ENHANCED SPARSE SPEECH PROCESSING STRATEGY FOR COCHLEAR IMPLANTS

Hongmei Hu[1,2], Guoping Li[1], Liang Chen[3], Jinqiu Sang[1], Shouyan Wang[1], Mark E Lutman[1], Stefan Bleeck[1]

1. Institute of Sound and Vibration Research, University of Southampton, SO17 1BJ, Southampton, UK
2. Department of Testing and Control, Jiangsu University, 212013, Zhenjiang, China
3. Department of Navigation and Positioning, Finnish Geodetic Institute, 02431, Masala, Finland

email: [hongmei.hu, lgp, js1e09, sy.wang, bleeck]@soton.ac.uk, mel@isvr.soton.ac.uk, liang.chen@fgi.fi[3]
web: http://www.audis-itn.eu/wiki/Hongmei%20Hu

ABSTRACT

To improve the performance of cochlear implants (CI) in noisy environments, a novel strategy is proposed by combining two established algorithms. By combining the Kalman Expectation-Maximize (KEM) algorithm for noise reduction and the SPARSE algorithm for information selection, the new "enhanced SPARSE" strategy gains from the benefits of both. In the new algorithm, the noisy speech is first transferred to the time-frequency domain via a 22- channel filter bank and the envelope in each frequency channel is extracted; secondly, KEM filtering is applied to the envelope in each channel; finally, the SPARSE strategy developed previously in our group is used to generate more sparse stimuli. Here we present results of objective and subjective experiments where KEM was applied to the standard CI speech strategy (Advanced Combination Encoder, ACE) and the novel SPARSE strategies for comparison. Sparseness, measured by kurtosis, of the KEM enhanced simulations was higher than the corresponding output without noise reduction. In subjective listening experiments, six normal hearing listeners were tested with sentences in noise at three signal-to-noise ratios (SNR): 0, 5, 10 dB. Speech intelligibility was better using the enhanced ACE algorithm than the original ACE for all these three SNR, enhanced SPARSE yielded better performance in low SNR (0 dB) than SPARSE, but the advantage was not obvious for the higher SNR (5, 10 dB). We conclude that the suggested strategy shows promise for achieving better speech perception for CI users in the future.

1. INTRODUCTION

Cochlear implants are electrical devices that help to restore hearing to the profoundly deaf. The main principle of cochlear implants is to stimulate the auditory nerve via electrodes surgically inserted in the inner ear. With the development of new speech processors and algorithms, the majority of implanted users benefit from this device, some of them to a degree that allows them to communicate via telephone without much difficulty. However, average performance of most cochlear implant users still falls below normal hearing (NH) listeners, and speech quality and intelligibility generally deteriorate in the presence of background noise. Specifically, users often complain that their cochlear implants do not work well in background noise. One of the most relevant differences between NH and CI users in terms of speech perception is the dynamic range: the dynamic range of the impaired ear is much smaller than that of the normal ear. Thus the electrical stimulation provides a severe bottleneck of the information transfer, which only allows limited acoustic information to be transmitted to the auditory neurons [1]. Our recently developed SPARSE speech processing strategies based on sparse coding theory [2] significantly improve the speech intelligibility in patients with cochlear implants by reducing the level of noise and increasing dynamic range simultaneously [3] to overcome the bottleneck of the information transmission.

There are currently two main ways how speech processing algorithms improve CI performance: one focuses on noise reduction by trying to enhance speech and suppress noise, such as model-based and non-model-based noise reduction algorithms [4-6]; the other focuses on redundancy reduction using cochlear coding strategies[3] [7, 8] to make good use of the limited dynamic range in the impaired auditory system. Most of speech enhancement algorithms are based on some assumptions on the noise distribution, for instance, our SPARSE algorithm assumes that the speech and noise have non-Gaussian distribution. In reality it is unlikely that the noise has either Gaussian or non-Gaussian distribution but mixture of two. Such multiple type of distribution of noise may reduce the performance and robustness of speech enhancement algorithms. This work aims to enhance our SPARSE algorithm by reducing Gaussian noise with a robust Kalman Expectation-Maximize (KEM) method [4, 5] and the enhanced algorithms will further improve the performance of cochlear implant users. The enhanced algorithms are evaluated by objective measures and subjective speech intelligibility test.

