

Article

Multi-Band Spectral Subtraction Method for Electrolarynx Speech Enhancement

Sheng Li ^{1,2}, MingXi Wan ^{1,*} and SuPin Wang ¹

¹ Key Laboratory of Biomedical Information Engineering of Ministry of Education, Department of Biomedical Engineering, School of Life Science and Technology, Xi'an Jiaotong University, Xi'an 710049, P.R. China; E-mails: sheng@mail.xjtu.edu.cn (S.L.); spwang@mail.xjtu.edu.cn (S.P.W.)

² Department of Biomedical Engineering, The Fourth Military Medical University, Xi'an 710032, P.R. China

* Author to whom correspondence should be addressed; E-mail: mxwan@mail.xjtu.edu.cn (M.X.W)

Received: 30 October 2008; in revised form: 6 February 2009 / Accepted: 25 February 2009 /

Published: 13 March 2009

Abstract: Although the electrolarynx (EL) provides an important means of voice reconstruction for patients who lose their vocal cords by laryngectomies, the radiated noise and additive environment noise reduce the intelligibility of the resulting EL speech. This paper proposes an improved spectrum subtract algorithm by taking into account the non-uniform effect of colored noise on the spectrum of EL speech. Since the over-subtraction factor of each frequency band can be adjusted in the enhancement process, a better noise reduction effect was obtained and the perceptually annoying musical noise was efficiently reduced, as compared to other standard speech enhancement algorithms.

Keywords: Electrolarynx, multi-band, speech enhancement, spectral subtraction method.

1. Introduction

It is well known that the larynx is the main organ of natural voice production for normal human beings to communicate, however, there are still many patients who suffer from larynx diseases or have no larynx (due to laryngectomies, which are the total removal of the larynx for the treatment of laryngeal cancer, other laryngeal lesions, and so on), which would lead to the loss of their vocal function. Therefore, various speech production substitutes have been developed and used practically in order to reconstruct speech functions [1-4], including the esophageal speech and the trachea-

esophageal speech production methods, Tapia's artificial larynx, and the electrolarynx (EL) [1]. Comparing these post-laryngectomy forms of speech production, the electrolarynx, which is a hand-held battery-powered device that is typically held against the neck at the level of the former glottis to excite the vocal tract acoustically [5], has several special advantages, such as being easier to use, production of longer sentences without special care, and being more effective in communication than other methods of voice rehabilitation in many situations. Therefore, the electrolarynx is the main method adopted for voice rehabilitation after laryngectomies, and has been clinically proven to be an important method of oral communication for laryngectomees for more than 50 years.

Unfortunately, both the EL device and the resulting speech have several serious shortcomings, especially the radiated noise during phonation, which is due to the fact that part of the sound energy produced is not able to pass through the neck tissue but rather is radiated directly from the instrument itself. Data from previous studies reported that this radiated noise was about 20-25 dB when the mouth was closed [5], and varied over 4-15 dB across subjects for the same device [6]. This radiated noise, therefore, may not only cause a stronger concentration of noise energy between 400-1 KHz and 2-4 kHz in EL speech, but also result in loss of speech intelligibility [7]. In addition, the masking effect of the noise can contribute to the unnaturalness and poor quality of EL speech.

Besides the acoustic shielding technology which is applied on the EL device itself [8], more and more signal processing techniques have been developed to improve EL speech [9-12]. To summarize these papers, two main speech enhancement algorithms have been developed to reduce the radiated noise of EL speech, and improve its intelligibility and naturalness. One method is the subtractive-type algorithm [12, 13], which is the most widely used, and has been shown to be an effective approach for noise canceling. Due to the simplicity of implementation, and low computational load, the spectral subtraction method is the primary choice for real time applications [9, 12]. In general, this method enhances the speech spectrum by subtracting an average noise spectrum from the noisy speech spectrum; here the noise is assumed to be uncorrelated and additive to the speech signal. The phase of the noisy speech is kept unchanged, since it is assumed that the phase distortion is not perceived by human ear. However, a serious drawback of this method is that the enhanced speech is accompanied by an unpleasant musical noise artifact which is characterized by tones with random frequencies. Although many solutions have been proposed to reduce the musical noise in the subtractive-type algorithms [11, 14-17], results performed with these algorithms show that there is a need for further improvement, especially under very low signal-noise ratio (SNR) conditions. Another method uses adaptive noise canceling [10, 18], which removes the noise components of the primary input signal that depend on the reference input signal and are based on second-order statistics. However, many other noise components in the primary input signal, which depend on the noise reference signal through higher-order statistics may exist in the enhanced speech, which affects the effect of noise reduction.

Also, a previous perceptual study of EL speech suggested that the intelligibility of EL speech would be decreased if the signal-to-noise ratio (SNR) decreased [19]. It is also noted that the low-energy EL speech can be easily masked by the different environment noises [12], the reduction of speech quality, furthermore, causes listener fatigue. Therefore, it is important to investigate a new efficiently algorithm to eliminate both additive noise and the radiated noise. This would not only improve the life quality of the laryngectomees, since the EL is the only equipment for them to communicate, but also

increase the quality of the EL speech for better understanding, especially in environment noise and electronic noise (low SNR) conditions.

Unlike white Gaussian noise, which has a flat spectrum, the spectrum of EL noise and the additive environment noise are not flat. Thus, the noise signal does not affect the speech signal uniformly over the whole spectrum. Some frequencies are affected more adversely than others. This means that this kind of noise is “COLORED”. In order to prevent destructive subtraction of the speech while removing most of the residual noise, it is necessary to propose a non-linear approach to improve the subtraction procedure.

Therefore, this investigation was motivated by the need to improve EL speech, especially in electronic environments. A multi-band spectral subtraction algorithm is proposed that takes into account the variation of signal-to-noise ratio across the speech spectrum using a different over-subtraction factor of each frequency band to reduce colored noise. According to the features of the EL speech, we also give some recommended over-subtraction factors to effectively improve the EL speech quality.

2. Method

2.1 Experiments

Two male laryngectomees with total removal of the larynx participated in the experiment. They were native male speakers of Mandarin Chinese, and were 57 and 70 years of age and had 6 and 2 years of experience using the EL, respectively. The participants had recovered from the fibrosis and edema resulting from radiation, and their neck tissue was supple enough so as to permit them to use EL effectively.

