Adaptive Sample Reduction Techniques for Continuous Interleaved Sampling Strategy in Multi Channel Cochlear Implant

Sonia Tabibi
Department of Biomedical Engineering,
Science and Research Branch, Islamic Azad University,
Tehran, Iran
e-mail: sonia1362@yahoo.com

Hamed Sadjedi
Technical and Engineering Faculty,
Shahed University,
Tehran, Iran
e-mail: sadjedi@shahed.ac.ir

Abstract—**CIS (Continuous Interleaved Sampling) strategy in which electrodes stimulus is determined based on extracted information from amplitude of different bands of input signal, is a basis in processing methods in cochlear implant systems. In this paper besides implementing this strategy, methods for improving sample reduction and computational efficiency are also presented. In this case adaptive method according to processing parameters and an input signal has been suggested that due to the use of fewer samples to synthesize the signal is better than the other methods and is significantly efficient.**

Keywords-cochlear implant; CIS strategy; amplitude modulation; signal processing; sample reduction

I. INTRODUCTION

Cochlear implant has been accepted as the treatment of people with profound deafness [1,2]. Although cochlear implant technology has improved, its speech processor is still one of the effective parts in system performance because of determining electrodes stimulus [3,4]. Speech processing in cochlear implant systems often based on amplitude information such as CIS (Continuous Interleaved Sampling) method [4,5]. Among processing strategies which are based on filter bank, CIS is one of the famous and effective methods and is used as the default strategy in cochlear implant systems [6].

In this method a signal is divided into some frequency band with a bank of bandpass filters. Then envelopes of the filtered signals are calculated for modulating trains of biphasic electrical pulses [4,6]. These pulses are delivered to the electrodes in a nonsimultaneous fashion [4]. The envelope signals extracted from the bandpass filters are compressed with a nonlinear mapping function prior to the modulation, in order to map the wide dynamic range of input signal (sound) in the environment into the narrow dynamic range of electrically evoked hearing [6].

The rate at which the pulses are delivered to the electrodes has been found to have a major effect on speech recognition [4]. Increasing the rate causes an increase in time accuracy but on the other hand increasing the rate or reducing the time interval between two successive stimuli, increase the possibility of applying stimulating pulses in a recovery phase; so excitation of neural fibers does not happen.

In CIS strategy if we reduce the number of samples of envelope signals which are modulated trains of pulses, the amount of information and the number of stimulating pulses will decrease. Therefore, employing sample reduction techniques for speech signal are profitable. In this paper these techniques are designed with the aim of transferring sufficient and useful information to patients. Choosing techniques is done through questionnaire and by comparing the spectrograms of a synthesized signal and an original signal.

II. MODELS & METHODS

In multi channel cochlear implant systems, main signal processing strategy was based on filter bank in which a signal was filtered into a number of frequency bands with a bank of bandpass filters. In CIS method after passing through a bank of filters, envelopes of the signals were extracted by full wave rectification and lowpass filtering. Then the envelopes were compressed to fit the patient's dynamic range and modulated with biphasic pulses. The compression function, maps the input acoustic range to the electrical range between threshold (THR) and maximum comfortable level (MCL) [4]. MCL was determined by increasing biphasic current pulses until the onset of discomfort, then reducing it a small amount until the loudness was acceptable [7]. The starting level for threshold detection was set to approximately a quarter of the patient's dynamic range [8].

The standard equation of the mapping function implemented for the CIS strategy was [9]:

$$Y = AX^{p} + B, \tag{1}$$

Where X was the acoustic amplitude (output of envelope detector), A and B were constants and Y was the electrical amplitude (output of compression function) [4]. The parameter p defined the shape of mapping function; when $p = 0.001$, the map function approximated the logarithmic function used in many of sound processing systems [9]. *A* and *B* were computed as follows in which $[X_{min}, X_{max}]$ was the input acoustic range [4]:

$$A = \frac{MCL - THR}{X_{max}^{p} - X_{min}^{p}}, \tag{2}$$

$$B = THR - AX_{min}^{p}, \tag{3}$$

The values THR and MCL may vary from electrode to electrode [4], but in this paper THR and MCL were set to 20 and 100dB.

978-1-4244-9709-6/11/$26.00 © 2011 IEEE

Number of filters used in the filter bank was eight because patients could not functionally use more than eight frequency bands [10-12]. These filters were tenth-order Butterworth bandpass filters, with frequencies Mel spaced [13] between 80 and 8800 Hz [14]. The lowpass filters that were used to extract amplitude modulation were fourth-order Butterworth filters. For synthesizing, all subband signals were summed to form an output signal [3,10,14].

