ITRW on Statistical and
Perceptual Audio Processing
ICC Jeju, Korea
October 3, 2004

MAP Estimation of Speech Spectral Component under GGD a Priori

Rajkishore Prasad, Hiroshi Saruwatari, Kiyohiro Shikano

Graduate School of Information Science,
Nara Institute of Science and Technology, Nara, Japan
{kishor-p, sawatari, shikano}@is.aist-nara.ac.jp

Abstract

This paper presents *Maximum A Posteriori* (MAP) estimation of the spectral components of clean speech from the observed data noised by the additive background noise having Gaussian or non-Gaussian statistical distribution. In the proposed algorithm MAP estimator for the spectral components of clean signal is derived using Generalized Gaussian Distribution (GGD) function as *a priori* statistical models for the spectral components of speech as well as noise. Since the spikiness of the GGD can be controlled by the shape parameter, it is possible to model Gaussian as well as non-Gaussian noise, corrupting the speech signal. The enhancement results for the speech signal corrupted by the Gaussian noise and non-Gaussian noise are presented to show the usefulness of the estimator. Denoising performance for the Laplacian noise and white Gaussian noise have also been compared with that of the conventional Wiener filtering, which assumes Gaussian distributions for both the speech and noise.

1. Introduction

Different algorithms have been proposed so far for the estimation of spectral components of clean speech in the Discrete Fourier Transform (DFT) domain from its noisy version under the additive background noise having Gaussian PDF [1]. In such algorithms magnitudes of speech spectral components of the clean signal are estimated from the same of the noisy speech signal with the help of some noise suppression rules and phase is taken same as that of the noisy spectral components [2]. Such phase restoration process is based on the fact that ear is insensitive to the phase, however, this is false for the time-variant signals. Speech enhancement has been an active area of research since long, and thus many noise suppression rules, using different heuristic ideas and mathematical formulations, have been proposed so far [3]. The first work reported in this direction appeared in [4] then there came Boll's spectral subtraction algorithm [5], which in turn followed by its different variants [1]. In the spectral subtraction, either the root mean square of estimated noise is subtracted from or power of the noisy spectral component is modified in each frequency bins while keeping phase unaltered. There have also been developments of methods, e.g. Wiener filtering, Short-Time Spectral Amplitude (STSA) estimator by Ephairim and Malah [6], based on statistical estimation techniques such as Maximum Likelihood (ML) and MAP estimation using statistical models for the spectral components. In most of the statistical estimation algorithms Gaussian Distribution (GD) function has been used to model the spectral components of the speech and noise signal under the advocacy of the Central Limit Theorem (CLT) as each Discrete Fourier Transform (DFT) coefficients is weighted sum of the random data samples. However, such assumption is not true for the DFT coefficients obtained for very short segments [6]. Since the speech signal is non-stationary, so its spectral components are obtained by Short-Time Fourier Transform (STFT) analysis in which signal is framed into overlapping pseudo-stationary segments of 20ms -40ms. The Gaussian assumption about PDF of such spectral components seems shifted from the reality. However, in addition to Wiener filtering, other algorithms such as in [6] also Gaussian statistical model has been assumed. Recently, Laplacian Distribution (LD) has been used for the clean speech signal and an algorithm has been derived for the speech enhancement [7]. The statistical model used in a speech enhancement algorithm matters and careful selection of the best matching theoretical PDF model is essential for better performance. There have been several studies on the statistical modeling of the DFT coefficients of speech and accordingly LD, Gamma distribution, Generalized Gaussian Distribution (GGD) etc. have been suggested [6, 8]. The other inconsistency arises due to use of GD for noise. All noise signals corrupting a speech signal are not Gaussian and their spectral components too have non-Gaussian distributions. There are several real world noise signals e.g. chair crack, page turning and tearing, babble (BAB) noise, speech like noise that are non-Gaussian [9]. In Fig.1 the scaled histograms for the real parts, imaginary parts and magnitudes of DFT coefficients of White Gaussian Noise (WGN) and BAB noise are shown along with the fittings of GD, LD and GGD from where it can be emphasized that the spectral components of noise are also not always Gaussian. The main objective of present study is inspired by such inconsistencies. As a solution, we propose a new algorithm under the Bayesian learning framework for the estimation of STSA of the clean speech using parametric GGD function as the statistical model for the DFT coefficients of clean speech as well as noise. We will derive MAP estimator for the STSA of the clean speech from the noisy spectral components.