A Hu-Die brand EL made in China was used in this experiment. This device can provide a frequency option of 60 and 90 Hz and a sound intensity of 70 – 80 dB. The recording procedure was carried out in a soundproof room. Speech samples were collected by using a microphone mounted at a distance of about 10 cm from the mouth of the laryngectomees, and amplified by a multi-channel conditioning amplifier (Brüel & Kjør, Denmark). Recordings were taken at a sampling frequency of 20 kHz with 16-bits per sample. Five Chinese sentences, each of which was composed of six words, were used as the speech materials for acoustic and perceptual analyses. Instructions were given to the speakers before the recording took place. The speakers were instructed to read the speech materials three times at normal loudness and speaking rate.

Two kinds of background noise: white Gaussian noise and speech babble noise, taken from the Noisex-92 database, were chosen in the experiments for the research of EL speech enhancement in noise, since these two representative noises have a greater similarity than the other noises to the actual talking conditions of the laryngectomees using the EL. Noise was added to the original EL speech signal with a varying SNR.

For the perceptual experiment, eight listeners were selected to evaluate the acceptability of each sentence based on the criteria of the mean opinion score (MOS), which is a five-point scale (1: bad; 2: poor; 3: common; 4: good; 5: excellent). All of the listeners were native speakers of Mandarin Chinese, had no reported history of hearing problems, as were unfamiliar with EL speech. Their ages

varied from 22 to 36, with a mean age of 26.37 (SD=4.63). The listening tasks took place in a sound-proof room, and the speech samples were presented to the listeners at a comfortable loudness level (60 dB SPL) via a high quality headphone. A 4-s pause was inserted before each citation word, and the order in which the speech samples were presented was randomized, to allow the listeners to respond and to avoid rehearsal effects.

2.2 Multi-band spectral subtraction method

The multi-band spectral subtraction method is based on the assumption that the additive noise will be stationary and uncorrelated with the clean speech signal. If $y(n)$, the noisy speech, is composed of the clean speech signal $s(n)$ and the uncorrelated additive noise signal $d(n)$, then:

$$y(n) = s(n) + d(n) \quad (1)$$

The power spectrum of the corrupted speech can be approximately estimated as:

$$|Y(\omega)|^2 \approx |S(\omega)|^2 + |D(\omega)|^2 \quad (2)$$

where $|Y(\omega)|^2$, $|S(\omega)|^2$ and $|D(\omega)|^2$ represent the noisy speech short-time spectrum, the clean speech short-time spectrum, and the noise power spectrum estimate, respectively.

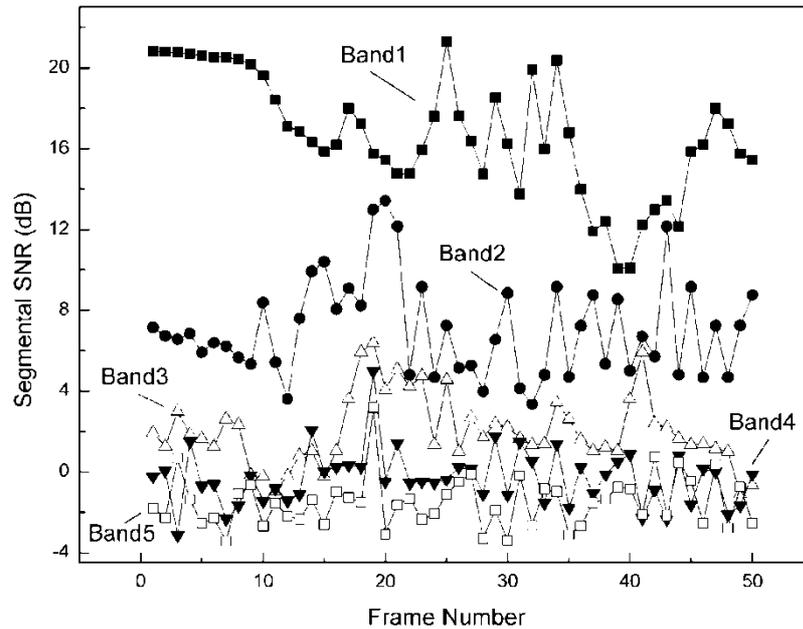
Most of the subtractive-type algorithms have different variations allowing for flexibility in the variation of the spectral subtraction. Berouti *et al.* [20] proposed the generalized spectral subtraction scheme is described as follows:

$$|\hat{S}(\omega)|^\gamma = \begin{cases} |Y(\omega)|^\gamma - \alpha |\hat{D}(\omega)|^\gamma, & \text{if } \frac{|\hat{D}(\omega)|^\gamma}{|Y(\omega)|^\gamma} < \frac{1}{\alpha + \beta} \\ \beta |\hat{D}(\omega)|^\gamma, & \text{otherwise,} \end{cases} \quad (3)$$

where α ($\alpha > 1$) is the over-subtraction factor [16][20], which is a function of the segmental SNR. β ($0 \leq \beta \leq 1$) is the spectral floor, and γ is the exponent determining the transition sharpness. Here we set $\gamma = 2$, and $\beta = 0.002$.

This implementation assumes that the noise affects the speech spectrum uniformly, the over-subtraction factor α , furthermore, subtracts an over-estimate of the noise over the whole spectrum. However, the noise in the EL speech maybe colored and does not affect the speech signal uniformly over the entire spectrum. Figure 1 shows the estimated segmental SNR for five frequency bands [0~300 Hz (Band 1), 300~1 KHz (Band 2), 1K~2K (Band 3), 2K~3K (Band 4), 3K~5K (Band 5)] of EL speech corrupted by radiated noise. It can be seen from Figure 1 that the SNR of the low frequency bands (Bands 1, 2) was significantly higher than the SNR of the high frequency bands (Bands 3~5). The largest SNR difference among the SNR was about 25 dB, a large difference. This phenomenon suggests that the noise signal does not affect the speech signal uniformly over the whole spectrum, therefore, subtracting a constant factor of noise spectrum over the whole frequency spectrum may remove speech also.

Figure 1. The segmental SNR of five frequency bands of EL speech.