2. KEM NOISE REDUCTION ALGORITHM

We aim to reduce the Gaussian noise prior to SPARSE processing. A model-based KEM algorithm is derived in which the noisy speech is assumed to be a stochastic process with an autoregressive (AR) clean speech source contaminated with additive white Gaussian noise (AWGN). The proposed KEM method is robust and provides less distortion than other similar Bayesian methods [6].

2.1 KEM method for white Gaussian noise reduction

2.1.1 Problem formulation

Suppose $z(t)$ is a measured noisy signal, which is mixture of a clean signal $s(t)$ and additive background noise $w(t)$:

$$z(t) = s(t) + w(t) \tag{1}$$

where $s(t)$ is modelled as a stochastic AR process, excited by Gaussian white noise with zero mean and variance g_s , i.e. $N(0, g_s)$. $w(t)$ conforms to a Gaussian process with distribution of $N(0, g_w)$.

$$\begin{aligned} s(t) &= -\sum_{k=1}^{p} a_k s(t-k) + \sqrt{g_s} u_s(t) \\ w(t) &= \sqrt{g_w} u_w(t) \end{aligned} \tag{2}$$

where $\alpha_p, \alpha_{p-1}, \ldots, \alpha_1$ are the AR coefficients and $u_w(t)$, $u_s(t)$ are white noise with Gaussian distribution of $N(0,1)$.

The corresponding state-space expressions for equations (1) and (2) are

$$\begin{aligned} \mathbf{x}_t &= \mathbf{A}\mathbf{x}_{t-1} + \mathbf{B}u_s(t) \\ z(t) &= \mathbf{H}\mathbf{x}_t + \sqrt{g_w} u_w(t) \end{aligned} \tag{3}$$

where

$$\mathbf{x_t} = \begin{bmatrix} s(t-p) \\ s(t-p+1) \\ \vdots \\ s(t) \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ & \ddots & \ddots & & \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \ddots & 1 \\ -a_p & -a_{p-1} & \cdots & -a_2 & -a_1 \end{bmatrix},$$

$\mathbf{B} = \begin{bmatrix} 0 & 0 & \cdots & \sqrt{g_s} \end{bmatrix}^T$, $\mathbf{H} = [0, \ldots, 0, 1]$, $u_s(t) \sim N(0,1)$, $u_w(t) \sim N(0,1)$.

In current application, the *priors* of signal and noise parameters are unknown. Thus, given a measured signal $\mathbf{z}_{1:t}$, the problem is to simultaneously infer the clean speech sequence $\hat{\mathbf{s}}_{1:t}$ and the parameters: $\lambda = \{\boldsymbol{\alpha}, g_s, g_w\}$, where $\boldsymbol{\alpha} = [-\alpha_p \quad -\alpha_{p-1} \quad \ldots \quad -\alpha_1]^T$.

2.1.2 Algorithm description

It is assumed that the signal and noise are stationary within one short analysis frame. Within every frame, we apply an EM method to solve the problem, which can be divided into the E-step and M-step.

(1) E-step: Based on the N measurements in every frame $\mathbf{z}_{1:N}$ and the l^{th} iterative estimation of model parameters λ^l , the state mean $\hat{\mathbf{x}}_t$ and the error covariance matrix $\mathrm{P_t}$ can be estimated by the forward and backward Kalman smoothing method [6].