Rest of the paper is organized as follows. Section II provides working signal model. Section III presents MAP estimation under the GGD prior and section IV procures experimental results. Then paper ends with a few references.

2. Problem Formulation

Let us consider the case of speech signal pick-up by a microphone under the additive background noise $d(n)$. The captured speech $y(n)$ in the presence of noise is given by

$$y(n) = x(n) + d(n), \qquad (1)$$

where $x(n)$ represents clean speech signal, n is the time-

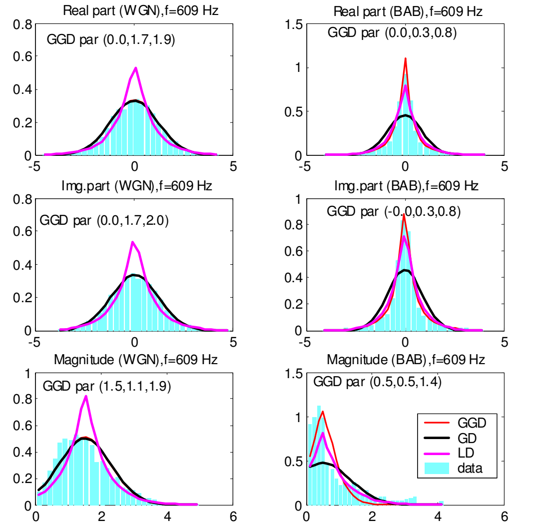


Figure1: Scaled relative frequency histograms of real part, imaginary part, and magnitude of spectral components, for f=609.37Hz, of the white Gaussian noise (subplots in left column) and babble noise (subplots in the right column) and fitted GGD, GD, and LD functions.

index, and random noise *d(n)* is uncorrelated with the clean speech signal. The speech enhancement algorithms estimate clean signal $\hat{x}(n)$ from the observed noisy signal $y(n)$. In the present study we will work in the frequency domain where DFT coefficients, as obtained by STFT method, of the clean speech are estimated. The STFT analysis of Eq.(1) gives following frequency domain additive signal model

$$Y(f) = X(f) + D(f). \qquad (2)$$

This indicates that the signal in any frequency bin is resultant of spectral components of the clean signal and noise and thus in any frequency bin their relative phase as well as magnitudes decide the content of the observed spectral components [10]. It is possible to access and manipulate, according to strength of contamination effect of noise, each spectral component to reduce the noise contained in. A speech enhancement algorithm in frequency domain modifies each spectral components $Y(f)$ by some function known as modification function aka noise suppression rule $G(f)$ to estimate clean signal spectral components $\hat{X}(f)$ given as

$$\hat{X}(f) = G(f).Y(f). \qquad (3)$$

The numerical value of modification function lies between 0 and 1 which implies that the noise suppression is more for the low SNR and less for higher SNR conditions. As stated earlier, GGD function will be used here as the statistical model for the spectral components of the speech and noise signals [8] assuming that the spectral components are mutually uncorrelated in any frequency bin. The GGD function is a parametric function defined in terms of mean μ, scale parameter *b*, and shape parameter β. The GGD PDF for an arbitrary zero mean random variable *z* is given by

$$f_{GG}(z;\mu,b,\beta) = \frac{\beta}{2b\Gamma(1/\beta)}\exp\left(-[|z-\mu|/b]^{\beta}\right) = A\exp\left(-[|z|/b]^{\beta}\right) \qquad (4)$$

where

$$A = \frac{\beta}{2\Gamma b(1/\beta)},\ \frac{1}{b} = \frac{1}{\sigma}\sqrt{\frac{\Gamma(3/\beta)}{\Gamma(1/\beta)}}; \qquad (5)$$

$\Gamma(x) = \int_0^{\infty} e^{-t} t^{x-1} dt$=Gamma PDF;

$-\infty<z<\infty$; $b>0$; $\beta>0$; σ=standard deviation(Stdv.)

The shape of the GGD function depends on the value of shape parameterβ. For β=1 and β=2, GGD represents LD and GD respectively and the shape of distribution tends to become uniform as$\beta\rightarrow\infty$. The GGD function for the different values of b and β are shown in Fig.2.

