Proceedings of the 28th IEEE
EMBS Annual International Conference
New York City, USA, Aug 30-Sept 3, 2006

Loudness Normalization for Cochlear Implant Using Pulse-Rate Modulation to Convey Mandarin Tonal Information: A Model-Based Study

Fei Chen [1] and Yuan-ting Zhang [1, 2]
[1] Shun Hing Institute of Advanced Engineering
[2] Joint Research Centre for Biomedical Engineering, Department of Electronic Engineering
The Chinese University of Hong Kong, Hong Kong
(Email: ytzhang@ee.cuhk.edu.hk)

Abstract – Cochlear implant (CI) devices employ electrical pulsatile stimulation of the auditory nerves (AN) to restore partial hearing to a profoundly deafened person. In order to improve the speech perception for CI users speaking tonal language, such as Mandarin, the pulse-rate has been suggested to be modulated according to the Mandarin tonal patterns to convey the Mandarin tonal information. However, recent psychological experiments have found that the pulse-rate modulation will produce accompanying variation of perceived loudness. The purpose of this paper is to introduce an amplitude compensation scheme to normalize the loudness perception when the pulse-rate is modulated to convey the Mandarin tonal information. Based on an integrate-and-fire AN model, a loudness perception model and a pitch perception were implemented. Result of model-based simulation showed that using the proposed amplitude compensation scheme, the estimated loudness was normalized while the Mandarin tonal information could still be efficiently transmitted. It is believed that, when the proposed electrical pulsatile stimulation incorporating both pulse-rate modulation and amplitude compensation is integrated with present CI devices, it would more efficiently enhance the speech identification for cochlear implantee speaking tonal languages, such as Mandarin.

Keywords – Electrical pulsatile stimulation, pulse-rate modulation, Mandarin tonal information, cochlear implant.

I. INTRODUCTION

It has been widely believed that cochlear implant (CI) is the only medical intervention that can restore partial hearing to a profoundly deafened person [1]. CI inserts the electrodes in the scala tympani of the cochlea, and directly employs electrical pulsatile stimulation of the auditory nerves (AN) to produce speech perception. The amplitude of the electrical pulses are modulated by that of the speech signal processed by CI speech processor, while the rate of electrical pulses is normally constant, exemplified by the well-known continuous-interleaved-sampling (CIS) strategy [2]. Up to date, CI technology has progressed to the stage that the average user can understand conversational speech in quite surroundings without the aid of lip reading. However, this good performance can be achieved mainly by CI users speaking mono-tonal languages, such as English and German [3]. It has also been reported that CI users who speak tonal language, such as Mandarin, showed poor results in speech identification [4]. Tonal languages are different from mono-tonal languages in that they also use different tones to express the lexical meaning of the pronounced words. Thereby, current CI devices need to be improved for a large population of users speaking tonal languages.

In order to enhance the speech identification of CI users speaking tonal language, novel speech processing strategies have been recently proposed, with attempts to extract tonal information from speech signal and encode it into the electrical pulse train to produce the perception of diverse tonal patterns [3, 5-7]. The usage of pulse-rate modulation to convey tonal information has attracted substantial attentions [3, 7]. The basis of this work is to extract the tonal pattern through the contour of fundamental frequency (F0) of speech signal, since it was found that the F0 variation contour could be used to indicate the tonal patterns. For instance, the F0 contour of Mandarin speech signal is able to represent the four tonal patterns pronounced in Mandarin, flat tone (–), rising tone (／), falling-rising tone (∨), and falling tone (＼). Then, the pulse-rate is modulated according to the extracted tonal pattern. Several studies have also shown that variable stimulation frequency can convey a certain amount of information on F0 in the speech signal [8-9].

Despite the improved speech perception performance from acoustic simulation experiments, not many results directly from CI users have been reported for the pulse-rate modulation strategy to convey Mandarin tonal information. On the other hands, recent studies from psychological experiments have found that the increment of pulse-rate would increase the loudness sensation perceived by the CI users [10]. Therefore, it might cause some detrimental effect on speech perception by using pulse-rate modulation, although it was originally expected to efficiently convey the tonal information. Besides, considering the limited dynamic hearing range of CI users [1], it is necessary to diminish such synthetic loudness variation. The purpose of this paper is to propose a strategy, i.e. amplitude compensation, to reduce the accompanying perceptual deficits from using pulse-rate modulation to convey Mandarin tonal information.

