2011 UKSim 13th International Conference on Modelling and Simulation

A New Adaptive Frequency and Amplitude Modulation Encoding 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—**In recent years, methods for speech processing in cochlear implant systems have been proposed in which frequency modulation is extracted as well as amplitude modulation in different bands. In this paper besides considering one of the above strategies, 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; amplitude modulation; frequency 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]. In this way 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].

Recent research [3,7,8] has shown that methods based on amplitude extraction cause difficulties for speech recognition in noise. So in recent years, methods have been proposed that extracting frequency information in addition to amplitude information. These methods are such as FAME (Frequency Amplitude Modulation Encoding) [7], MCFA (Multi Carrier Frequency Algorithm) [8] and Temporal Fine Structure extraction using Hilbert transform [3].

In MCFA method the most prominent frequency in each band is estimated by the short-time Fourier transform (STFT). Although the STFT provides a computationally fast and simple method of estimating frequency, its resolution is limited by the sample duration which is quite short on the order of milliseconds for cochlear implant systems [8].

Hilbert transform is a common method to extract amplitude and frequency modulations, but the estimated instantaneous frequency usually varies rapidly and over a broad range so producing values that often have no clear physical meaning. Also this method is difficult to apply directly to cochlear implant systems because the extracted frequency modulation generally varies too widely in range and too rapidly in rate [7].

So in this paper we consider FAME method in which amplitude and frequency modulations are extracted in two separate ways. According to Nie et al. discoveries, FAME method contains more acoustic information than previously used methods such as CIS [7]. By comparing the spectrograms of a synthesized signal and an original signal, you can find out that the FAME spectrogram most closely matches the original spectrogram visually [7,8]. So for improving speech processor in cochlear implant systems, one suggestion is the usage of FAME method instead of CIS.

Considering that in FAME method frequency modulation is extracted as well as amplitude modulation, the amount of information increases. Increasing the amount of information or increasing the number of samples causes an increase in the number of stimulating pulses, so complexity of computation and hardware will follow. Therefore, employing sample reduction techniques for speech signal are profitable. In this paper adaptive methods are designed with the aim of transferring sufficient and useful information to patients.

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 [4]. In FAME method amplitude and frequency modulations were extracted in two parallel pathways for each band. The amplitude modulation pathway extracted the slowly varying envelope, while the frequency modulation pathway extracted the slowly varying frequency modulation [7].

Number of filters used in the filter bank was eight because patients could not functionally use more than eight frequency bands [8-10]. These filters were tenth-order Butterworth bandpass filters, with frequencies Mel spaced [11] between 80 and 8800 Hz [7]. The lowpass filters that were used to extract amplitude and frequency modulations were fourth-order Butterworth filters. For

978-0-7695-4376-5/11 $26.00 © 2011 IEEE
DOI 10.1109/UKSIM.2011.70

synthesizing, all subband signals were summed to form an output signal [3,7,8].

Same as many other researches [8,12-15], benchmarks which were used in this study were “Choice” and “Hood” signals. For evaluating the methods, 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.

Both methods (CIS and FAME) were 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 and frequency modulations and signal synthesis were implemented in the Matlab (R2008a) environment.

III. DIFFERENT SAMPLE REDUCTION RATES

FAME method works better in synthesizing a signal because of extracting more information. Although this approach leads to better performance but increase the amount of information, so employing methods to reduce the number of samples and computational efficiency seems to be useful. For this purpose the amplitude and frequency modulations 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 in each band, the frequency signal multiplied by the amplitude signal to recover the original signal of band and the final output speech is obtained by summing these signals. So we have the following relation (N=number of bands):

$$Y(t)=\sum_{k=1}^{N} A_k(t)\mathrm{Cos}\left[2\pi f_{ck}t+2\pi\int_0^t g_k(\tau)d\tau+\theta_k\right], \qquad (1)$$

Where $A_k(t)$ and $g_k(t)$ are the k th band’s amplitude and frequency modulations, whereas f_{ck} and θ_k are the k th band’s center frequency and initial phase [7]. $Y(t)$ is the synthesized signal based on extracted information in FAME 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.

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

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:

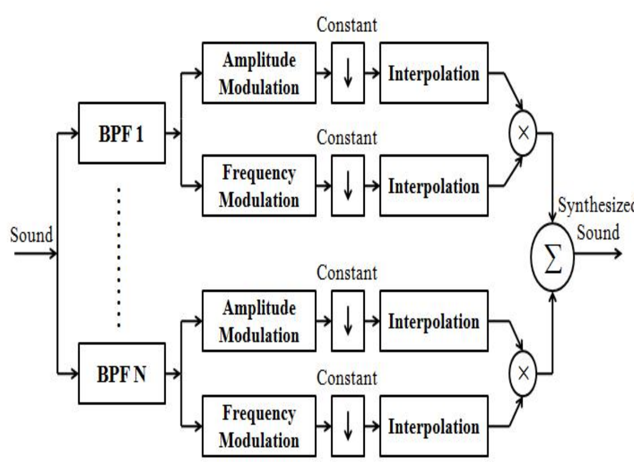

Figure 1. Block diagram of sample reduction with a fixed rate

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

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 30 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 30.

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.

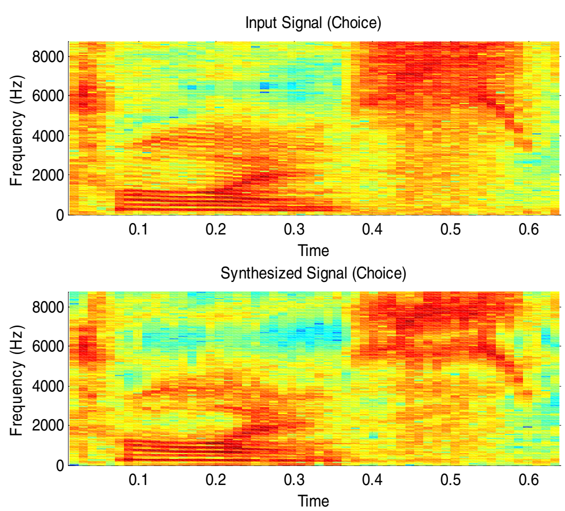

Figure 2. Spectrogram of the synthesized signal with sample reduction rate of 30

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.

V. DISCUSSION

Although Shannon et al. in 1995 found that cochlear implant users could achieve good speech recognition scores in quiet with only four bands, their performance in noise is compromised and noise is one of the major problems in cochlear implant systems [7,8]. So in recent years methods have been proposed that extract more information from the input signal in order to overcome this problem. In most of these methods in addition to amplitude information in different bands, frequency information is also extracted that causes an increase in the amount of information.

So if we want to use these methods in cochlear implant systems, we must seek ways to reduce the amount of information. We can use sample reduction techniques discussed in this paper and among them the method which is proportional to the signal is better.

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.

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	970	305	65	23	55	22
Samples for band 2	970	305	162	56	134	55
Samples for band 3	970	305	294	102	241	101
Samples for band 4	970	305	477	164	399	97
Samples for band 5	970	305	728	250	630	178
Samples for band 6	970	305	1078	362	1005	327
Samples for band 7	970	305	1531	525	1119	344
Samples for band 8	970	305	2238	761	1478	684
Total samples	7760	2440	6573	2243	5061	1808

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.

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] 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.

[8] 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.

[9] 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.

[10] 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.

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

[12] 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.

[13] 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.

[14] 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.

[15] 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.