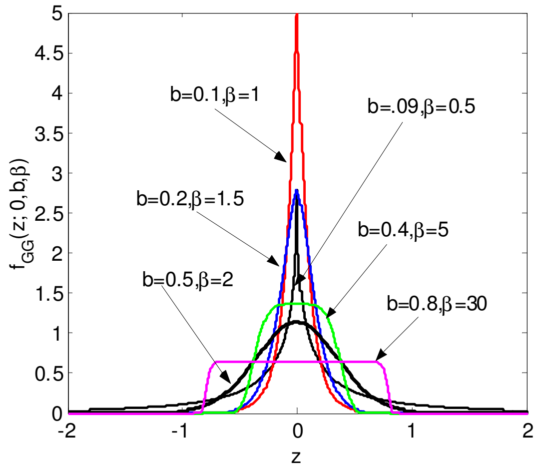


Figure 2: GGD distribution functions for different values of scale parameter *b* and shape parameter β. This family of distribution includes Gaussian, sub-Gaussian and super-Gaussian distributions.

3. MAP Estimation of DFT Coefficients

Maximum a posteriori (MAP) is one of the most widely used statistical estimation techniques. As its name indicates, it estimates an unknown variable as the maximum of its posterior PDF. MAP estimation uses past knowledge as well as new information about the variable to be estimated. The past statistical knowledge is incorporated into the prior PDF and new information in the current data is included in the likelihood function. The known prior PDF and likelihood function are combined together to obtain posterior PDF. The maximum of this posterior PDF is taken as the most possible or estimated value of the unknown variable. Thus informativeness of the used *a priori* is crucial as the posterior PDF is only improved version of the priori and, therefore, it should be chosen carefully. In order to formulate the problem of estimation of the clean speech spectral component, we represent the spectral components of the speech signal and noise in the polar form and then will do estimation of phase and magnitude. Let in the kth frequency bin $Y_k = R_k e^{i\nu_k}$ represents a spectral component of the noisy signal and

$X_k = a_k e^{i\alpha_k}$ represents that of the clean signal in the polar form. The MAP estimators of magnitude a_k and phase α_k are given as the maximum of the posterior PDF $p(a_k, \alpha_k \mid Y_k)$, which is given by Bayes' theorem in terms of likelihood function $p(Y_k \mid a_k, \alpha_k)$ and a prior PDF $p(a_k, \alpha_k)$ as follows

$$p(a_k, \alpha_k \mid Y_k) = p(X_k \mid Y_k) = \frac{p(Y_k \mid a_k, \alpha_k)\, p(a_k, \alpha_k)}{p(Y_k)}. \tag{6}$$

Since $p(Y_k)$ is constant with respect to (w.r.t) spectral magnitude a_k and phase α_k, only numerator of Eq.(6) is significant in the optimization landscape and denominator will be dropped hereafter. The natural logarithmic function of only numerator is optimized, which is given by

$$J = \ln[p(Y_k \mid a_k, \alpha_k)\, p(a_k, \alpha_k)]. \tag{7}$$

The MAP estimators of magnitude a_k and phase α_k are given by

$$(\hat{a}_k, \hat{\alpha}_k) = \arg\max_{(a_k, \alpha_k)} \{J\} = \arg\max_{(a_k, \alpha_k)} [\ln\{p(Y_k \mid a_k, \alpha_k)\, p(a_k, \alpha_k)\}]. \tag{8}$$

Obviously, MAP estimation needs knowledge of conditional probability $p(Y_k \mid a_k, \alpha_k)$ and prior probability $p(a_k, \alpha_k)$ of the spectral components of clean speech for which GGD will be used here. The GGD model for the magnitude of the DFT coefficients of the clean speech signal in the kth frequency bin, is given by

$$p(a_k) = \begin{cases} A_x e^{-\left[\frac{|a_k|}{b_x}\right]^{\beta_x}}, & for\ 0 \le a_k \le \infty \\ 0 & \text{else} \end{cases}, \tag{9}$$

where b_x is the scale parameter, β_x is shape parameter,

$$A_x = \frac{\beta_x}{2b_x \Gamma(1/\beta_x)} = \frac{\beta_x}{2\Gamma(1/\beta_x)} \frac{1}{\sigma_x} \sqrt{\frac{\Gamma(3/\beta_x)}{\Gamma(1/\beta_x)}}; \tag{10}$$

and, $\sigma_x = \text{Stdv.of clean speech} = \sqrt{E\{a_k^2\}}$. (11)