Same as many other researches [10,15-18], benchmarks which were used in this study were “Choice” and “Hood” signals. For evaluating the CIS method, the spectrograms of the synthesized signal and the original signal were compared. In addition, this was done by comparing a synthesized sound and an input sound. In this case twenty individuals, six male and fourteen female adults from 25 to 40 years old participated in experiment and listened to the sounds. All the individuals had normal hearing and speech signal evaluation was done through questionnaire.

CIS method was programmed and run using Dell computer with Intel(R) Core(TM) 2 Duo CPU T9300 @ 2.50 GHz, 3 GB RAM and 320 GB Hard disk. All the procedures including filtering, amplitude modulation and signal synthesis were implemented in the Matlab (R2008a) environment.

III. DIFFERENT SAMPLE REDUCTION RATES

As we mentioned before, employing methods to reduce the number of samples and computational efficiency seems to be useful. For this purpose the amplitude modulation signals of different bands downsampled with the desired rate and then interpolated to synthesize the signal. In this case reducing the sampling rate is done up to the rate in which the spectrogram of the synthesized signal is similar to the spectrogram of the original signal and the synthesized sound is the same as the input sound. Finally the amplitude signals of all bands were summed to obtain the final output speech. So we have the following relation (N=number of bands):

$$Z(t) = \sum_{k=1}^{N} A_k(t) \mathrm{Cos}(2\pi f_{ck} t + \theta_k), \tag{4}$$

Where $A_k(t)$ is the k th band’s amplitude modulation, whereas f_{ck} and θ_k are the k th band’s center frequency and initial phase [14]. $Z(t)$ is the synthesized signal based on extracted information in CIS method.

A. Sample reduction with a fixed rate

In this case, all frequency bands in terms of sample reduction rates are treated alike. As in the previous section we explained, reducing the sampling rate is done with the rate in which the spectrograms of the original signal and the synthesized signal are similar visually. Fig. 1 shows block diagram of sample reduction with a fixed rate.

B. Sample reduction rate proportional to the band

In sample reduction with a fixed rate if the rate is appropriate for low frequency bands, causing deterioration and loss of information in high frequency bands. If the sample reduction rate is suitable for high frequency bands, the rate used for low frequency bands is high and leading to increase in computational cost. So in these cases variable rate is considered and determining the proper value for the rate is very important.

Method which is discussed in this section, reduces the sampling rate for each band according to the center frequency of that band. So reducing the sampling rate is proportional to the frequency information of each band in which f_{ck} is the k th band’s center frequency.

$$\begin{gathered}\text{Sample reduction rate}(SR_k) = \eta \times Fs/f_{ck} \\ Fs = 44100,\end{gathered} \tag{2}$$

It should be noted that η is an adjustable parameter and you can change it until achieving proper rate for sample reduction. The advantage of this method than the previous method is that the reduction in sampling rate per band is proportional to the frequency information of that band, so for high frequency bands a smaller rate and thus a greater number of samples are used for signal synthesis.

C. Sample reduction rate proportional to the signal

In this case to obtain sample reduction rates of different bands, in addition to frequency information of each band an envelope signal is also considered. Thus the mean value of envelopes is used as a criterion for maintaining those samples which are used in signal synthesis. In fact, this approach does not consider samples for signals which have small amplitudes. So we have:

$$\begin{gathered}M = \text{mean(envelope signals)} \\ \text{Sample reduction rate}(SR_k) = \eta \times Fs/f_{ck} \\ d_k = \text{downsample}(M, SR_k) \\ \text{if envelope}_k \succ \tau \times d_k \rightarrow \text{accept that sample,}\end{gathered} \tag{3}$$

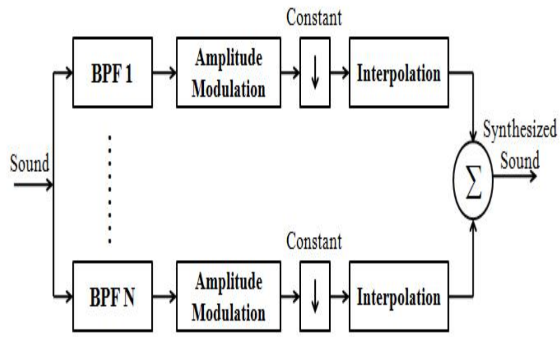

Figure 1. Block diagram of sample reduction with a fixed rate

τ is also an adjustable parameter that works like a kind of threshold. The idea of this method is the same as “n of m” strategy in which the signal is filtered into m frequency bands and n bands of envelope signals with greater energy are selected.

