

Title	Psychoacoustically-motivated adaptive $n$ -order generalized spectral subtraction based on data-driven optimization
Author(s)	Li, Junfeng; Jiang, Hui; Akagi, Masato
Citation	Proceedings of INTERSPEECH 2008: 171-174
Issue Date	2008-09-23
Type	Conference Paper
Text version	publisher
URL	<a href="http://hdl.handle.net/10119/9984">http://hdl.handle.net/10119/9984</a>
Rights	Copyright (C) 2008 International Speech Communication Association. Junfeng Li, Hui Jiang, Masato Akagi, Proceedings of INTERSPEECH 2008, pp.171-174.
Description	

# Psychoacoustically-motivated Adaptive $\beta$ -order Generalized Spectral Subtraction Based on Data-driven Optimization

Junfeng Li<sup>1</sup>, Hui Jiang<sup>2</sup>, Masato Akagi<sup>1</sup>

<sup>1</sup> School of Information Science, Japan Advanced Institute of Science and Technology, Japan

<sup>2</sup> Department of Computer Science and Engineering, York University, Canada

Email: junfeng@jaist.ac.jp, hj@cse.yorku.ca, akagi@jaist.ac.jp

## Abstract

To mitigate the performance limitations caused by the constant spectral order  $\beta$  in the traditional spectral subtraction methods, we previously presented an adaptive  $\beta$ -order *generalized spectral subtraction* (GSS) in which the spectral order  $\beta$  is updated in a heuristic way [10]. In this paper, we propose a psychoacoustically-motivated adaptive  $\beta$ -order GSS, by considering that different frequency bands contribute different amounts to speech intelligibility (i.e., the band-importance function). Specifically, in this proposed adaptive  $\beta$ -order GSS, the tendency of spectral order  $\beta$  to change with the input local *signal-to-noise ratio* (SNR) is quantitatively approximated by a sigmoid function, which is derived through a data-driven optimization procedure by minimizing the intelligibility-weighted distance between the desired speech spectrum and its estimate. The inherent parameters of the sigmoid function are further optimized with the data-driven optimization procedure. Experimental results indicate that the proposed psychoacoustically-motivated adaptive  $\beta$ -order GSS yields great improvements over the traditional spectral subtraction methods with the intelligibility-weighted measures.

**Index Terms:** Psychoacoustically-motivated adaptive  $\beta$ -order GSS, Data-driven optimization, Importance function, Speech intelligibility, Sigmoid function.

## 1. Introduction

Background noises severely degrade performance of state-of-the-art speech applications, for example, in the reduced intelligibility of perceived speech for cochlear implant users. As a pre-processor, noise reduction has been shown to be effective in improving the quality and/or intelligibility of noisy signal, and feeding the enhanced signal to implant users results in significant benefits in sentence/word recognition [1].

Research studies on speech intelligibility show that various factors affect speech intelligibility to different degrees [2]. These factors include the audibility in each subband derived from the corresponding subband SNR and the listener's hearing threshold, and the band-importance functions that indicate to which degree each frequency band contributes to intelligibility [3, 4]. Speech intelligibility was traditionally computed from the long-term speech and noise spectra in stationary noise [2, 3]. This calculation approach was recently extended for non-stationary noise conditions by first calculating the speech intelligibility in each short-term frame followed by averaging these intelligibility values across frames, yielding the overall intelligibility for that particular condition [4].

To improve the speech intelligibility of the received signal in noisy conditions, a number of noise reduction algo-