Since the positions of analysis window in STFT analysis are arbitrary, the PDF of phase α_k, follows uniform distribution and is expressed as

$$p(\alpha_k) = Uniform\ \text{PDF} = \begin{cases} \frac{1}{2\pi} & \text{for } -\pi \le \alpha_k \le \pi \\ 0 & \text{else} \end{cases}. \tag{12}$$

The joint PDF of the magnitude a_k and phase α_k is given as

$$p(a_k, \alpha_k) = \frac{A_x}{2\pi} e^{-\left[\frac{|a_k|}{b_x}\right]^{\beta_x}}, \tag{13}$$

$$for\ 0 \le a_k \le \infty \text{ and } -\pi \le \alpha_k \le \pi.$$

The conditional probability $p(Y_k \mid X_k)$ of the observed data, given clean signal inherits randomness of noise and can be given as the PDF of noise, which is also modeled by GGD as follows

$$p(Y_k \mid X_k) = p(Y_k \mid a_k, \alpha_k) = Noise\ \text{PDF} \tag{14}$$

$$= (A_N/2\pi)\, e^{-\left[\frac{|Y_k - X_k|}{b_n}\right]^{\beta_n}},$$

$$for\ 0 \le |Y_k - X_k| \le \infty, -\pi \le \gamma_k \le \pi$$

where b_n and β_n are scale and shape parameters, respectively, for the GGD distribution for noise spectral component in the frequency bin k, and γ_k is the corresponding phase for noise.

$$A_N = \frac{\beta_n}{2b_n \Gamma(1/\beta_n)} = \frac{\beta_n}{2\Gamma(1/\beta_n)} \frac{1}{\sigma_n} \sqrt{\frac{\Gamma(3/\beta_n)}{\Gamma(1/\beta_n)}}, \tag{15}$$

and, $\sigma_n = \text{Stdv.of noise} = \sqrt{E\{|D_k^2|\}}$. (16)

Now the desired posterior density in Eq.(6) can be given by using Eq.(13) and Eq.(14) and dropping denominator as follows

$$p(a_k, \alpha_k \mid Y_k) \propto p(Y_k \mid a_k, \alpha_k)\, p(a_k, \alpha_k)$$

$$= (A_N A_x / 4\pi^2)\, e^{-\left[\frac{|Y_k - X_k|^{\beta_n}}{b_n^{\beta_n}}\right] - \left[\frac{|a_k|^{\beta_x}}{b_x^{\beta_x}}\right]}. \tag{17}$$

Using Eq.(17) in Eq.(7) gives

$$J = -\frac{|R_k e^{iv_k} - a_k e^{i\alpha_k}|^{\beta_n}}{b_n^{\beta_n}} - \frac{|a_k|^{\beta_x}}{b_x^{\beta_x}} + \ln\frac{A_N A_X}{4\pi^2}. \tag{18}$$

Now, in order to locate, say at $\hat{\alpha}_k$ and $\hat{a}_k$, the highest of the posterior PDF, differentiating Eq.(18) w.r.t. phase α_k and spectral amplitude a_k, and equating derivatives to zero gives

$$\partial J/\partial \alpha_k \big|_{\alpha_k = \hat{\alpha}_k} = B[R_k^2 + a_k^2 - 2R_k a_k \cos(v_k - \hat{\alpha}_k)]^{\beta} [2R_k a_k \sin(v_k - \hat{\alpha}_k)],$$

Equating it with zero and further simplification gives

$$\sin(v_k - \hat{\alpha}_k) = 0 \Rightarrow \hat{\alpha}_k = v_k, \tag{19}$$

where $B = 0.5\beta_n / b_n^{\beta_n}; \beta = 0.5\beta_n - 1$. Eq.(19) gives MAP estimated phase of the spectral components of clean signal which is same as that of that of the spectral components of the noisy speech. Similar treatment of Eq.(18) w.r.t spectral amplitude a_k along with use of Eq.(19) gives

$$\partial J/\partial a_k \mid_{a_k = \hat{a}_k} = 0$$

which further gives

$$2B[R_k^2+\hat{a}_k^2-2R_k a_k]^{\beta}[-R_k+\hat{a}_k] \\ -\beta_x \hat{a}_k^{\beta_x-1} b_x^{-\beta_x}[sign(\hat{a}_k)]^{\beta_x}=0, \quad (20)$$