(2) M-step: Parameters λ^l are updated in this step [see Appendix for more details].

$$\boldsymbol{\alpha}^l = \left[\sum_{t=1}^{N} \widehat{\mathbf{x}_{t-1}\mathbf{x}_{t-1}} \right]^{-1} \sum_{t=1}^{N} \widehat{\mathbf{x}_{t-1} s_t} \tag{4}$$

$$g_s^l = \frac{1}{N} \sum_{t=1}^{N} (\hat{s}_t^2 + (\boldsymbol{\alpha}^l)^T \widehat{\mathbf{x}_{t-1} s_t}) \tag{5}$$

$$g_w^l = \frac{1}{N} \sum_{t=1}^{N} (z_t^2 - 2 z_t \hat{s}_t + \hat{s}_t^2) \tag{6}$$

A distinct advantage of the proposed algorithm compared to alternative algorithms is that it enhances the SNR of the speech, while preserving its intelligibility and natural sound quality [6].

In this paper, each channel of the ACE spectrum envelope was modelled by a linear dynamic state-space equation, and applying the KEM method.

3. ENHANCED SPARSE STRATEGY

The dynamic range for electrical stimulation for CI users is much smaller than acoustic dynamic range in the normal ear. Thus the electrical stimulation has a severe bottleneck to overcome, which only allows limited acoustic information to be transmitted to auditory neurons. However, many experiments have showed that speech has a high degree of redundancy and only few components are needed to allow people to understand speech [9, 10]. This ability to understand speech based on partial information has been explained by various theories, such as glimpsing theory [10]. This redundancy property of speech was investigated in the SPARSE strategy[3, 7].

3.1 Sparse strategy

Existing CI strategies, such as Continuous Interleaved Sampling (CIS), Spectral Peak (SPEAK) and Advanced Combination Encoder (ACE) indeed already take advantage of the redundancy property of speech by selecting only few channels or only using envelope information to stimulate auditory neurons. Li [3] suggested that the success of these strategies is due to that, in fact, they help to solve the information bottleneck problem by stimulating auditory neurons sparsely and efficiently. A Principal Component Analysis (PCA) and Independent Component Analysis (ICA) based sparse algorithm working on the spectral envelope for CI, called SPARSE, was proposed and tested at various SNR [7]. To deal with the limited dynamic range of CI users, the idea of reducing the redundancy of the stimuli by PCA and then making the electrical stimuli more sparse by applying thresholding after ICA was used. Since ICA extracts independent components out of the speech spectral envelope, the mutual information between these independent channels is minimal[3]. The envelope after SPARSE is more sparse than the ACE output. The electrical pulse trains driving the stimulation channels are modulated by the envelopes of the signals in the corresponding band pass filters. In addition, the pulse trains are separated in time and interleaved in order to avoid interaction among the electrodes. Previous experimental work using subjective listening tests has shown that the SPARSE strategy achieves improvements for CI users [3].

3.2 Enhanced sparse strategy

KEM can be applied in the time or frequency domain. To make the proposed strategy more compatible with current algorithms and easier to implement in real time, in this paper the enhanced SPARSE algorithm is applied to the spectrum envelope.

Figure 1 shows the idea of the KEM enhanced SPARSE strategy for CI stimulation. The pre-emphasis filter in figure 1 is to compensate for the –6 dB/octave natural slope in the long term speech spectrum, starting at 500 Hz. After transforming the input speech signal into spectrogram by Fourier analysis, the envelope is extracted in 22 frequency bands by summering the power within each band. These three steps are same as standard ACE strategy, hence we define it as ACE (although ACE has additional steps such as channel selection). Then the KEM is applied to the spectrum envelope on a frame by frame basis in each channel to reduce the Gaussian noise. The enhanced ACE spectral information of speech then can be further analysed by PCA and ICA. In order to produce stimuli for CI, inverse ICA is used to transform the modified independent components back to envelopes for CI stimulation. Finally, appropriate electrodes are selected and used to stimulate the auditory neurons.