rithms have been proposed in the past several decades [5]. As a well-known noise reduction method, *spectral subtraction* (SS) has been used widely due to its simplicity in implementation and improved to overcome its shortcomings in different ways [5, 6, 7, 8]. Boll applied several secondary procedures to the processed signal after SS to further attenuate the residual noise [5]. To enhance noise reduction and mitigate “musical” noise, Berouti *et al.* introduced two additional parameters, an oversubtraction factor that controls the amount of noise to be subtracted, and a spectral flooring factor that mitigates “musical” noise [6]. Schless *et al.* suggested to set both oversubtraction factor and spectral flooring factor based on the *signal-to-noise ratio* (SNR) [7]. Kamath *et al.* proposed to empirically set different oversubtraction factors in different subbands [8]. Sim *et al.* derived a short-time spectral amplitude estimator of the speech signal based on a parametric formulation of *generalized spectral subtraction* (GSS) [9]. In all SS methods mentioned above, however, the spectral order  $\beta$  is always fixed to some constants which result in performance limitation to a certain degree. More recently, therefore, Li *et al.* suggested an adaptive  $\beta$ -order GSS in which the spectral order  $\beta$  is updated in each subband according to the input local SNRs in a heuristic way [10].

In this paper, we propose a psychoacoustically-motivated adaptive  $\beta$ -order GSS based on the research results on speech intelligibility and a data-driven optimization procedure. Considering the different contributions of different frequency bands to speech intelligibility which is quantified by the band-importance function [2, 3], we propose to quantitatively determine the tendency of spectral order  $\beta$  to change with the input local SNR in MMSE sense, and introduce a data-driven optimization procedure that quantifies this change tendency as a sigmoid function and optimizes the inherent parameters. Experimental results in various noise conditions show that the proposed method outperforms the traditional SS algorithms in terms of the intelligibility-weighted measures.

## 2. $\beta$ -order generalized spectral subtraction

### 2.1. Signal model

Suppose the observed noisy signal to be the sum of the clean speech signal and the uncorrelated additive noise signal. Applying the *short-time Fourier transform* (STFT), the observed signal in the time-frequency domain is represented as

$$X(k, \ell) = S(k, \ell) + N(k, \ell), \quad (1)$$

where  $k$  and  $\ell$  are the frequency bin index and the time frame index, respectively;  $X(k, \ell)$ ,  $S(k, \ell)$  and  $N(k, \ell)$  are the STFTs of the noisy signal, the clean signal and the noise signal.

## 2.2. $\beta$ -order generalized spectral subtraction

The  $\beta$ -order generalized spectral subtraction is defined as [9]

$$|\hat{S}_\beta(k, \ell)|^\beta = a_\beta(k, \ell) |X(k, \ell)|^\beta - b_\beta(k, \ell) E \left[ |N(k, \ell)|^\beta \right], \quad (2)$$

where  $\beta$  denotes the spectral order,  $\hat{S}_\beta(k, \ell)$  is the spectral estimate of the speech enhanced by the  $\beta$ -order GSS, and  $a_\beta(k, \ell)$  and  $b_\beta(k, \ell)$  are two parameters. Note that the speech spectrum estimate  $\hat{S}_\beta(k, \ell)$  is dependent not only on time and frequency, but also on the spectral order  $\beta$ .

The parameters  $a_\beta(k, \ell)$  and  $b_\beta(k, \ell)$  in the  $\beta$ -order GSS are determined and optimized by minimizing the MSE of the  $\beta$ -order speech spectrum amplitude  $|S_\beta(k, \ell)|^\beta$  and its estimate  $|\hat{S}_\beta(k, \ell)|^\beta$ . Under the complex Gaussian assumption and substituting the two newly-derived optimal coefficients into Eq. (2), the gain function of the  $\beta$ -order GSS is derived as [9]

$$\hat{G}_\beta(k, \ell) = \left\{ \frac{[\xi_\beta(k, \ell)]^\beta}{1 + [\xi_\beta(k, \ell)]^\beta} \right\}^{\frac{1}{\beta}} \left\{ 1 - \left( 1 - [\xi_\beta(k, \ell)]^{\frac{-\beta}{2}} \right) \Gamma\left(\frac{\beta}{2} + 1\right) \left( \frac{1}{\gamma(k, \ell)} \right)^{\frac{\beta}{2}} \right\}^{\frac{1}{\beta}}, \quad (3)$$

where  $\xi_\beta(k, \ell)$  and  $\gamma(k, \ell)$  are the *a priori* SNR and the *a posteriori* SNR as defined in [11], and  $\Gamma(\cdot)$  denotes the Gamma function. The estimate of  $\xi_\beta(k, \ell)$  is updated in a decision-directed scheme, greatly decreasing the residual “musical” noise [11].