In order to avoid singularity when $0<\beta_x<1$, and $a_k=0$, $\hat{a}_k^{\beta_x-1}$ in Eq.(20) is replaced by $\hat{a}_k^{\beta_x-1}+\delta$, where δ is very small ($<10^{-4}$) number. Further simplification of Eq.(20), results in the following radical (power) equation

$$\beta_x \hat{a}_k^{\beta_x-1} b_x^{-\beta_x} sign(\hat{a}_k)^{\beta_x} = 2B(R_k-\hat{a}_k)^{2\beta+1} \\ \Rightarrow \hat{a}_k^{\beta_x-1}=P(R_k-\hat{a}_k)^{\beta_n-1}, \quad (21)$$

where $P = b_x^{\beta_x}\beta_n / b_n^{\beta_n}\beta_x$. It may be very difficult to find an analytical solution of the Eq.(21), however, its numerical solution can be easily obtained by Newton-Rapshon's method under which numerical solution after the *i*th iteration is given as

$${}^{i+1}\hat{a}_k = {}^{i}\hat{a}_k - \frac{{}^{i}\hat{a}_k^{\beta_x-1}-P(R_k-{}^{i}\hat{a}_k)^{\beta_n-1}}{(\beta_x-1)\,{}^{i}\hat{a}_k^{\beta_x-2}+P(\beta_n-1)(R_k-{}^{i}\hat{a}_k)^{\beta_n-2}}. \quad (22)$$

This solution gives MAP estimator of the spectral magnitude which is further combined with the phase of a related noisy spectral component to get a spectral component of the clean signal. The solution in Eq.(22) is sensitive to the used initial value. The good initial values can be obtained as the special case solutions of the Eq.(21) as described below

Initial Values

1. For $\beta_x=\beta_n=2$, the spectral components of both the noise and speech signal have Gaussian (assumption working under the Wiener filtering) PDF under which solution of the Eq.(21) using Eq.(5) is given by

$$\hat{a}_k=\frac{b_x^2}{b_x^2+b_n^2}R_k=\frac{\sigma_x^2}{\sigma_x^2+\sigma_n^2}R_k, \quad (23)$$

which is Wiener filter and can be used as an initial value for the iterative solution in Eq.(22).

2. When $\beta_x=1$ i.e. the clean speech spectral component has Laplacian PDF and that of noise is GGD, the solution to Eq.(22) is given by

$$\hat{a}_k=R_k-\left(\frac{1.4142\sigma^{\beta_n}{}_n}{\sigma_x\beta_n}\right)^{\frac{1}{\beta_n-1}}\left[\frac{\Gamma\left(1/\beta_n\right)}{\Gamma\left(3/\beta_n\right)}\right]^{\frac{0.5\beta_n}{\beta_n-1}} \quad (24)$$

in which further if the PDF of noise spectral components is assumed to be Gaussian, we have $\beta_n=2$ and Eq.(24) simplifies into

$$\hat{a}_k=R_k-1.4142\frac{\sigma_n^2}{\sigma_x}=R_k-1.4142\frac{\sigma_x}{\xi} \quad (25)$$

where $\xi=\sigma_x^2/\sigma_n^2$ is the spectral SNR of the noisy speech signal. It is interesting to note that for $\beta_x=\beta_n=1$, Eq.(21) collapses and it can be shown that under such a condition Eq.(21) leads to $\sigma_x=\sigma_n$. Out of above mentioned two initial values, Eq.(24) seems more natural in the light of results reported in [8] and will be used here, too.