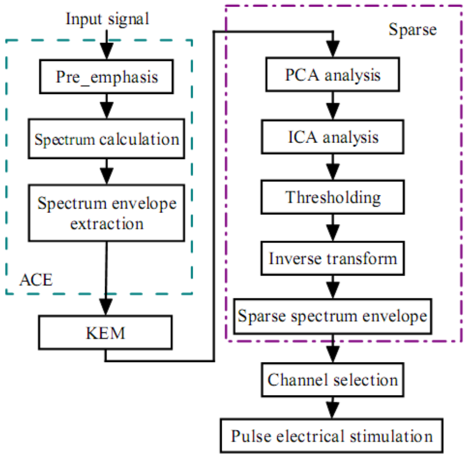

Figure 1 – Enhanced SPARSE strategy

4. EXPERIMENTS AND RESULTS

Objective and subjective experiments were performed to evaluate the enhanced SPARSE algorithm. In subjective listening tests, Bamford-Kowal-Bench (BKB) sentences [11] were used as the clean speech. BKB sentence lists are standard British speech materials with 21 lists and each list contains 50 keywords in 16 sentences. Babble noises at three different signal to noise ratios (SNR) (0, 5, 10 dB) was added to the speech material. Figure 2 shows the structure of experiments.

The method to get the spectral envelopes of the input signal is the same as that in ACE as described in section 3. For comparison, four different strategies were designed by applying different combination of ACE, KEM and SPARSE processes on the spectrum envelope. The ACE envelope and SPARSE envelope are the original spectrum envelope matrices obtained from ACE and SPARSE respectively; enhanced ACE is the KEM enhanced ACE spectrum envelope matrix, and enhanced SPARSE is the matrix obtained with KEM and SPARSE together. These matrices are all in the spectro-temporal domain and thereafter appropriate electrodes can be selected to stimulate the auditory neurons. Here, in order to simulate the perception of a CI user, the output signal is a reconstruction of the acoustical signal based on the spectrum envelope by using a vocoder [12].

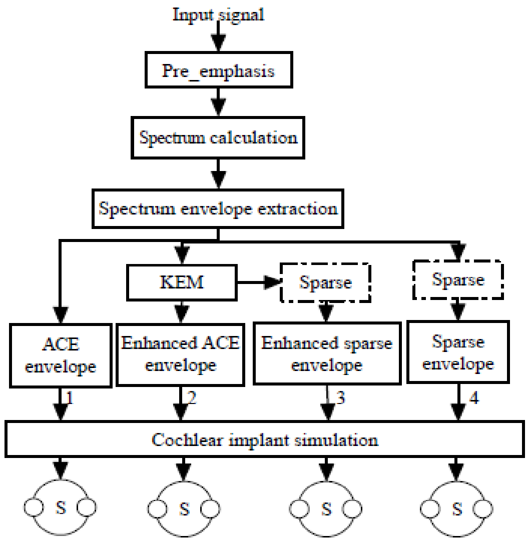

Figure 2 – Flow chart showing the structure of experiments. Four different strategies were compared

4.1 Objective experiments and results

To improve the performance of CI users in noisy environments, two factors are considered in the proposed algorithm: one is speech enhancement (noise reduction) and the other is increased sparseness of the reconstructed signal. An important goal of these algorithms is to transform the stimuli to be in a more sparse distribution in order to resemble the natural code of auditory neurons better. Sparseness can be quantified by the kurtosis of the signal [7] based on equation (7):

$$K = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i - \mu}{\sigma}\right)^4 - 3 \qquad (7)$$

where x is the amplitude of the signal, μ is the mean and σ is the standard deviation. For a normalised Gaussian (non-sparse) distribution with $\mu = 0$ and $\sigma = 1$, the kurtosis is (by definition) $K = 0$; for other signals the kurtosis may be larger than zero for a super-Gaussian or smaller than 0 for a sub-Gaussian process. If the kurtosis becomes larger than the sparseness of the stimuli is increased.