## 3. Psychoacoustically-motivated adaptive $\beta$ -order GSS using data-driven optimization

Considering the research findings on speech intelligibility, in this section, we propose a psychoacoustically-motivated adaptive  $\beta$ -order GSS in which the dependence of  $\beta$  on the local SNR is derived and optimized through a data-driven optimization procedure.

### 3.1. Subband-processing-based determination of spectral order $\beta$

The value of spectral order  $\beta$  has been shown to be largely dependent on the SNR in the current condition [10]. Furthermore, the SNRs vary greatly with time due to the time-varying characteristics of speech and noise signals, and also significantly vary in different subbands because of the colorfulness of noise signals and the non-uniform spectral energy distribution of speech signals. As a result, speech signal corrupted by real-world noises is characterized by different local SNRs in different partitions in the time-frequency domain. Further taking the strong correlations of spectral components between adjacent frequency bins into account, the appropriate value of  $\beta$  should be determined depending not only on knowledge of the current frequency bin under consideration, but also on knowledge of the neighboring bins. As a result, the appropriate value of spectral order  $\beta$  must be adaptively determined according to the local SNRs calculated in subbands instead of the instantaneous SNR in each individual frequency bin in the time-frequency domain.

### 3.2. Psychoacoustically-motivated derivation of spectral order $\beta$

Improving speech intelligibility motivates us to derive the appropriate value for spectral order  $\beta$  by integrating the psychoacoustical research findings on speech intelligibility. There are

a number of aspects which play an important role in speech intelligibility enhancement. For example, it is understood that the short-term spectrum is of primary importance in the perception of speech, and different frequency bands contribute different amounts to speech intelligibility [2, 3]. Speech intelligibility index (SII) in ANSI S.35-1997 standard quantifies the degree of speech intelligibility in the presence of background noise [3]. Audibility is measured as the ratio of time-averaged speech power and time-averaged noise power in a set of frequency bands, followed by weighting with the so-called band-importance function. The band-importance function indicates to which degree each frequency band contributes to intelligibility. Finally, the SII is determined by accumulation of the audibility across the different frequency bands. As a result, the band-importance functions should be integrated when determining the appropriate value of spectral order  $\beta$ , as shown in the following section.

### 3.3. Data-driven optimization of spectral order $\beta$

In our previous study, the change tendency of the spectral order  $\beta$  along with the input SNR was qualitatively analyzed and intuitively described by a sigmoid function, where no theoretical/experimental evidences were provided [10]. With the consideration of the non-uniform effect of noise on speech in Section 3.1 and the psychoacoustical findings on speech intelligibility in Section 3.2, in this section, we aim to quantify the dependency of the spectral order  $\beta$  on the input local SNR using a data-driven optimization approach.

In our data-driven optimization procedure, ten speech sentences were randomly selected from the NTT database [12], and two noise signals (“car” and “babble”) were taken from the NOISEX-92 database [13]. The speech and noise signals were first downsampled to 8kHz and then mixed with the global SNR ranging from  $-40$  to  $40$  dB. We assume that the noise spectrum is a known prior in this optimization procedure. For a given value of  $\beta$ , the gain function of the  $\beta$ -order GSS is calculated using Eq. (3), and then used to enhance the target speech signal. Furthermore, based on the discussions in Sections 3.1 and 3.2, we propose to optimize the spectral order  $\beta$  by minimizing the overall intelligibility-weighted distance between the spectral amplitude  $|S(k, \ell)|$  of the clean signal and that of its estimate  $|\hat{S}_\beta(k, \ell)|$  summed across all subbands, that is,