3.1. GGD parameter estimation

The solutions of Eq. (21), require scale and shape parameters of clean speech and noise signals which are not known, however, can be estimated from the noisy data only. The GGD parameters of noise can be estimated, using Maximum Likelihood (ML) approach [8, 11], from the noise only portion e.g. a few samples from the beginning or other silent parts of the noisy data can be taken. The estimated noise parameters and statistics of the noisy data can be used to estimate GGD parameters for the clean signal. The shape parameter β_x can be estimated from kurtosis K_x by inverting the following relation for GGD

$$K_X=\frac{\Gamma(1/\beta_x)\Gamma(5/\beta_x)}{\Gamma(3/\beta_x)^2}=F(\beta_x),$$

which gives

$$\hat{\beta}_x=F^{-1}(K_x). \quad (26)$$

However, this needs kurtosis of the clean signal, which can be estimated from the higher order statistics of spectral components of the noisy speech and estimated noise as follows

$$K_x=\frac{\left[K_y\sigma_y^4-4S_x\sigma_x^3\mu_n-4S_n\sigma_n^3\mu_x-6\sigma_y^2\sigma_n^2-(K_n-6)\sigma_n^4\right]}{(\sigma_y^2-\sigma_n^2)^2}, \quad (27)$$

where

$$S_x=\frac{\left[S_y\sigma_y^3-3(\sigma_y^2-\sigma_n^2)\mu_n-3(\mu_y-\mu_n)\sigma_n^2-S_n\sigma_n^3\right]}{\left(\sigma_y^2-\sigma_n^2\right)^{3/2}},$$

$\sigma_x^2=\max\left(\sigma_y^2-\sigma_n^2,0\right)$; and $\mu_x=\mu_y-\mu_n$,

in which the subscripts *n* denotes noise and *y* denotes noisy speech signal; *S*, σ and μ denote skewness, standard deviation and mean of the signals indicated by the subscripts. These estimated GGD parameters for noise and clean speech signals can be used in Eq.(22) to estimate spectral magnitudes which in combination with the noisy phases give spectral components of the clean speech signal. The Wiener filtering can be done using Eq.(23). In order to evaluate performance of the proposed denoising algorithm, global SNR and segmental SNR of the estimated signal will be measured. The global SNR provides error measurement over time and frequency and is defined as

$$SNR_{dB}=10\log_{10}\frac{\sum x^2(t)}{\sum[x(t)-\hat{x}(t)]^2}, \quad (28)$$

where $x(t)$ and $\hat{x}(t)$ represent clean signal and estimated clean signal. Segmental SNR represents SNR for each frames. The mean square errors (MSE) between estimated and clean signal will also be computed for both the Wiener estimated and MAP estimated clean signals.

4. Experiments and Results

In the experiments the clean speech data was taken from the ASJ continuous speech corpus for the research [8] and noise data from the NOISEX-92 database freely available at http://mi.eng.cam.ac.uk/comp.speech/Section1/Data/noisex.html. The clean speech signal was degraded by WGN and BAB noise to the different SNR levels ranging between –5db and 20 dB. In the enhancement process only noised speech signals were used. In the STFT analysis, hanning window of 20 ms with step size 10 ms, sampling frequency 8000 Hz and 512 point DFT were used. For the estimation of noise parameters a few samples, depending upon the length of initial silence period, were used. First, GGD parameters for the spectral components of the noise were estimated, using ML approach, from a first few frames corresponding to noise only samples. Then higher order statistics of the noisy signal and GGD parameters of noise were used to estimate GGD parameters for the clean speech using Eq.(26)- Eq.(27). The estimated parameters for the speech signal, from a female speaker degraded by the babble noise, are shown in Fig.3. The shown parameters are averaged over the number of frequency bins. The speech enhancements using MAP estimation and Wiener filtering were done on the four speech signals degraded by WGN and BAB noise. The SNRs of the input and estimated clean signals are shown in the Fig.4. It is evident from there that the proposed method outperforms Wiener filtering. MAP estimation provides better SNR than modified Wiener filter [12]. The normalized MSE between the estimated and original signals are also shown in Fig.5. It can be observed that the noise suppression capacity of GGD based algorithm is superior to that of the Wiener filtering. The segmental SNR for the speech signal from the female speaker are shown in Fig. 6a and Fig.6b. It is evident from these figures