It should be noted that the number of samples used in this method is far less than previous methods, so the amount of information is reduced.

IV. RESULTS

As it mentioned before, reducing the sampling rate is done up to the rate in which the spectrogram of the synthesized signal is similar to the spectrogram of the original signal. In this case when you listen to the synthesized signal, the sound is the same as the input signal. You can see the spectrogram of "Choice" with the sample reduction rate of 20 in Fig. 2. As you see in the figure the spectrograms are the same up to this rate, so reducing the sampling rate for "Choice" is possible with the rate of 20.

The number of samples used for different bands is listed in the Table I. Sample reduction in three methods is done until all the individuals who were participated in experiment, stated that the synthesized sound is the same as the input sound.

As we expected, the number of samples decreases with sample reduction methods and choosing the rate proportional to the signal works better than the other methods. It should be noted that the total number of samples for "Choice" is 29084 and for "Hood" is 15203.

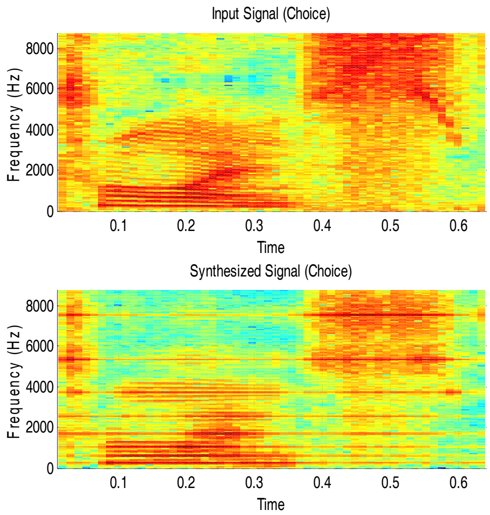

Figure 2. Spectrogram of the synthesized signal with sample reduction rate of 20

V. DISCUSSION

When we use sample reduction techniques, the amount of information or the number of samples for synthesizing the signal, the number of stimulating pulses and complexity of computation will decrease. Among these techniques the method which is proportional to the signal is better than the other methods.

Sample reduction techniques were evaluated on words such as "Choice" and "Hood" and results showed that the number of samples was reduced.

In sample reduction method with a fixed rate because of identical sampling rate for all bands, the loss of information in bands with greater frequency content is more. Therefore, if the input signal contains less information on these bands, the sampling rate can be reduced with greater rate.

In sample reduction method proportional to the band, the sampling rate is reduced according to the center frequency of each band. So in this case for high frequency bands with greater frequency content, reducing the sampling rate is done with a smaller rate and thus a greater number of samples are used for signal synthesis.

In sample reduction method proportional to the signal because envelope signal is considered in addition to frequency information, number of samples is not always ascendant. It means that, this method does not use more samples for high frequency bands necessarily and in this case the selection criterion is the envelope signal.

VI. CONCLUSION

In this paper adaptive method proportional to processing parameters and the input signal was presented which can synthesize speech signal with fewer samples, whereas transferring useful information to the patients.

Employing this method can lead to reduce the computational cost and if use to determine the appropriate rate of stimulation, the fewer number of stimulating pulses are used. It should be noted that the use of variable pulse rate requires further investigation, because the audition pattern varies according to individuals and the variable rate may sometimes leads to hearing impairment for patients.

TABLE I. THE NUMBER OF SAMPLES WHICH IS USED FOR SIGNAL SYNTHESIS

Number of samples	**Sample reduction with a fixed rate**		**Sample reduction rate proportional to the band**		**Sample reduction rate proportional to the signal**	
	Choice	***Hood***	***Choice***	***Hood***	***Choice***	***Hood***
Samples for band 1	1455	507	74	23	58	28
Samples for band 2	1455	507	186	58	140	58
Samples for band 3	1455	507	339	105	265	98
Samples for band 4	1455	507	549	169	437	108
Samples for band 5	1455	507	831	258	665	149

Samples for band 6	1455	507	1212	381	1088	294
Samples for band 7	1455	507	1711	543	1178	411
Samples for band 8	1455	507	2424	761	1523	357
Total samples	11640	4056	7326	2298	5354	1503

ACKNOLEDGMENT

The authors would like to thank Dr. Mohammad Farhadi for his helpful advice.