$$\beta^{\text{opt}} = \arg \min_{0.1 \leq \beta \leq 3.0} \left( \sum_{m=1}^M \sum_{k=\omega_m}^{\omega_{m+1}} I_m \left[ |S(k, \ell)| - |\hat{S}_\beta(k, \ell)| \right]^2 \right), \quad (4)$$

where  $I_m$  is the band-importance function in the  $m$ -th subband [3],  $M$  is the number of subbands,  $\omega_m$  denotes the boundary frequency of the  $m$ -th subband, and the range of  $\beta$  is empirically confined to  $[0.1, 3.0]$ . Note that the importance functions are normally given as constant values [3] and each term in the overall intelligibility-weighted spectral amplitude distance in Eq. (4) are positive, therefore, the minimization problem in Eq. (4) can be simplified to the minimization problem performed in each independent subband. That is, the minimization of the overall intelligibility-weighted spectral amplitude distance can be achieved by the value that yields the minimal intelligibility-weighted spectral amplitude distance of the clean signal and its estimate in each individual subband, given by

$$\beta_m^{\text{opt}} = \arg \min_{0.1 \leq \beta \leq 3.0} \left( \sum_{k=\omega_m}^{\omega_{m+1}} \left[ |S(k, \ell)| - |\hat{S}_\beta(k, \ell)| \right]^2 \right). \quad (5)$$

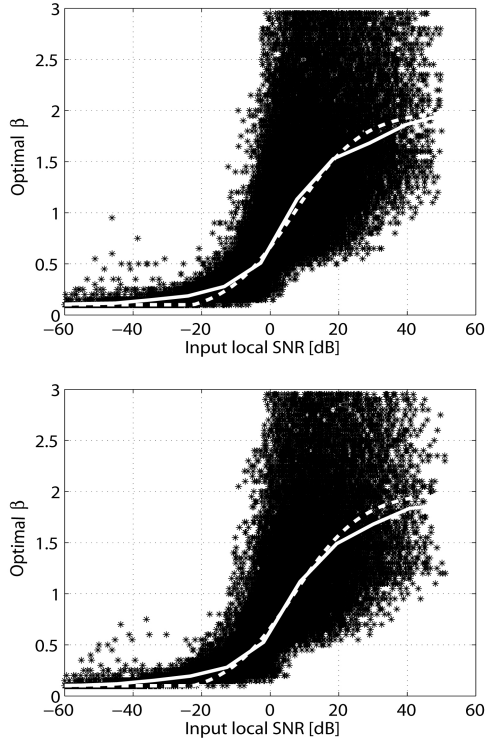


Figure 1: Scatter plots of the optimized  $\beta$  value with respect to the input local SNR, the mean of the scattered data (solid line) and the fitted sigmoid function (dashed line) with the parameters  $A = 0.1$ ,  $B = 2.0$ ,  $D = 7.0$ , in the car noise condition (upper panel) and in the babble noise condition (lower panel).

Though only ten speech sentences are used in this optimization procedure, we should note that the optimization is performed in the following scenarios: in each frames (each speech sentence is divided into 220 ~ 350 overlapping frames by windowing before Fourier transform.), in each critical subband (e.g., 18 subbands) and at the different global SNR conditions (that is, -40 to 40 dB in 10 dB increments). It is therefore believed that the optimization of spectral order  $\beta$  is sufficient in the statistical sense, and the parameters obtained from this optimization procedure might also be applicable in other different conditions.

Fig. 1 shows the scatter plots of the optimized  $\beta$  value against the local SNR (defined in Eq. (6)), the mean curve as well as the fitted sigmoid function in the “car” and “babble” noise conditions. The results shown in Fig. 1 indicate that: (1) The optimal  $\beta$  value increases (decreases) as the local input SNR increases (decreases), which proves our previous qualitative analysis results [10]; (2) Most importantly, the change tendency of the appropriate value of  $\beta$  with the change of the local SNR can be approximated by a sigmoid function (defined in Eq. (7)); (3) Moreover, the inherent parameters are also determined through this data-driven optimization.