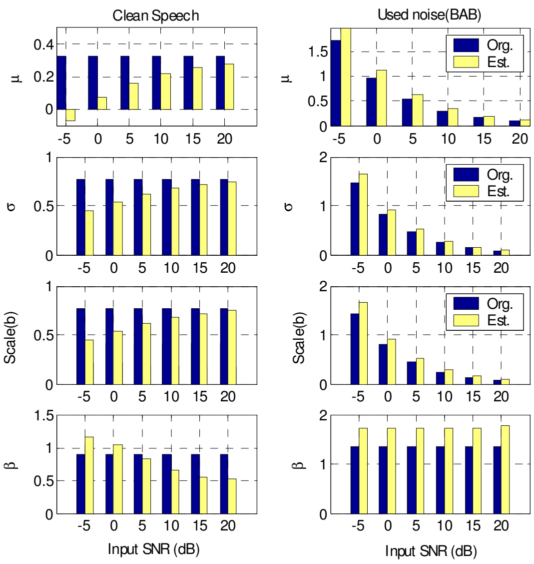


Figure 3: Estimated GGD parameters for the clean speech and noise from noisy speech data degraded by BAB noise. The estimated values are averaged over the frequency bins. Subplots in the left column are for clean speech while the same in the right column are for the BAB noise.

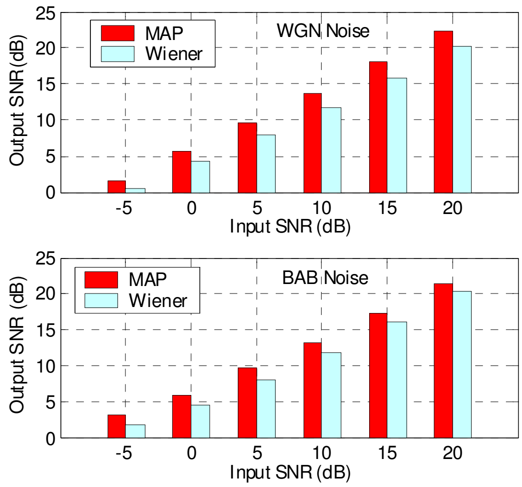


Figure 4: SNRs of the estimated speech under different degraded conditions of the input by the WGN (upper subplot) and BAB noise (lower subplot). The result is averaged for the speech data for the four speakers (two male and two females).

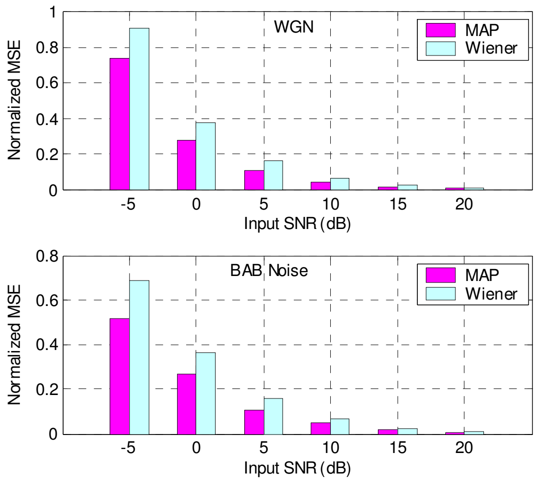
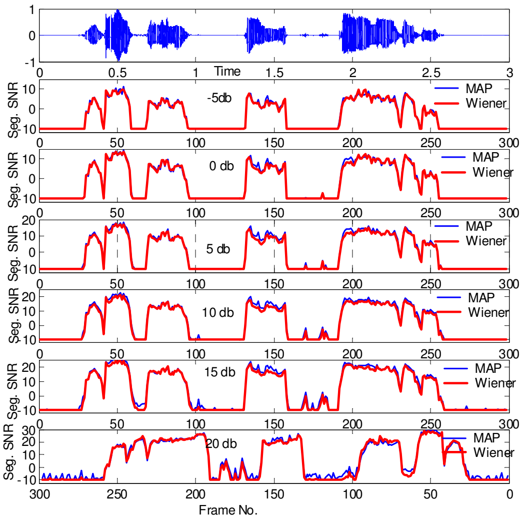
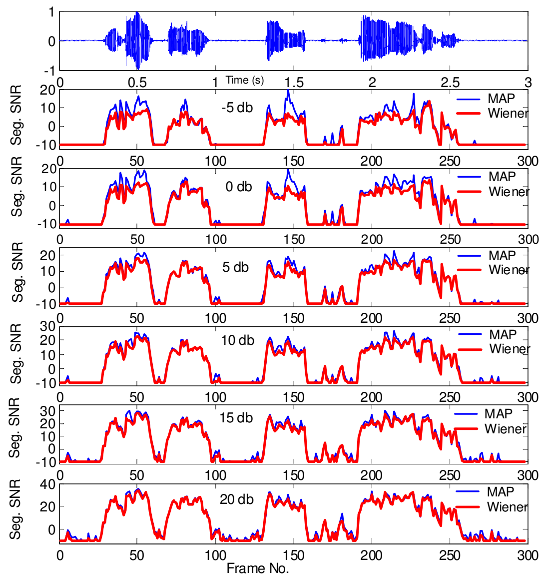