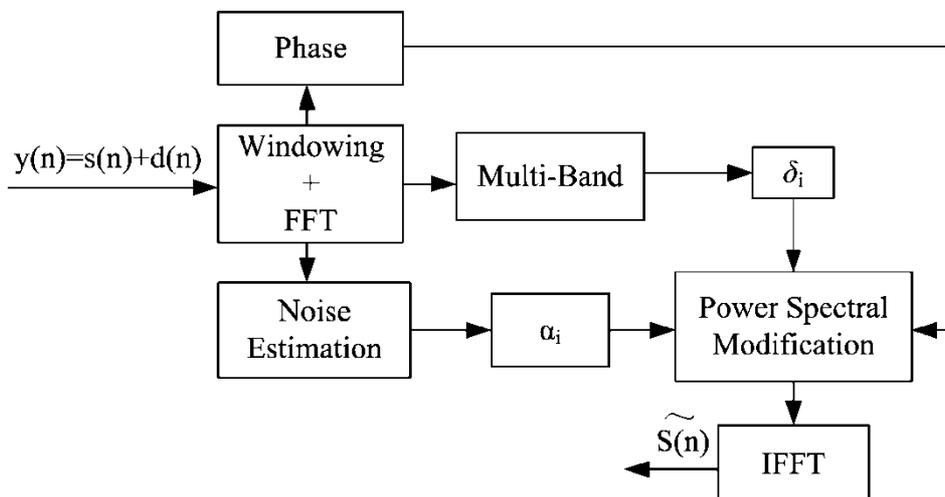


In order to take into account the fact that colored noise affects the EL speech spectrum differently at various frequencies, the multi-band spectral subtraction technique [16] is used in this proposed approach, which estimates a suitable factor that will subtract just the necessary amount of the noise spectrum from each frequency sub-band. In this study, the speech spectrums was divided into N (N=5) non-overlapping bands, and spectral subtraction was performed independently in each band. Hence the estimate of the clean speech spectrum in the i th band is obtained by:

$$|\hat{S}_i(\omega)|^2 = |Y_i(\omega)|^2 - \alpha_i \delta_i |\hat{D}_i(\omega)|^2, \quad b_i \leq \omega \leq e_i \quad (4)$$

where α_i is the over-subtraction factor of the i th frequency band, and δ_i is a tweaking factor that can be individually set for each frequency band to customize the noise removal properties. b_i and e_i are the beginning and ending frequency of the i th frequency band. The whole algorithm is as shown in Figure 2.

Figure 2. The proposed speech enhancement scheme.



The band specific over-subtraction factor α_i is a function of the segmental noisy signal to noise ratio SNR_i of the i th frequency band which is calculated as:

$$SNR_i (dB) = 10 \log_{10} \frac{\sum_{k=b_i}^{e_i} |Y_i(\omega)|^2}{\sum_{k=b_i}^{e_i} |\hat{D}_i(\omega)|^2} \quad (5)$$

According to the SNR_i value calculated in Eq. (5) and shown in Figure 1, also consistent with Kamath *et al.* [16] and Udrea *et al.* [17], the over-subtraction factor α_i is calculated as:

$$\alpha_i = \begin{cases} 5 & SNR_i < 5 \\ 4 - \frac{3}{20}(SNR_i) & -5 \leq SNR_i \leq 20 \\ 1 & SNR_i > 20 \end{cases} \quad (6)$$

The use of this over-subtraction factor α_i can provides a degree of control over the noise subtraction level in each band. Another factor δ_i , which is shown in Equation (4) can be used to provide an additional degree of control within each band, since most of the speech energy is present in the lower frequencies, smaller δ_i values were used for the low frequency bands in order to minimize speech distortion. The values of δ_i were empirically determined and set to:

$$\delta_i = \begin{cases} 1 & 60Hz \leq f_i \leq 300Hz \\ 1.3 & 0.3KHz < f_i \leq 1KHz \\ 1.6 & 1KHz < f_i \leq 2KHz \\ 1.8 & 2kHz < f_i \leq 3kHz \\ 1.3 & 3kHz < f_i \leq 5kHz \end{cases} \quad (7)$$

Both factors, α_i and δ_i can be adjusted for each band for different speech conditions to get better speech quality.

2.3 Noise estimation

The noise in the EL speech, which included the additive radiated noise and environment noise, is a highly nonstationary noise, so it is imperative to update the estimate of the noise spectrum frequently. This study adopted the minimum-statistics method proposed by Cohen and Berdugo [21] for noise estimation, since this method is computationally efficient, robust with respect to the input SNR, and has an ability to quick follow any abrupt changes in the noise spectrum. The minimum tracing is based on a recursively smoothed spectrum which is estimated using first-order recursive averaging:

$$|\hat{D}_{(k,l)}(\omega)|^2 = \lambda_D |\hat{D}_{(k-1,l)}(\omega)|^2 + (1 - \lambda_D) |\hat{Y}_{(k,l)}(\omega)|^2 \quad (8)$$

$0 < \lambda_D < 1$, where $|\hat{D}_{(k,l)}(\omega)|^2$ and $|\hat{Y}_{(k,l)}(\omega)|^2$ are the k th components of noise spectrum and noisy speech spectrum at the frame l , and λ_D is a smooth parameter. Let $p'(k,l)$ denote the conditional signal presence probability in Cohen and Berdugo (2002)[21], then Eq. (8) implies:

$$\left| \hat{D}_{(k,l)}(\omega) \right|^2 = \hat{\lambda}_D(k,l) \left| \hat{D}_{(k-1,l)}(\omega) \right|^2 + (1 - \hat{\lambda}_D(k,l)) \left| \hat{Y}_{(k,l)}(\omega) \right|^2 \quad (9)$$

where $\hat{\lambda}_D(k,l) \triangleq \lambda_D + (1 - \lambda_D)p'(k,l)$ is a time-varying smoothing parameter. Therefore, the noise spectrum can be estimated by averaging past spectral power values. For more descriptive details of this algorithm, the reader is referred to [21, 22].