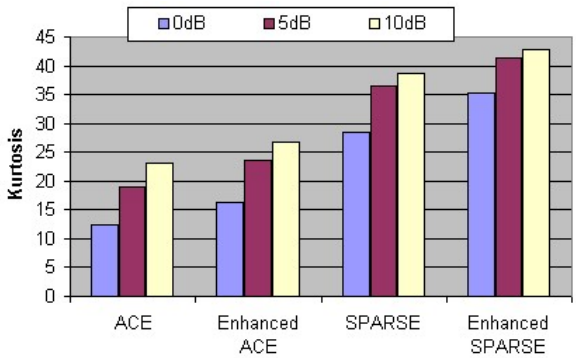

Figure 3 – Kurtosis of speech processed by four strategies at three SNR levels of 0, 5 and 10 dB

Figure 3 shows the kurtosis of the simulated sounds of speech processed by four strategies. The evaluation of sparseness takes the simulated output waveforms as a whole and calculates the kurtosis of the entire time series. These results are consistent with the results of [7] in that the output of the SPARSE algorithm is more sparse than the output of ACE algorithm. Figure 3 also shows that the kurtosis values of enhanced SPARSE and enhanced ACE are higher than their corresponding non-enhanced algorithms. The KEM enhanced signal is therefore more sparse than the output of the non-enhanced algorithm. We expect the enhanced SPARSE algorithm should further improve the performance of CI users.

Besides kurtosis, the other objective criterion used in this paper is the SNR matrix which is calculated as

$$SNR_{\hat{\mathbf{X}}} = 10\log 10 \frac{(\mathbf{X}_{clean})^2}{(\mathbf{X}_{clean} - \hat{\mathbf{X}})^2} \tag{8}$$

where $\mathbf{X}_{clean}$ and $\hat{\mathbf{X}}$ are the clean and estimated spectrum envelope matrices. For example, for ACE, $\mathbf{X}_{clean}$ is the clean speech ACE spectrum envelope matrix and $\hat{\mathbf{X}}$ can be either the ACE spectrum envelope of noisy speech or that of the KEM enhanced speech, which will result in the noisy ACE SNR matrix (ACE (noisy)) and the enhanced ACE SNR matrix (enhanced ACE (noisy)) respectively as shown in figure 4 (b).

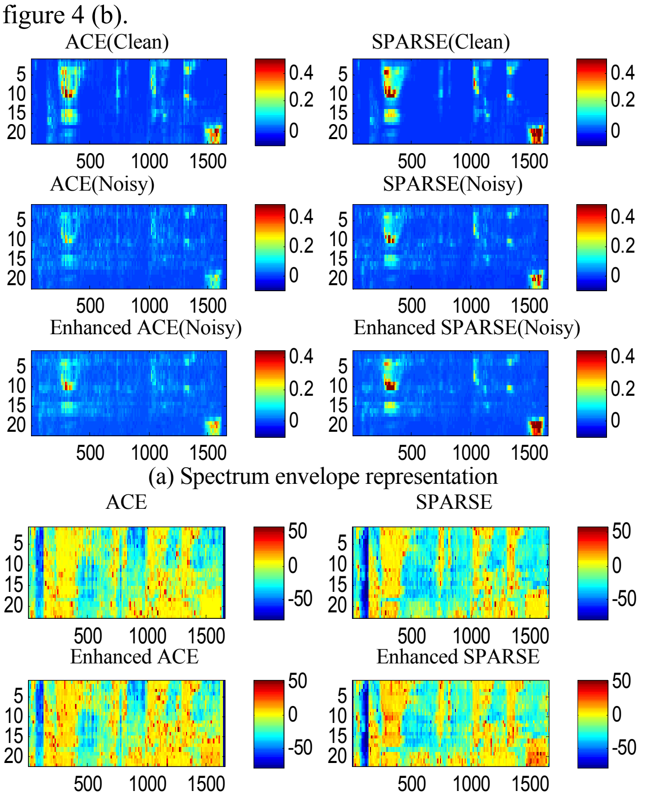

Figure 4 – (a) Spectrum envelope representation (b) SNR matrix after processing by four strategies in the condition of 0 dB SNR input. The sentence is "the clown had a funny face". The vertical axis is the channel index and the horizontal axis is the time index

