hearing, 27(2), pp.208-217, 2006.
[4] B. Wilson, C. Finley, D. Lawson, R. Wolford, D. Eddington, and W. Rabinowitz. Better speech recognition with cochlear implants. Nature, 352(7), pp.236-238, 1991.
[5] P. Loizou, M. Dorman, and Z. Tu. On the number of channels needed to understand speech. Journal of Acoustics Society of American, 106(4), pp. 2097-2103, 1999.
[6] Z-X. Hou. The all phase DFT filter. Digital Signal Processing Workshop, Proceedings of 2002 IEEE 10th, 2002
[7] K-B Nie, G.Stickney, and F-G. zeng. Encoding frequency modulation to improve cochlear implant performance in noise. IEEE Trans. Biomed. Eng., 51(1), pp.64-73, 2005
[8] F-G. Zeng. Trends in Cochlear implant. Trends In Amplification, 8(1), pp.1-34, 2004