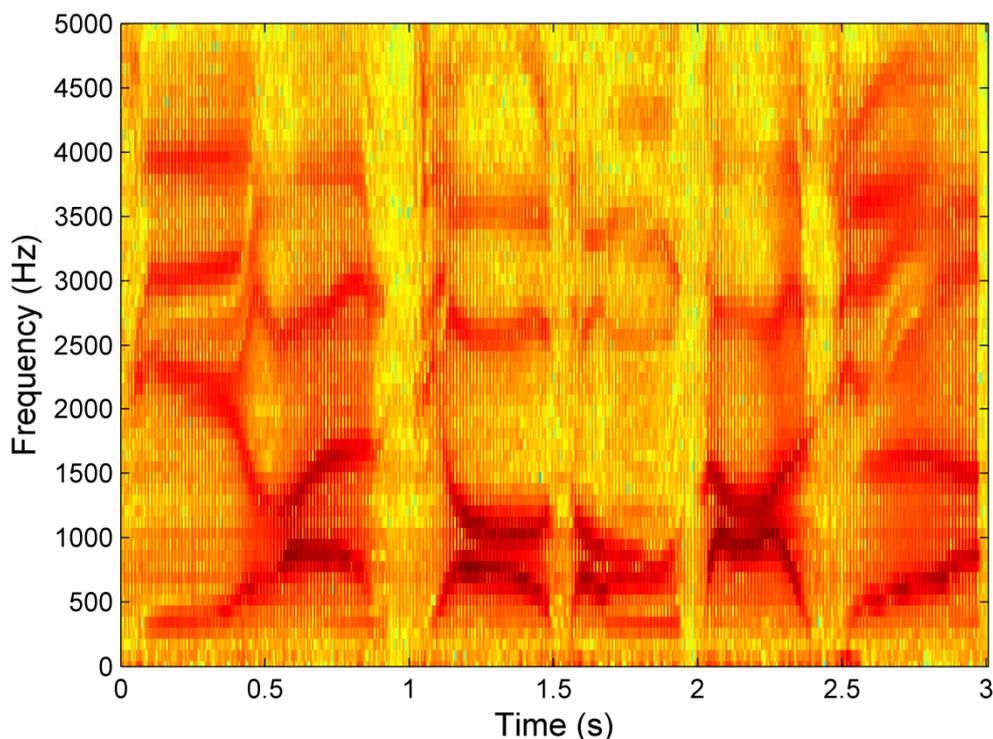
3. Results and Discussions

In order to evaluate and compare the performance of the proposed enhancement algorithm, three other algorithms are performed in this study, they are: traditional spectral subtraction method, basic Wiener filtering, and a noise-estimation algorithm [23]. For the purpose of analyzing the time-frequency distribution of the original/enhanced speech, speech spectrograms were provided since they have been identified as a well-suited tool for observing both the residual noise and speech distortion. In addition, results are also measured objectively by Signal-to Noise ratio (SNR) and subjectively by Mean Opinion Score (MOS) in conditions of different additive white Gaussian noise as well as Bobble noise (for MOS) for the algorithm evaluation.

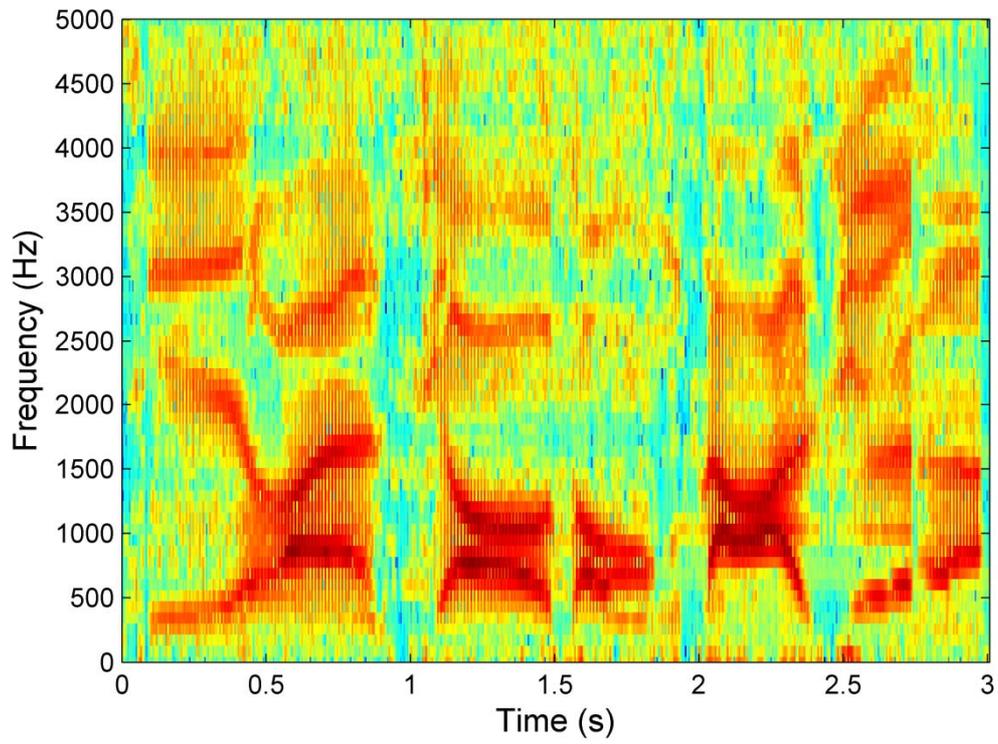
Figure 3 shows the spectrograms of the original EL speech (a), and the enhanced speech using traditional spectral subtraction algorithm (b), basic Wiener filtering (c), noise-estimation algorithm (d), and the proposed multi-band spectral subtraction algorithm in this study (e). The speech material is a Chinese sentence “Xi An Jiao Tong Da Xue” (‘Xi’an Jiaotong University’ in English).

Figure 3. Spectrogram of the sentence “Xi An Jiao Tong Da Xue”.

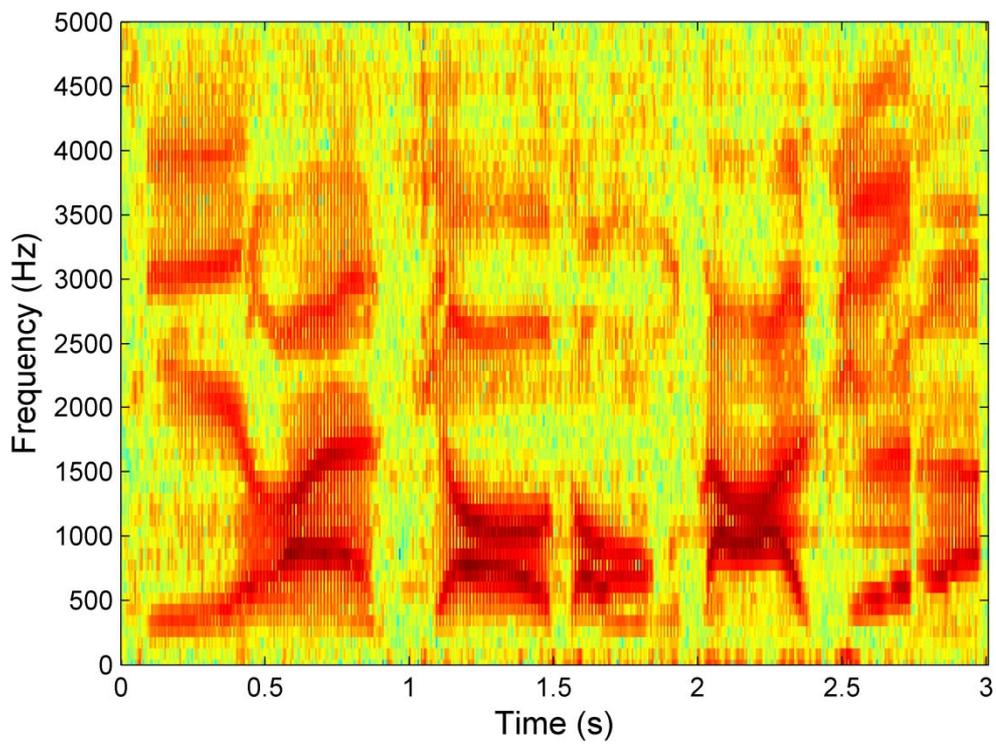
(a) The original EL speech.