Figure 4(a) shows the spectrogram of the 22-channel spectrum envelope under different conditions. Figure 4(b) is the corresponding SNR matrix. In figures 4, the left column is calculated using ACE, and the right column is calculated using SPARSE. In figure 4(a), the first row shows the spectral envelope for clean speech, the second row for noisy speech at 0 dB and the third row is for the KEM de-noised speech. In figure 4(b), the top row shows the SNR matrix for noisy speech and the bottom row for enhanced speech.

Comparing both the spectral envelope and the SNR matrices of enhanced ACE and enhanced SPARSE in figure 4 to their corresponding originals, it is evident that the enhanced speech has more energy in the speech segments than the non-enhanced speech. The SNR at the speech segments are also improved. This is encouraging, as it might imply potential for higher speech recognition rates for CI users.

4.2 Subjective experiments and results

Speech intelligibility is often used as a criterion to evaluate the performance of speech perception [13]. In this paper, the speech intelligibility was tested on speech processed by four strategies at three SNR levels of 0, 5 and 10 dB, where babble noise was used. The experiment was approved by The Human Experimentation Safety and Ethics Committee, Institute of Sound and Vibration Research, University of Southampton, Southampton, UK and consent forms were obtained from all subjects. The experiments consisted of two parts: measuring the percentage of correctly identified keywords and an open questionnaire, in which the participants were asked to describe the sound quality of different speech signals in their own words.

Six normal hearing subjects with no previous experience of the BKB sentence lists (5 males, 1 female, aged between 21 to 28 years) participated in the experiments. The participants were trained with three BKB lists of clean speech processed by ACE or SPARSE strategies to familiarise with the test procedure. In the formal test, each participant was presented with four different classes of simulated speech-in-noise material with different SNR (0, 5, 10 dB) in babble noise. Correct keyword recognition rates were calculated. To avoid memorising keywords, each subject listened at each condition (4 algorithms × 3 SNR) to different, randomly selected sentence lists (one or two). To full use the BKB sentences and put more emphasis on SPARSE strategies, 2 BKB sentences lists were used in the SPARSE and in the enhanced SPARSE strategies while only one list for ACE and enhanced ACE in each condition. The average score of all tested sentence lists in the same condition was recoded as the percent correct for each subject in each condition. The listeners sat in a sound-isolated room in front of a computer screen and listened to the sentences through headphones (Sennheiser HD380 Pro). The presentation level for each participant was set to 60 dB SPL. Participants were also asked to fill in a questionnaire to describe the sound quality in their own words.

Table 1 shows the results of the subjective listening experiments. The enhanced ACE produced better performance than the original ACE for all three SNR, but the improvement is slight and decreased with increasing SNR. The enhanced SPARSE produced better performance at low SNR (0 dB) than SPARSE, but the advantage is not obvious for the higher SNR (5, 10 dB). Although the results reflect the general trend that the recognition rate increases with increasing SNR, the differences were not statistically significant different (2-way-ANOVA, $p > 0.05$, partial Eta squared = 0.32, power = 0.48). Experiments with more participants are needed to confirm the trends. Subjects appeared to have reached a performance limit in the two higher SNR (5, 10 dB) conditions, explaining the small difference between ACE and SPARSE, enhanced ACE and ACE, enhanced SPARSE and SPARSE in these two SNR.