II. METHODOLOGY

1) Auditory Nerve Model

An integrate-and-fire based AN model has been recently developed [11], as shown in Fig. 1. The pulsatile stimulating current, *I*, would increase the potential of the capacitor *C*, which simulated the depolarization of membrane potential. A Gaussian noise, V_{noise}, was introduced to present the membrane potential fluctuation. When the membrane potential *V* exceeded the threshold, an action potential (AP) was produced, and the membrane potential *V* was reset to zero. Besides the fixed threshold V_{th}, an additional threshold function V_{Ref} was used to simulate complex dynamic of neural threshold, incorporating both absolute refractory period and relative refractory period during neural firing, as

$$V_{Ref} = \begin{cases} \infty, & t - t_i < 0.6ms \\ V_{th} \times 0.97 e^{-\frac{t-t_i}{1.32ms}}, & t - t_i \geq 0.6ms \end{cases}, \quad (1)$$

where t_i was the occurrence time of the *i*th neural spike. For up to 0.6 ms after the firing of an AP, it was impossible to evoke another AP, no matter how large the depolarizing stimulus. This absolute refractory period was followed by a relative refractory period, during which the stimulus must be larger than normal to evoke an AP.

The whole ANs distributed along cochlea were modeled by groups of ANs. When electrical pulsatile current was applied to an electrode, only those AN groups close to the electrode would be activated. The spatial attenuation of current was taken to be 4 dB/mm for biphasic electrical stimulation, by which we might calculate the stimulation current at the site of each AN group for a given stimulus intensity.

2) Loudness Perception Model and Pitch Perception Model

The loudness perception model was built up based on neural spike trains from a population of stimulated ANs. The neural spike trains were fed to a temporal integration window. The output from the integrator was used to indicate the estimated loudness [10].

The pitch perception model was implemented by combining the interpulse intervals from the whole active ANs, searching for the common time intervals, and picking the most prominently intervals. The estimated pitch was corresponding to the reciprocal of the interval selected [12].

3) Amplitude Compensation Scheme

To convey the 2nd, 3rd and 4th Mandarin tones, the pulse-rates had been suggested to be modulated as their tonal patterns, as shown in Fig. 2. The basis for amplitude compensation was to adjust the amplitude of pulsatile train accompanying the pulse-rate modulation.

Regarding the pulse-rates modulated for the 2nd and 4th Mandarin tones, the following two functions were used to compensate the amplitude of pulsatile train, as

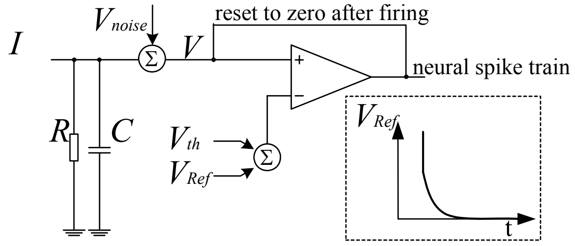


Fig. 1. The integrate-and-fire based AN model.

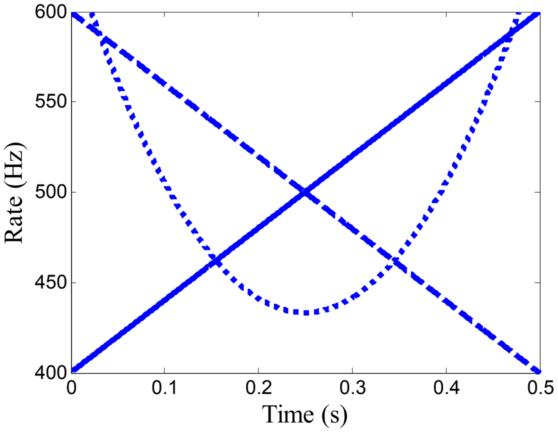


Fig. 2. The modulated pulse-rates of electrical pulsatile stimulation to convey the 2nd (solid line), 3rd (dotted line), and 4th (dashed line) Mandarin tonal patterns, respectively.