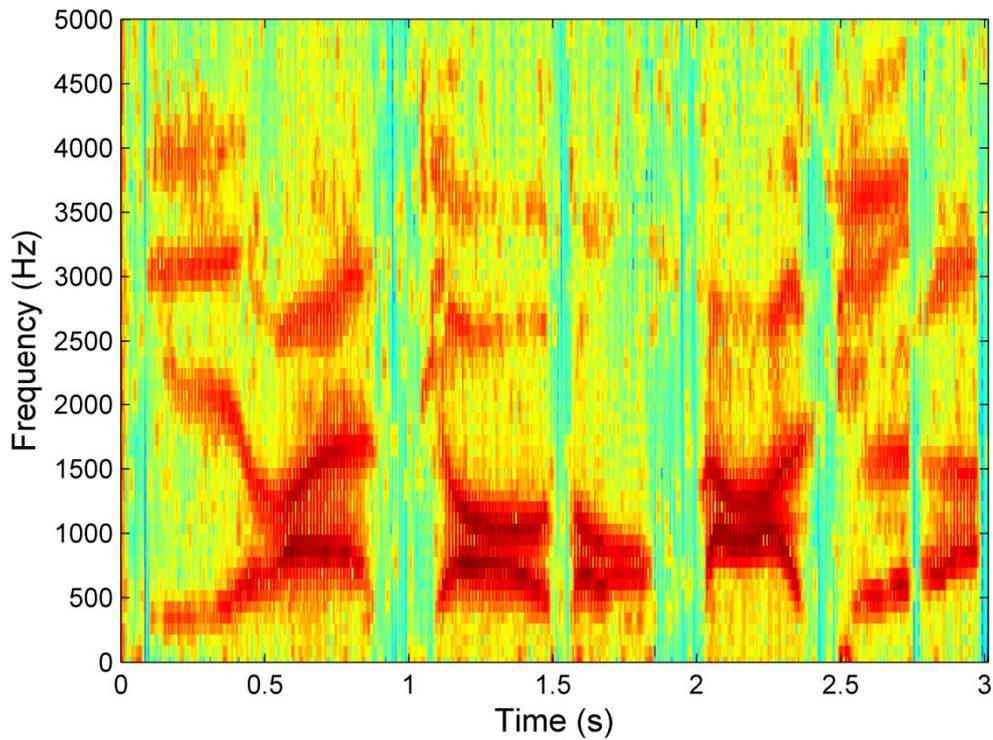
(b) Enhanced speech obtained by the traditional spectral subtraction method.



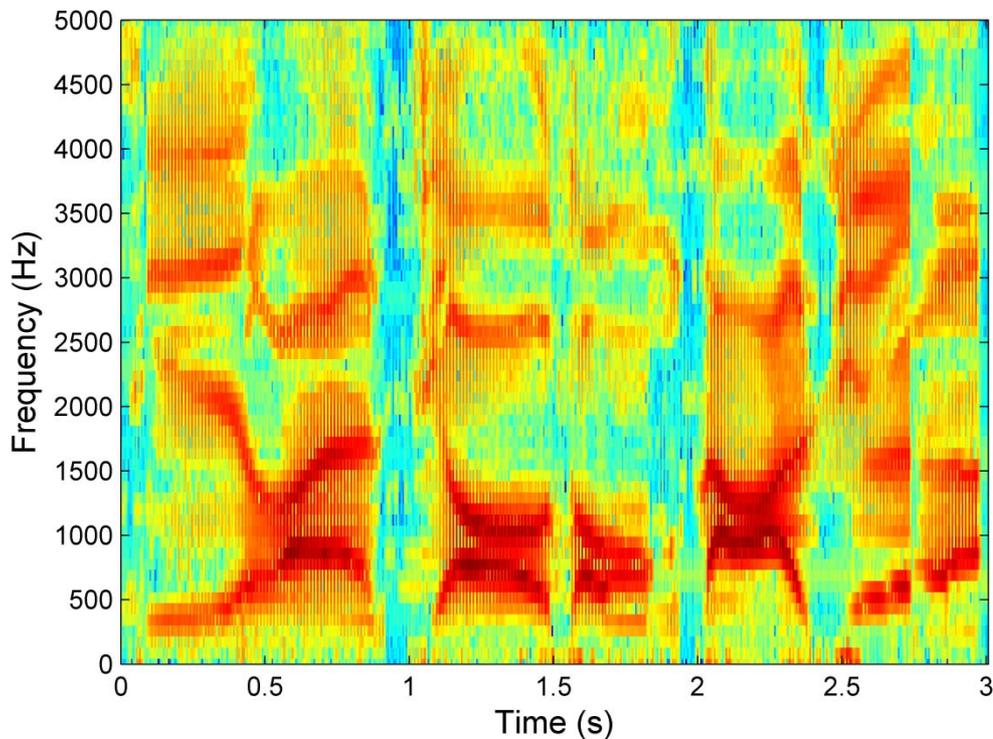
(c) Enhanced speech obtained by the basic Wiener filtering.



(d) Enhanced speech obtained by the noise-estimation algorithm.



(e) Enhanced speech obtained by the proposed multi-band spectral subtraction algorithm.