REFERENCES

[1] P.C. Loizou, "Introduction to cochlear implants", IEEE Eng. Med Biol. Mag. 18 (1), pp. 32-42, 1999.

[2] F.-G. Zeng, "Trends in cochlear implants", Trends Amplif. 8, pp. 1-34, 2004.

[3] F. Chen, Y.-T. Zhang, "A novel temporal fine structure-based speech synthesis model for cochlear implant", Elsevier Signal Proc. 88, pp. 2693-2699, 2008.

[4] P.C. Loizou, "Mimicking the human ear", IEEE Signal Proc. Mag., pp. 101-130, September 1998.

[5] B.S. Wilson, C.C Finley, D.T. Lawson, R.D. Wolford, D.K. Eddington, W.M. Rabinowitz, "Better speech recognition with cochlear implants", Nature 352 (6332), pp. 236-238, July 1991.

[6] B.S. Wilson, M.F. Dorman, "Cochlear implants: a remarkable past and a brilliant future", Elsevier Hear. Res. 242, pp. 3-21, 2008.

[7] L.T. Cohen, L.M. Richardson, E. Saunders, R.S.C. Cowan, "Spatial spread of neural excitation in cochlear implant recipients: comparison of improved ECAP method and psychophysical forward masking", Elsevier Hear. Res. 179, pp. 72-87, 2003.

[8] A.v. Wieringen, R.P. Carlyon, J. Laneau, J. Wouters, "Effects of waveform shape on human sensitivity to electrical stimulation of the inner ear", Elsevier Hear. Res. 200, pp. 73-86, 2005.

[9] D.K. Eddington, C. Boex-Spano, J. Tierney, V. Noel, M. Whearty, "Speech processors for auditory prostheses", Seventh quarterly progress report. Massachusetts Institute of Technology, Research Laboratory of Electronics, Cambridge, 1997.

[10] C.S. Throckmorton, M.S. Kucukoglu, J.J. Remus, L.M. Collins, "Acoustic model investigation of a multiple carrier frequency algorithm for encoding fine frequency structure: Implications for cochlear implants", Elsevier Hear. Res. 218, pp. 30-42, 2006.

[11] K. Fishman, R.V. Shannon, W.H. Slattery, "Speech recognition as a function of the number of electrodes used in the SPEAK cochlear implant speech processor", J. Speech Hear. Res. 40 (5), pp. 1201-1215, October 1997.

[12] L. Friesen, R.V. Shannon, D. Baskent, X.S. Wang, "Speech recognition in noise as a function of the number of spectral channels: comparison of acoustic hearing and cochlear implant", J. Acoust. Soc. Am. 110 (2), pp. 1150-1163, 2002.

[13] F. Zheng, G. Zhang, Z. Song, "Comparison of different implementation of MFCC", J. Comp. Sci. Tech., 16(6), pp. 582-589, September 2001.

[14] K. Nie, G. Stickney, F.-G. Zeng, "Encoding frequency modulation to improve cochlear implant performance in noise", IEEE Trans. Biomed. Eng., vol. 52, no. 1, pp. 64-73, January 2005.

[15] F.-G. Zeng, K. Nie, G.S. Stickney, Y.-Y. Kong, M. Vongphoe, A. Bhargave, C. Wei, K. Cao, "Speech recognition with amplitude and frequency modulations", PNAS (Proceedings of the National academy of Science of the United States of America), vol. 102, no. 7, pp. 2293-2298, February 2005.

[16] R.A. Fearn, "Music and pitch perception of cochlear implant recipients". A thesis for the degree of Doctor of Philosophy, University of New South Wales, December 2001.

[17] C.-H. Chang, G.T. Anderson, P.C. Loizou, "A neural network model for optimizing vowel recognition by cochlear implant listeners", IEEE Trans. Neural Systems and Rehabilitation Eng., vol. 9, pp. 42-48, August 2002.

[18] J.H. Won, S.M. Schimmel, W.R. Drennan, P.E. Souza, L. Atlas, J.T. Rubinstein, "Improving performance in noise for hearing aids and cochlear implants using coherent modulation filtering", Elsevier Hear. Res. 239, pp. 1-11, 2008.