### 3.4. Adaptive scheme for an appropriate value of spectral order $\beta$

As stated in Sections 3.1 and 3.2, the appropriate value of spectral order  $\beta$  vary with the change of the local SNR in each subband. Further considering the mechanism of human perception, the whole spectrum is first divided into subbands according to

the critical-band scale [14]. Then, the local SNR  $\rho(m, \ell)$  in the  $m$ -th critical band and the  $\ell$ -th frame is calculated as

$$\rho(m, \ell) = 10 \log_{10} \left( \frac{\sum_{k=\omega_m}^{\omega_{m+1}} |X(k, \ell)| - |\hat{N}(k, \ell)|^2}{\sum_{k=\omega_m}^{\omega_{m+1}} |\hat{N}(k, \ell)|^2} \right), \quad (6)$$

where  $\omega_m$  is specified as the boundary frequency of the  $m$ -th critical band.

Based on the data-driven optimization results in Section 3.3, we propose to determine the optimized value of the spectral order  $\tilde{\beta}(m, \ell)$  according to the local SNR  $\rho(m, \ell)$  in each critical band, frame by frame, by the use of the sigmoid function, given by

$$\tilde{\beta}(m, \ell) = \frac{B}{1 + e^{-A[\rho(m, \ell) - D]}}, \quad (7)$$

where the parameter  $A$  controls the changing speed of the value of  $\tilde{\beta}(m, \ell)$  with respect to the local SNR  $\rho(m, \ell)$ ,  $B$  determines the range of the value of  $\tilde{\beta}$ , and  $D$  denotes the shift along the SNR axis. In order to avoid severe speech distortion caused by an extremely low  $\beta$  value, we further confine the value of  $\tilde{\beta}(k, \ell)$  with a minimum value  $\beta_{min}$ . As a result, the optimized value of the spectral order  $\hat{\beta}(m, \ell)$  is finally determined as

$$\hat{\beta}(m, \ell) = \max \left[ \tilde{\beta}(m, \ell), \beta_{min} \right]. \quad (8)$$

## 4. Experiments and results

### 4.1. Experimental configuration

Performance of the proposed psychoacoustically-motivated adaptive  $\beta$ -order GSS was assessed by the following experiments. We randomly selected 40 clean continuous speech sentences produced by two females and two males from NTT speech database [12]. Two types of noise source, “car” and “babble”, were chosen from NOISEX-92 database [13]. The clean speech and noise signals were first downsampled to 8kHz. We generated noisy speech signals artificially by adding various noise signals to the clean signals at different SNRs ranging from 0 to 15 dB in 5 dB increments. Note that the car noise was a stationary signal, whereas the babble noise was a non-stationary signal. Based on the optimization results shown in Fig. 1, the parameters in our experiments were set as follows:  $A = 0.1$ ,  $B = 2.0$ ,  $D = 7.0$  and  $\beta_{min} = 0.1$ . Note that the spectral order  $\beta$  in the current experiments ranges from 0.1 to 2.0. Performance of the proposed adaptive  $\beta$ -order GSS was further compared with that of the traditional SS methods (power SS, amplitude SS and SS when  $\beta = 0.1$ ) that were implemented by setting the spectral order  $\beta$  to 2.0, 1.0 and 0.1 in Eq. (3) derived in [9], respectively.

### 4.2. Experimental results

Performance of the studied algorithms was objectively examined in terms of *intelligibility-weighted SNR* ( $\text{SNR}_{\text{int}}$ ) and *intelligibility-weighted log-spectral distance* ( $\text{LSD}_{\text{int}}$ ), as defined in [15]. The experimental results of  $\text{SNR}_{\text{int}}$  and  $\text{LSD}_{\text{int}}$ , averaged across all sentences in two noise conditions, are shown in Tables 1 and 2, respectively.