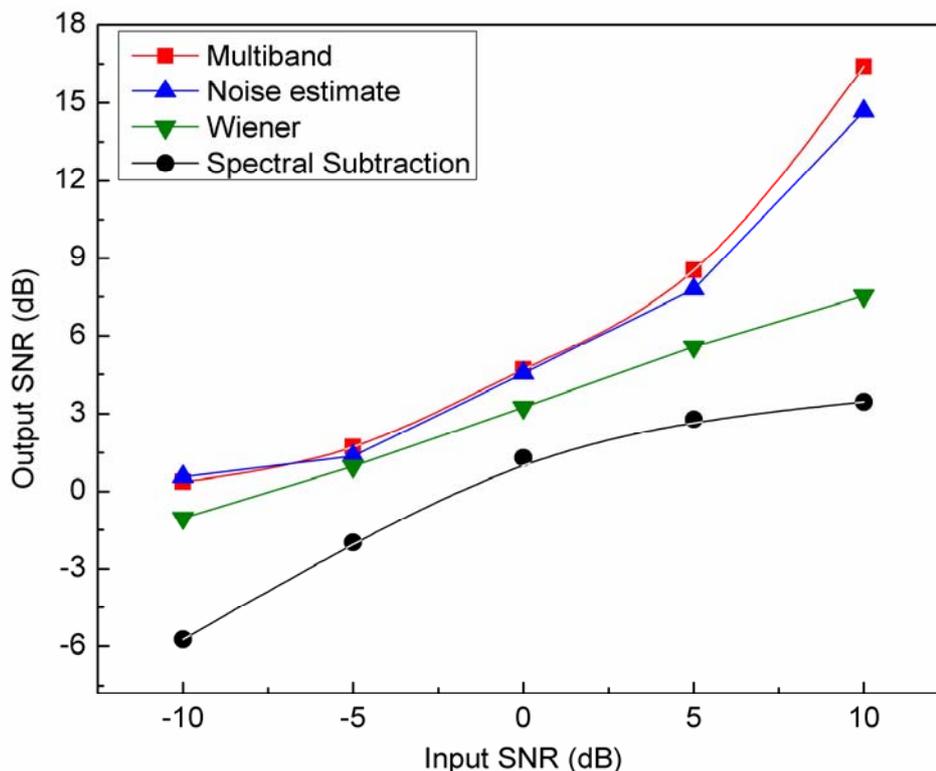


Due to its different speech production theory and working conditions, EL speech has some special attributes. As stated before, the most special is that radiated noise, the additive environment noise may also combined in the origin EL speech. These noises can be clearly seen from Figure 3(a), especially in speech-pause region. Figure 3(b) shows that the spectral subtraction algorithm is effective in reducing

the radiated noises, both in the speech and non-speech section. However, there are still too much residual noise remains in the enhanced speech, especially in the high-frequency section, suggesting that the noise reduction is not satisfactory. Figures 3(c)-(d) indicate that much noise still remains in the EL speech enhanced by the basic Wiener filtering and the noise-estimation algorithm, suggest that the noise reduction is not satisfactory. The proposed algorithm appears to be much better in reducing radiated noise as shown in Figure 3(e), since which can not only greatly reduces the low-frequency noise, but also eliminates the high-frequency noise completely. Observing the speech-pause regions, the residual noise is almost absent. Moreover, it is clearly visible that the residual noise is reduced to a large extent and has lost its structure. These results suggest that the proposed algorithm achieves a better reduction of the whole-frequency noise for EL speech than other enhancement methods.

The SNR measures were used for evaluation of the proposed multi-band spectral subtraction algorithm objectively. The results of SNR for the white noise experiments are shown in Figure 4, for five sentences produced by two male laryngectomees (these five SNRs are averaged). Methods compared included the traditional spectral subtraction algorithm (spectral subtraction), basic Wiener filtering (Wiener), noise-estimation algorithm (noise-estimation), and the proposed multi-band spectral subtraction algorithm (multi-band). From the figure, the proposed multi-band spectral subtraction method and the noise-estimation method have the best performance for this noise condition. Each of these gave about 6 dB improvement at the lower SNRs increasing to about 12 dB improvement at the higher SNRs than traditional spectral subtraction method. The multi-band method shows the best SNR improvement, and did slightly better than the noise-estimation method, especially for the higher SNR noise cases.

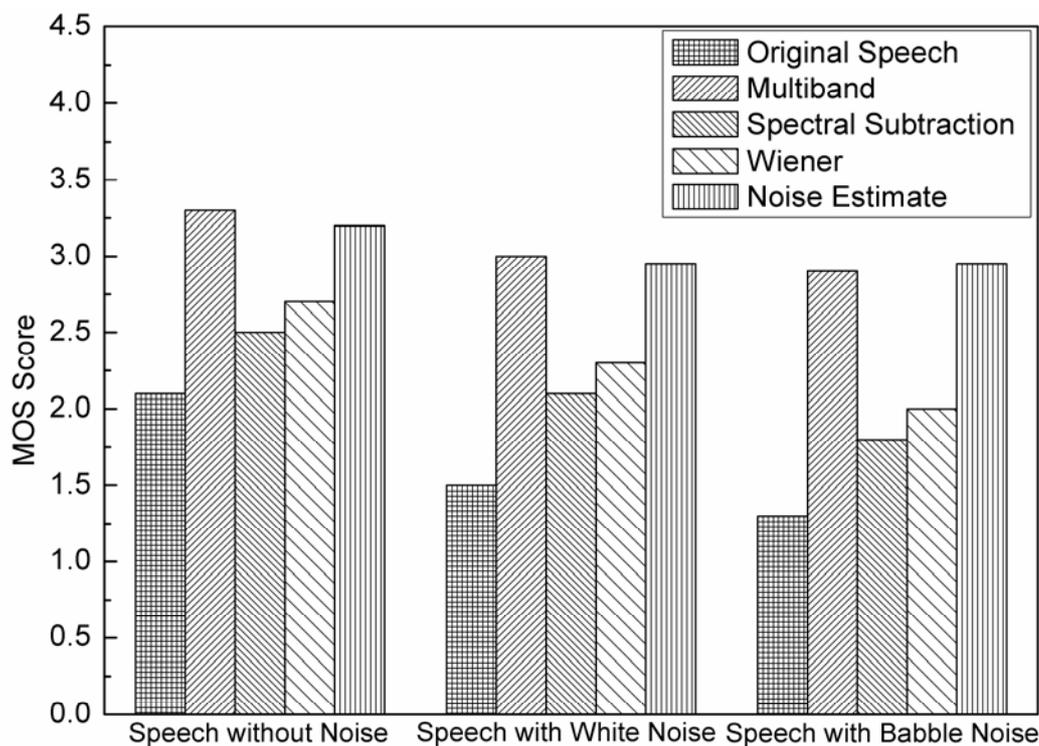
Figure 4. SNR results for white noise case at -10, -5, 0, +5, and +10 dB SNR levels



Subjective results using Mean Opinion Scores (MOS) for these enhancement algorithms for five sentences produced by two male laryngectomees (averaged) are shown in Figure 5. The noisy speech in the case of additive White and Babble noise has an input SNR of 0 dB. It can be seen from figure that the score of the enhanced speech obtained from using the multi-band algorithm is the highest, followed by that from the noise estimation and the Wiener filtering algorithms. This is true for both the original speech and the noisy speech. There is an interesting difference between the MOS results and the SNR results. The noise estimation algorithm received higher scores than would have been suggested by SNR, especially for the Babble noise, its score is little higher than the proposed method. The most likely explanation for this phenomenon is that the presence or absence of the residual musical noise and speech distortion in the enhanced speech, which are known to have significant impact on perception but have only mild influence on SNR values. It also can be seen from the speech spectrograms (Figure 3) that the spectrogram of the noise-estimation algorithm is more “smooth” than that from the proposed algorithm, which means less residual musical noise and speech distortion. Actually, the noise-estimation algorithm may suit for highly non-stationary noise environments, and the proposed algorithm may suit for colored noise environments.