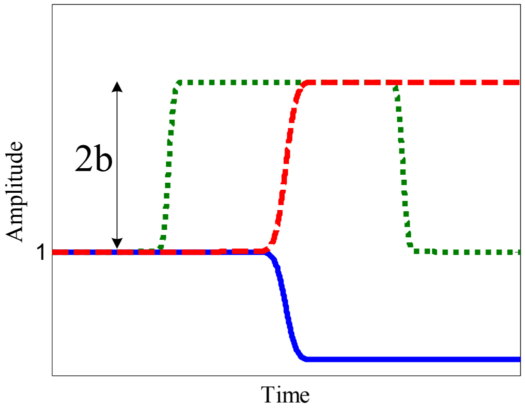


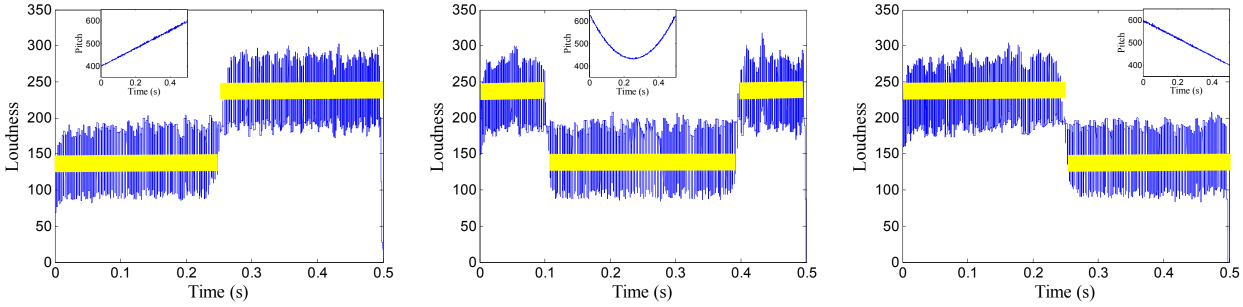
Fig. 3. The amplitude compensation schemes for pulse-rate modulated to convey the 2nd (solid line), 3rd (dotted line), and 4th (dashed line) Mandarin tonal patterns.

$$y = erf(-a \cdot t) \cdot b + 1 - b, \quad (2)$$

and

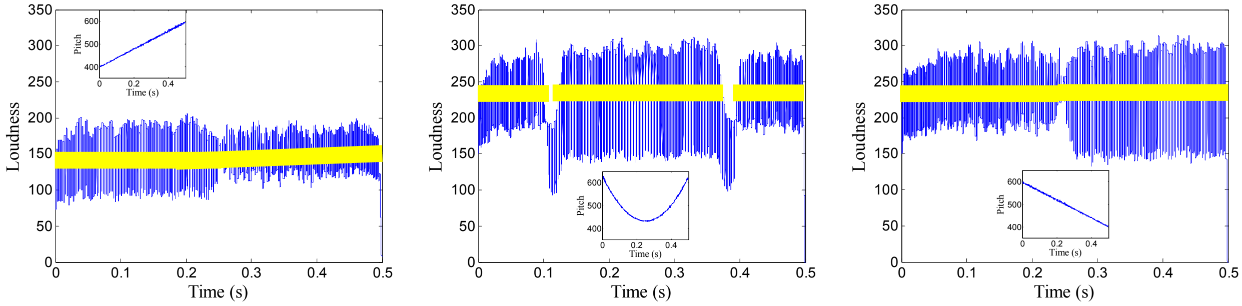
$$y = erf(a \cdot t) \cdot b + 1 + b, \quad (3)$$

where *erf*(•) was the error function, *a* controlled the slope of amplitude change, and *b* denoted the degree of amplitude compensation.



(a) (b) (c)

Fig. 4. The estimated loudness patterns for electrical pulsatile stimulation using pulse-rate modulation to convey the 2nd (a), 3rd (b), and 4th (c) Mandarin tonal patterns, respectively. The inserted plot in each figure indicates the corresponding estimated pitch pattern.



(a) (b) (c)

Fig. 5. The estimated loudness patterns after amplitude compensation for electrical pulsatile stimulation using pulse-rate modulation to convey the 2nd (a), 3rd (b), and 4th (c) Mandarin tonal patterns, respectively. The inserted plot in each figure indicates the corresponding estimated pitch pattern.