Table 1 illustrates that the proposed psychoacoustically-motivated adaptive  $\beta$ -order GSS consistently provides the highest  $\text{SNR}_{\text{int}}$  improvements, compared to the traditional SS methods, for all conditions at all SNRs. While the SS method with

Table 1: Intelligibility-weighted SNR [dB] of the noisy signal, the power SS ( $\beta = 2.0$ ) output, the amplitude SS ( $\beta = 1.0$ ) output, the SS output when  $\beta = 0.1$ , and the proposed adaptive  $\beta$ -order GSS output.

Global SNR	0	5	10	15
Condition	car noise			
Noisy	-11.51	-6.26	-1.15	3.85
$\beta = 2.0$ (Power SS)	-6.02	-2.35	1.06	4.02
$\beta = 1.0$ (Amplitude SS)	-3.56	-0.54	2.50	5.32
$\beta = 0.1$	-5.33	-5.54	-5.35	-5.33
<b>Adaptive-<math>\beta</math></b>	<b>-2.12</b>	<b>0.89</b>	<b>3.76</b>	<b>6.04</b>
Condition	babble noise			
Noisy	-13.96	-8.61	-3.41	1.65
$\beta = 2.0$ (Power SS)	-9.54	-5.31	-1.42	1.95
$\beta = 1.0$ (Amplitude SS)	-6.56	-3.05	0.28	3.32
$\beta = 0.1$	-5.22	-5.28	-5.31	-5.32
<b>Adaptive-<math>\beta</math></b>	<b>-4.38</b>	<b>-1.33</b>	<b>1.58</b>	<b>4.22</b>

Table 2: Intelligibility-weighted LSD [dB] of the noisy signal, the power SS ( $\beta = 2.0$ ) output, the amplitude SS ( $\beta = 1.0$ ) output, the SS output when  $\beta = 0.1$ , and the proposed adaptive  $\beta$ -order GSS output.

Global SNR	0	5	10	15
Condition	car noise			
Noisy	18.93	15.96	13.43	11.30
$\beta = 2.0$ (Power SS)	15.29	13.22	11.46	10.00
$\beta = 1.0$ (Amplitude SS)	13.97	12.15	10.57	9.29
$\beta = 0.1$	52.78	54.33	55.24	55.69
<b>Adaptive-<math>\beta</math></b>	<b>13.69</b>	<b>11.63</b>	<b>10.05</b>	<b>8.97</b>
Condition	babble noise			
Noisy	20.47	17.25	14.51	12.21
$\beta = 2.0$ (Power SS)	17.41	14.89	12.76	11.04
$\beta = 1.0$ (Amplitude SS)	15.75	13.62	11.79	10.27
$\beta = 0.1$	50.37	52.43	53.84	54.69
<b>Adaptive-<math>\beta</math></b>	<b>14.78</b>	<b>12.82</b>	<b>11.13</b>	<b>9.81</b>

$\beta = 0.1$  yields greatly reduced  $\text{SNR}_{\text{int}}$  results because it introduces severe speech distortion due to the too small value of  $\beta$  (i.e., 0.1). The highest  $\text{SNR}_{\text{int}}$  by our proposed algorithm indicates high noise reduction ability corresponding to high speech intelligibility in some sense [15]. This might be attributed to the use of low gains in speech-absence periods due to the low values of the spectral order  $\beta$ .

Concerning the results of  $\text{LSD}_{\text{int}}$  shown in Table 2, we can observe that all tested algorithms decrease the  $\text{LSD}_{\text{int}}$  in all conditions, except for the SS algorithm with  $\beta = 0.1$  that markedly increases  $\text{LSD}_{\text{int}}$  (i.e., high speech distortion and low intelligibility). The enhanced speech signal processed by the proposed algorithm involves the lowest speech distortion, which corresponds to high speech intelligibility to a certain degree. This achievement can specifically be attributed to the use of high gains in speech-presence periods, due to the high values of the spectral order  $\beta$ .

As a result, the proposed psychoacoustically-motivated adaptive  $\beta$ -order GSS outperforms the traditional SS algorithms in speech intelligibility enhancement, due to the time-varying frequency-dependent spectral order  $\beta$  that is derived and optimized through a data-driven optimization procedure based on the consideration of psychoacoustic research findings on speech intelligibility (e.g., the band-importance function).