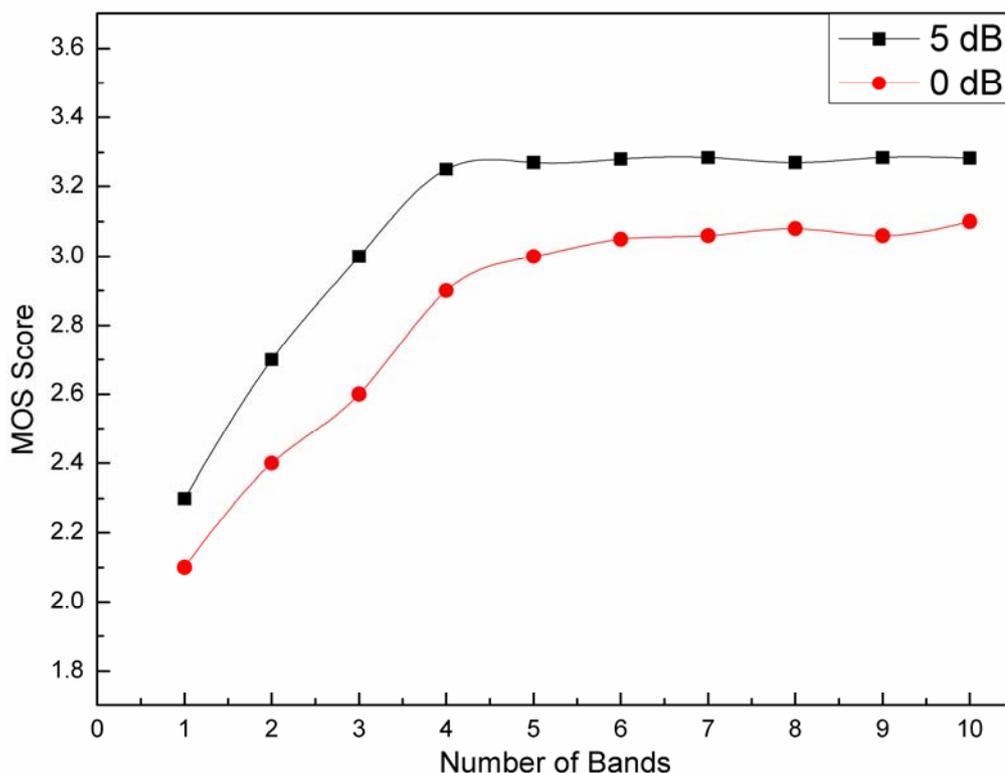
The results of the perceptual experiment also suggested that the enhanced speech with the proposed multi-band spectral subtraction algorithm is more pleasant, the residual noise is better reduced, and with minimal, if any, speech distortion. This is because that the over-subtraction factor of each frequency band can be adjusted, which can prevent speech from quality deterioration during spectral subtraction process.

Figure 5. Acceptability scores of the original and enhanced EL speech. The noisy speech in the case of additive noise has an input SNR of 0 dB.



It can be seen from theory of the multi-band algorithm that the number of bands may have important effects on the quality of enhanced speech. Therefore, we varied the number of bands from 1 to 10 in this study to determine the optimal number of bands. Linear frequency spacing was used, and the performed speech quality was examined using Subjective measures (MOS). Figure 6 plots the averaged MOS scores for five enhanced sentences (produced by two male laryngectomees) for different number of bands. It can be seen from the figure that the MOS score has a great improvement when the number of bands increased from 1 to 5 for both 5 and 0 dB. For the bands larger than 5, the score does show a slight increase in performance, however, it is too slight to be perceivable in speech quality. Therefore, the optimal number of bands can be determined as 5 for EL speech. This number is almost consistent with the result of Kamath *et al.*, which for normal speech was 4. In addition, when the total number of bands is one, then the approach of multi-band spectral subtraction algorithm reduces to the traditional power spectral subtraction approach[20].

Figure 6. Acceptability scores of the performance of the multi-band spectral subtraction approach for different number of bands for the enhanced speech. The noisy speech in the case of additive White noise has an input SNR of 5 and 0 dB.



Because the subtraction parameters are fixed for a given frame, the traditional spectral subtraction algorithm cannot reduce the noise effectively, especially for the colored noise. These limitations will be worse for the enhancement of EL speech in the case of combined electronic noise. With regard to the multi-band spectral subtraction algorithm, the over-subtraction factor of each frequency band can be adjusted, so that this algorithm can realize a good tradeoff between reducing noise, increasing intelligibility, and keeping the distortion acceptable to a human listener. The results also indicate that

the proposed algorithm can not only reduce the residual noise but also improve the low-frequency deficit of EL speech.

The proposed algorithm also has strong flexibility to adapt complicated speech environment for EL device users, because the over-subtraction factor of each frequency band can be adjusted, so that the proposed algorithm is able to fit other different or complex speech environment. This makes it possible to obtain better speech quality via speech enhancement under some rigorous speech environment.

As a single channel subtractive-type speech enhanced methods, the proposed algorithm in this paper can be applied into the enhancement of EL speech in a practical electronic situation. For example, an EL speech enhanced system embedded with this algorithm can be developed. With the help of digital signal processing (DSP) technology, the speech enhancement function can be realized with a microprocessor and implanted into an EL-telephone, EL-microphone, or other electronic media. Different enhancement algorithms can be selected through the switch based on different noisy conditions. Along with the development of efficient enhancement methods, the quality of EL speech will be extensively improved for better perception.

4. Conclusions

In order to remove the radiated noise (during phonation) and the other additive environment noise from the EL speech, an improved spectral subtraction method, multi-band spectral subtraction algorithm is investigated in this study in order to takes into account the non-uniform effect of colored noise on the spectrum of EL speech. Because the over-subtraction factor of each frequency band can be adjusted in the enhancement process, both the objective and subjective test results suggest that a better noise reduction effect was obtained and the perceptually annoying musical noise was efficiently reduced (especially in the high-frequency regions), with little distortion to speech information as compared to the other standard speech enhancement algorithm. Furthermore, the proposed algorithm has strong flexibility to adapt complicated and rigorous speech environment by adjusted the over-subtraction factor of each frequency band.