Table 1 – Correct key word rates for different algorithms

Correct rate (%) / SNR	ACE	Enhanced ACE	SPARSE	Enhanced SPARSE
0 dB	52.7	56.7	62.8	68.3
5 dB	92.7	94.3	92.0	92.5
10 dB	97.0	97.7	97.8	98.3

In the open questionnaires, three out of six NH participants rate the quality of the noisy SPARSE higher than noisy ACE; enhanced ACE was rated higher than noisy ACE, especially for low SNR. As to the noisy SPARSE and the enhanced SPARSE, most of them described that they were subjectively indistinguishable except for the 0 dB SNR.

5. CONCLUSIONS

Normal hearing listeners understand speech well in a noisy environment, but this is a very challenging situation for cochlear implant users. We propose here an enhanced SPARSE algorithm for cochlear implant speech processing, which aims to reduce noise with multiple types of distribution as well as using sparsification to deliver key information to CI users via limited frequency channels. The algorithm sequentially combines two algorithms of KEM and SPARSE. The objective and subjective experiments showed some improvement from this combination. Further research is required into the family of sparsifying algorithms, as the most suitable objective evaluation criterion for sparse processing is still unknown. Finally, it is essential to perform experiments with experienced CI users to evaluate the described algorithms in a real world listening environment. Since higher kurtosis has been associated with higher speech recognition rates for CI users, we expect that the enhanced SPARSE algorithm should improve the speech perception of CI users; this will be evaluated in future work.

6. ACKNOWLEDGEMENTS

This work was supported by the European Commission within the ITN AUDIS (grant agreement number PITN-GA-2008-214699). The authors appreciate Cochlear Europe Ltd. providing the NIC software and participants' hard work in subjective experiments.

APPENDIX

The logarithm of joint probability distribution function (pdf) $p(\mathbf{x}_{1:N}, \mathbf{z}_{1:N})$ can be expressed as:

$$\log p(\mathbf{x}_{1:N}, \mathbf{z}_{1:N}) = \sum_{t=1}^{N} \log p(\mathbf{x}_t \mid \mathbf{x}_{t-1}) + \sum_{t=1}^{N} \log p(z_t \mid \mathbf{x}_t)$$

$$= \frac{1}{2g_s} \sum_{t=1}^{N} (s_t + \boldsymbol{\alpha}^T \mathbf{x}_{t-1}) - \frac{N}{2} \log g_s - \frac{1}{2g_v} \sum_{t=1}^{N} (z_t - s_t)^2 - \frac{N}{2} \log g_v + \text{const}$$

The optimization function is:

$$Q(\lambda, \lambda^{old}) = E_x \log[p(\mathbf{x}_{1:N}, \mathbf{z}_{1:N} \mid \lambda) \mid \mathbf{z}_{1:N}, \lambda^{old}]$$

$$\begin{aligned} &= -\frac{N}{2} \log g_v - \frac{1}{2g_v} \sum_{t=1}^{N} (z_t^2 - 2 z_t \hat{s}_t + \hat{s}_t^2) \\ &\quad -\frac{N}{2} \log g_s - \frac{1}{2g_s} \sum_{t=1}^{N} (\hat{s}_t^2 + 2\boldsymbol{\alpha}^T \widehat{\mathbf{x}_{t-1} s_t} + \boldsymbol{\alpha}^T \widehat{\mathbf{x}_{t-1}\mathbf{x}_{t-1}} \boldsymbol{\alpha}) \end{aligned} \quad (9)$$