For the pulse-rate modulated to convey the 3rd Mandarin tone, a complex function was determined jointly by (2) and (3) in different time regions, as illustrated in Fig. 2. Since the 1st Mandarin tone was characterized by a flat tone, the pulse-rate did not need modulation, and no specific compensation function was designed.

III. SIMULATION RESULTS

In the simulation study, only one electrode was activated. For the grouping of ANs, the group closest to the electrode consisted of 20 ANs, and each of other groups contained 15 ANs. We used electrical pulse train with duration of 0.5 s to stimulate the ANs. The pulse-rate was centered at 500 Hz, and modulated according to the Mandarin tonal patterns in a range between 400 Hz and 600 Hz, as shown in Fig. 2. A reference period of 4 ms was chosen to implement the temporal integration for loudness estimation.

We first estimated the loudness and pitch when the pulse-rate was modulated according to the three Mandarin tonal patterns and the amplitude of electrical pulse train was constant. Results in Fig.4 (a)-(c) demonstrated that, although the estimated pitch resembled the corresponding tonal pattern, the estimated loudness also varied in a similar mode to that of pulse-rate. When the amplitudes of electrical pulse train were compensated, the estimated loudness patterns were normalized, as shown in Fig. 5 (a)-(c); meanwhile, the Mandarin tonal information could still be efficiently transmitted.

IV. DISCUSSION

The loudness sensation is an important factor to be considered during the design of CI devices, owing to the factor that many parameters of the pulsatile simulation may control the resultant loudness percept. Besides stimulation rate, psychological experiments have found that stimulation level, electrode separation, stimulation mode and so on, may affect the perceived loudness [10, 13]. McKay and McDermott have suggested that the effect of increasing current level on loudness perception resulted more from recruitment of auditory neurons than any increment in the average spiking probability [10]. The model-based study in this paper supported this view, which might also help us to understand the mechanism behind the amplitude compensation scheme for loudness normalization. Figure 6 showed the spread of the action potentials as the pulse-rate was modulated to convey the 4th Mandarin tone. When the pulse-rate was decreased, it initially would reduce the count of APs within the time window for loudness estimation. However, after amplitude compensation by increasing the pulse amplitude, the electrical current was strong enough to

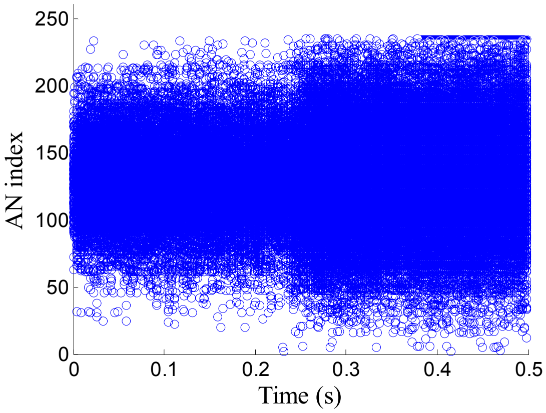


Fig. 6. The spread of action potentials across AN position and time. Each circle in the figure indicates one AP. The stimulation was centralized around the 130th AN, and the stimulation level decayed at both sides away from the stimulation position in the figure.

spread to a large scope and activate more ANs. Therefore, the amount of AP increment from more excited ANs counterbalanced that reduced from pulse-rate decrement, which enabled loudness constancy on the whole.

The proposed scheme could be integrated with the existing CI devices, which attempt to use pulse-rate modulation to convey Mandarin tonal information [3]. On the other hand, although the majority of CI users speaking non-tonal languages were able to understand the conversational speech in quite environments and talk on the phone, recent work has reported that they had difficulties in music appreciation, speaker recognition, or other tasks needing pitch perception [14-15]. Therefore, the proposed scheme is still useful to the non-tonal language CI users for the situation that the pulse-rate of the electrical stimulation is modulated to deliver the pitch information.