Future work on this approach will include the development of time-varying over-subtraction factors for each frequency band based on tracking the characteristics of the input EL signal, such as those used in some wavelet denoising techniques[24]. Furthermore, this algorithm is not consider the fact that the human auditory system has different sensitivities for different frequencies, therefore, the perceptual weighting technique[25] should also be added into the algorithm so that take into account the frequency-domain masking properties of the human auditory system. Continued work in this direction is expected to lead to additional improvement in overall EL speech enhancement.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (NSFC, No. 30770544, and No. 10874137), and the National postdoctoral Science Foundation of China (No. 20070411131). We also want to thank the participants from the E.N.T. Department, the Second Hospital, Xi'an Jiaotong University, for helping with data acquisition and analysis.

References and Notes

1. Ooe, K.; Fukuda, T.; Arai, F. A new type of artificial larynx using a PZT ceramics vibrator as a sound source. *IEEE/ASME Transactions on Mechatronics* **2000**, *5*, 221 - 225.
2. Lowry, L. D. A review: Development of a prototype self-contained intra-oral artificial larynx. *Laryngoscope* **1981**, *91*, 1332-1355.
3. Li, S.; Scherer, R. C.; Wan, M.; Wang, S.; Wu, H. The effect of glottal angle on intraglottal pressure. *Journal of the Acoustical Society of America* **2006**, *119*, 539-548.
4. Li, S.; Scherer, R. C.; Wan, M.; Wang, S.; Wu, H. Numerical study of the effects of inferior and superior vocal fold surface angles on vocal fold pressure distributions. *Journal of the Acoustical Society of America* **2006**, *119*, 3003-3010.
5. Barney, H. L.; Haworth, F. E.; Dunn, H. K. An experimental transistorized artificial larynx. *Bell System Technical Journal* **1959**, *38*, 1337-1356.
6. Weiss, M. S.; Yeni-Komshian, G. H.; Heinz, J. M. Acoustic and perceptual characteristics of speech produced with an electronic artificial larynx. *Journal of the Acoustical Society of America* **1979**, *65*, 1298-1308.
7. Knox, A. A.; Anneberg, M. The effects of training in comprehension of electrolaryngeal speech. *Journal of Communication Disorders* **1973**, *6*, 110-120.
8. Norton, R. L.; Bernstein, R. S. Improved Laboratory Prototype electrolarynx (LAPEL): using inverse filtering of frequency response function of the human throat. *Annals of Biomedical Engineering* **1993**, *21*, 163-174.
9. Boll, S. F. Suppression of acoustic noise in speech using spectral subtraction. *IEEE Transactions on Acoustics, Speech and Signal Processing* **1979**, *27*, 113-120.
10. Espy-Wilson, C. Y.; Chari, V. R.; MacAuslan, J.; Walsh, M. Enhancement of electrolaryngeal speech by adaptive filtering. *Journal of Speech, Language, and Hearing Research* **1998**, *41*, 1253-1264.
11. Liu, H.; Zhao, Q.; Wan, M.; Wang, S. Application of spectral subtraction method on enhancement of electrolarynx speech. *Journal of the Acoustical Society of America* **2006**, *120*, 398-406.
12. Liu, H.; Zhao, Q.; Wan, M.; Wang, S. Enhancement of electrolarynx speech based on auditory masking. *IEEE Transactions on Biomedical Engineering* **2006**, *53*, 865 - 874.
13. Pandey, P. C.; Bhandarkar, S. M.; Bachher, G. K.; Lehana, P. K. Enhancement of alaryngeal speech using spectral subtraction. *Digital Signal Processing* **2002**, *2*, 591-594.
14. Lockwood, P.; Boudy, J. Experiments with a nonlinear spectral subtractor (NSS), hidden Markov models and projection, for robust recognition in cars. *Speech Communication* **1992**, *11*, 215-228.
15. Hansen, J. H. L. Morphological constrained feature enhancement with adaptive cepstral compensation (MCE-ACC) for speech recognition in noise and Lombard effect. *IEEE Transactions on Speech and Audio Processing* **1994**, *2*, 598-614.
16. Kamath, S.; Loizou, P. A multi-band spectral subtraction method for enhancing speech corrupted by colored noise. *IEEE International Conference on Acoustics, Speech, and Signal Processing* **2002**, *4*, 4160-4164.

17. Udrea, R. M.; Ciochina, S.; Vizireanu, D. N. Multi-band Bark Scale Spectral over-subtraction for Colored Noise Reduction. *International Symposium on Signals, Circuits and Systems* **2005**, *1*, 311-314.
18. Niu, H. J.; Wan, M. X.; Wang, S. P.; Liu, H. J. Enhancement of electrolarynx speech using adaptive noise cancelling based on independent component analysis. *Medical and Biological Engineering and Computing* **2003**, *41*, 670-678.
19. Holly, S. C.; Lernman, J.; Randolph, K. A comparison of the intelligibility of esophageal, electrolarynx, and normal speech in quiet and in noise. *Journal of Communication Disorders* **1983**, *16*, 143-155.
20. Berouti, M.; Schwartz, R.; Makhoul, J. Enhancement of speech corrupted by acoustic noise. *IEEE International Conference on Acoustics, Speech, and Signal Processing* **1979**, 208-211.
21. Cohen, I.; Berdugo, B. Noise Estimation by Minima Controlled Recursive Averaging for Robust Speech Enhancement. *IEEE Signal Processing Letters* **2002**, *9*, 12-15.
22. Cohen, I.; Berdugo, B. Speech enhancement for non-stationary noise environments. *Signal Processing* **2001**, *81*, 2403-2418.
23. Rangachari, S.; Loizou, P. C. A noise-estimation algorithm for highly non-stationary environments. *Speech Communication* **2006**, *48*, 220-231.
24. Cohen, I. Enhancement of speech using bark-scaled wavelet packet decomposition. *7th European Conference on Speech Communication and Technology* **2001**, 1933-1936.
25. Hu, Y.; Loizou, P. C. A perceptually motivated approach for speech enhancement. *IEEE Transactions on Speech and Audio Processing* **2003**, *11*, 457-464.

© 2009 by the authors; licensee Molecular Diversity Preservation International, Basel, Switzerland. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/3.0/>).