Differentiation of Equation (9) to the parameters $\lambda = \{\boldsymbol{\alpha}, g_s, g_v\}$ respectively,

$$\frac{\partial Q(\lambda, \lambda^{old})}{\partial \boldsymbol{\alpha}^T} = -\frac{1}{2g_s} \sum_{t=1}^{N} (2\widehat{\mathbf{x}_{t-1} s_t} + 2\widehat{\mathbf{x}_{t-1}\mathbf{x}_{t-1}} \boldsymbol{\alpha}) \quad (10)$$

$$\frac{\partial Q(\lambda, \lambda^{old})}{\partial g_s} = -\frac{1}{2g_s} + \frac{1}{2g_s^2} \sum_{t=1}^{N} (\hat{s}_t^2 + 2\boldsymbol{\alpha}^T \widehat{\mathbf{x}_{t-1} s_t} + \boldsymbol{\alpha}^T \widehat{\mathbf{x}_{t-1}\mathbf{x}_{t-1}} \boldsymbol{\alpha}) \quad (11)$$

$$\frac{\partial Q(\lambda, \lambda^{old})}{\partial g_s} = -\frac{1}{2g_v} - \frac{1}{2g_v^2} \sum_{t=1}^{N} (z_t^2 - 2 z_t \hat{s}_t + \hat{s}_t^2) \quad (12)$$

Equations (4)(5)(6) can be derived by setting equations (10)-(12) equal to zero.

REFERENCES

[1] S. Greenberg, W. A. Ainsworth, A. N. Popper *et al.*, "Speech Processing in the Auditory System: An Overview," *Speech Processing in the Auditory System*, Springer Handbook of Auditory Research, pp. 1-62, New York: Springer, 2004.

[2] B. A. Olshausen, and D. J. Field, "Sparse coding of sensory inputs," *Current Opinion in Neurobiology,* vol. 14, pp. 481-487, 2004.

[3] G. Li, "Speech perception in a sparse domain," PhD thesis, Institute of Sound and Vibration, University of Southampton, Southampton, 2008.

[4] A. Hussain, M. Chetouani, S. Squartini *et al.*, "Nonlinear Speech Enhancement: An Overview " *Lecture Notes in Computer Science*, Progress in Nonlinear Speech Processing Y. Stylianou, M. Faundez-Zanuy and A. Esposito, eds., Vol. 4391, pp. 217-248, New York: Springer, 2007.

[5] W. J. J. Roberts, Y. Ephraim, and H. Lev-Ari, "A Brief Survey of Speech Enhancement," *Microelectronics (2nd Edition)*, J. C. Whitaker, ed., pp. 20-1–20-11: CRC Press, 2006.

[6] S. Gannot, D. Burshtein, and E. Weinstein, "Iterative and sequential Kalman filter-based speech enhancement algorithms," *Speech and Audio Processing, IEEE Transactions on,* vol. 6, no. 4, pp. 373-385, 1998.

[7] G. Li, and M. E. Lutman, "Sparse stimuli for cochlear implants," in 16th European Signal Processing Conference (EUSIPCO 2008), Lausanne, Switzerland, 2008.

[8] K. Nie, W. Drennan, and J. Rubinstein, "Cochlear Implant Coding Strategies and Device Programming," *Ballenger's otorhinolaryngology: head and neck surgery*, J. B. Snow, P. A. Wackym and J. J. Ballenger, eds., pp. 389-394: People's Medical Publishing House, 2009.

[9] K. Kasturi, P. C. Loizou, M. Dorman *et al.*, "The intelligibility of speech with ``holes" in the spectrum," *The Journal of the Acoustical Society of America,* vol. 112, no. 3, pp. 1102-1111, 2002.

[10] M. Cooke, "A glimpsing model of speech perception in noise," *J. Acoust. Soc. Am.,* vol. 119, pp. 1562-1573, 2006.

[11] J. Bench, A. Kowal, and J. Bamford, "The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children," *Br J Audiol,* vol. 13, no. 3, pp. 108-12, Aug, 1979.

[12] R. V. Shannon, F.-G. Zeng, V. Kamath *et al.*, "Speech Recognition with Primarily Temporal Cues," *Science* vol. 270, no. 5234, pp. 303-304 1995.

[13] H. Hu, J. Sang, M. Lutman *et al.*, "Simulation of hearing loss using compressive gammachirp auditory filters," in ICASSP 2011, Prague, Czech Republic, pp. 5428-5431, 2011.