Using the amplitude compensation functions, the stimulation level in each electrode could be compensated before it is further modulated according to the speech signal processed by CI speech processor. On the other hand, since the proposed compensation scheme may circumvent the loudness variation accompanying pulse-rate modulation, it prevents loudness variation being used as a cue, and provides an objective way to evaluate the efficiency of using pulse-rate modulation to convey tonal information.

V. CONCLUSION

In this paper, a simple and efficient amplitude compensation scheme was introduced to normalize the loudness perception when the pulse-rate of electrical pulsatile stimulation was modulated to convey the Mandarin tonal information. Result of model-based simulation shown that using the proposed amplitude compensation scheme, the estimated loudness was normalized while the Mandarin tonal information could still be efficiently transmitted. It is believed that, when the proposed electrical pulsatile stimulation incorporating both pulse-rate modulation and amplitude compensation is integrated with present CI devices, it would more efficiently enhance the speech identification for cochlear implantee speaking tonal languages, such as Mandarin.

ACKNOWLEDGMENT

The authors would like to acknowledge the support of Shun Hing Institute of Advanced Engineering.

REFERENCES

1. F.G. Zeng, “Trends in cochlear implants,” *Trends Amplif.*, vol. 8, pp.1-34, 2004.
2. P.C. Loizou, “Signal-processing techniques for cochlear implants,” *IEEE Eng. Med. Biol. Mag.*, vol. 18, pp. 34-46, 1999.
3. N. Lan, K.B. Nie, S.K. Gao, and F.G. Zeng, “A novel speech-processing strategy incorporating tonal information for cochlear implants,” *IEEE Trans. Biomed. Eng.*, vol. 5, pp. 752-760, 2004.
4. C.G. Wei, K.L Cao, and F.G. Zeng, “Mandarin tone recognition in cochlear-implant subjects,” *Hear. Res.*, vol. 197, pp. 87-95, 2004.
5. X. Luo and Q.J. Fu, “Enhancing Chinese tone recognition by manipulating amplitude envelope: implications for cochlear implants,” *J. Acoust. Soc. Amer.*, vol. 116, pp. 3659-3667, 2004.
6. K. Nie, G. Stickney, and F. G. Zeng, “Encoding frequency modulation to improve cochlear implant performance in noise,” *IEEE Trans. Biomed. Eng.*, vol. 52, pp. 64-73, 2005.
7. H. McDermott and C. McKay, “Cochlear implant sound processing method and system,” U.S. Patent Application Publication US 2005/0107843 A1, May 19, 2005.
8. B. Townshend, N. Cotter, and D. V. Compernoll, “Pitch perception by cochlear implant subjects,” *J. Acoust. Soc. Amer.*, vol. 82, pp. 106-115, 1987.
9. P. A. Busby, Y. C. Tong, and G. M. Clark, “Electrode position, repetition rate, and speech perception by early and late-deafened cochlear implant patients,” *J. Acoust. Soc. Amer.*, vol. 93, pp. 1058-1067, 1993.
10. C.M. McKay and H.J. McDermott, “Loudness perception with pulsatile electrical stimulation: The effect of interpulse intervals,” *J. Acoust. Soc. Amer.*, vol. 104, pp. 1061-1074, 1998.
11. F. Chen and Y.T. Zhang, “An integrated-and-fire based auditory nerve model and its response to high-rate pulse train,” *Neurocomputing*, acceptance subject to revision, 2006.
12. B.C.J. Moore, “Pitch perception,” in *An introduction to the psychology of hearing*, B.C.J. Moore, Ed. San Diego, Calif.: Academic Press, pp. 177-212, 1997.
13. C.M. McKay, M.D. Remine, and H.J. McDermott, “Loudness summation for pulsatile electrical stimulation of the cochlea: Effects of rate, electrode separation, level, and mode of stimulation,” *J. Acoust. Soc. Amer.*, vol. 110, pp. 1514-1524, 2001.
14. Y.Y. Kong, R. Cruz, J.A. Jones, and F.G. Zeng., “Music perception with temporal cues in acoustic and electric hearing,” *Ear. Hearing*, vol. 25, pp. 173-185, 2004.
15. M. Vongphoe and F.G. Zeng, “Speaker recognition with temporal cues in acoustic and electric hearing,” *J. Acoust. Soc. Amer.*, vol. 118, pp. 1055-1061, 2005.