## 5. Conclusion

In this paper, we proposed a psychoacoustically-motivated  $\beta$ -order generalized spectral subtraction for speech intelligibility enhancement, in which the dependence of the spectral order  $\beta$  on the local SNRs is derived and optimized through a data-driven optimization procedure based on consideration of the different contributions of different frequency bands to speech intelligibility. Experimental results illustrate that the proposed algorithm outperforms the traditional SS algorithms in various noise conditions in terms of intelligibility-weighted measures. As a future work, the effectiveness of our proposed algorithm in improving speech intelligibility for cochlear implant users (using state-of-the-art cochlear implants) should be assessed in daily-life noise conditions.

## 6. Acknowledgement

This study was supported by a Grant-in-Aid for Young Scientists (B) (No. 19700156) from the Ministry of Education, Science, Sports and Culture of Japan.

## 7. References

- [1] P.C. Loizou, "Introduction to cochlear implants," *IEEE Engineering in Medicine and Biology Magazine*, vol. 18, pp. 32-42, 1999.
- [2] C.V. Pavlovic, "Derivation of primary parameters and procedures for use in speech intelligibility prediction," *J. Acoust. Soc. Am.*, vol. 82, no. 2, pp. 413-422, 1987.
- [3] ANSI S3.5-1997, "American National Standard Methods for Calculation of the Speech Intelligibility Index," 1997.
- [4] K.S. Rhebergen and N.J. Versfeld, "A speech intelligibility index-based approach to predict the speech reception threshold for sentences in fluctuating noise for normal hearing listeners," *J. Acoust. Soc. Am.* vol. 117, no. 4, pp. 2181-2191, 2005.
- [5] S. F. Boll, "Suppression of acoustic noise in speech using spectral subtraction," *IEEE Trans. ASSP*, vol. ASSP-27, no. 2, pp. 113-120, 1979.
- [6] M. Berouti, R. Schwartz, and J. Makhoul, "Enhancement of speech corrupted by acoustic noise," In *Proc. ICASSP*, pp. 208-211, 1979.
- [7] V. Schless and F. Class, "SNR-dependent flooring and noise over-estimation for joint application of spectral subtraction and model combination," In *Proc. ICSLP*, pp. 721-725, 1998.
- [8] S.D. Kamath and P.C. Loizou, "A multi-band spectral subtraction method for enhancing speech corrupted by colored noise," In *Proc. ICASSP*, pp. 4164-4167, 2002.
- [9] B.L. Sim, *et al.* "A parametric formulation of the generalized spectral subtraction," *IEEE Trans. SAP*, vol. 6, no. 4, pp. 328-337, 1998.
- [10] J. Li, *et al.*, "Noise reduction based on adaptive  $\beta$ -order generalized spectral subtraction for speech enhancement," In *Proc. Interspeech2007*, pp. 802-805, 2007.
- [11] Y. Ephraim and D. Malah, "Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator," *IEEE Trans. ASSP*, vol. 32, no. 6, pp. 1109-1121, 1984.
- [12] [http://www.ntt-at.com/products\\_e/speech2002/](http://www.ntt-at.com/products_e/speech2002/).
- [13] A. Varga and H.J.M. Steeneken, "Assessment for automatic speech recognition: II. NOISEX-92: A database and an experiment to study the effect of additive noise on speech recognition systems," *Speech Communication*, vol. 12, pp. 247-251, 1993.
- [14] E. Zwicker and E. Terhardt, "Analytical expressions for critical band rate and critical bandwidth as a function of frequency," *J. Acoust. Soc. Am.*, vol. 68, pp. 1523-1525, 1980.
- [15] J.E. Greenberg, P.M. Peterson and P.M. Zurek, "Intelligibility-weighted measures of speech-to-interference ratio and speech system performance," *J. Acoust. Soc. Am.*, vol. 94, no. 5, pp. 3009-3010, 1993.