Figure 5: MSE between estimated signal and original speech signal for different SNR conditions of the input signal degraded by WGN and babble noise. The shown results are averaged for the four speakers (two male and two females).

that the segmental performance of the GGD based MAP estimator is also better than that of the Wiener filter. The shown segmental SNR were limited between –10 db to 35 db. The spectrograms of the noisy signals, under WGN, estimated clean speech signals by Wiener filtering and proposed method are also shown in Fig.7. For different SNR levels of the input signal. With the increasing input SNR, difference in segmental SNR of both gets lessened. In all these experiments GGD parameters for the noise has been estimated using a few samples from the initial silence part of the noisy signal. However, silence parts are also available at other time instances that can be further utilized to increase the number of noise only frames which in turn can improve the ML estimation of the GGD parameters of the noise as well as of clean speech. In the MAP estimation process we



(a)



(b)

Figure 6: Segmental SNR for MAP estimation and Wiener filtering computed for female speech signal degraded by (a) WGN and (b) babble noise.

have used the initial value based on Eq.(24) , however, its comparative study of performance with Wiener based initial values is left unexplored. Looking effectiveness of Eq.(24), we are studying its use as adaptive spectral subtraction.

5. References

[1] Furui,S., Sondhi, M.M., (Eds.), *Advances in Speech Signal Processing,* Narcel Dekker Inc.,USA, 1991.

[2] McAulay, R.J., Malpass, M.L., "Speech Enhancement using Soft Decision Noise Suppression filter. IEEE Trans. Acoust., Speech, Signal Process. ASSP-28, 137-145, 1980.

[3] Simon J. G. and Peter J.W. R, *Digital Audio Restoration*, Springer, London, UK, 1998.

[4] Weiss, M.R., Aschkenasy,E., and Parsens T.W., " Study and Development of the INTEL Technique for Improving Speech Intelligibility", Report NSC- FR/4023,Nicolet Scientific Corp., December, 1974.

[5] Boll S., "Suppression of Acoustic Noise in Speech Using Spectral Subtraction", IEEE Trans. Acoust., Speech, Signal Processing, Vol. ASSP-27(2), 113-120, 1979.

[6] Ephraim A., Malah D., "Speech Enhancement Using a Minimum Mean Square Error Short-Term Spectral Amplitude Estimator", IEEE Trans. Acoust., Speech, Signal Processing, Vol. ASSP-32, no. 6 (Dec.1984) 1109-21,1984.

[7] Martin R., Colin B., "Speech Enhancement in the DFT Domain Using Laplacian Priors", Proc. IWAENC, 87-90, Kyoto, Japan, 2003.

[8] Prasad R.K., Saruwatari H., Shikano K., "Probability Distribution of Time-Series of Speech Spectral Components", IEICE Trans. Fundamental of Elec.,Com., CS., Vol.E87- A,No.3, 584-597, 2004.

[9] Panayotis G.G., Panagiotis T., Chris K., "Alpha-Stable Modeling of Noise and Robust Time Delay Estimation in the Presence of Impulsive Noise", IEEE Trans. on Multimedia, Vol.1, No.3, 291-301, 1999.

[10] Zhu Q., Alwan A., "The Effect of Additive Noise on Speech Amplitude Spectra: A Quantitative Analysis",IEEE, SPL, vol.9,275-277, Sept. 2002.

[11] Varanasi M.K. and Aazhang B., "Parametric Generalized Gaussian Density Estimation", J. of Acost. Society of America, Vol. 86(4):1404-1414, (1989)

[12] Ogata, S., Shimamura, T., "Improvement of Wiener filter for Noise Reduction", Proc. WESTPRA VII, vol.1:477-480, Kumamoto, Japan, 2000.

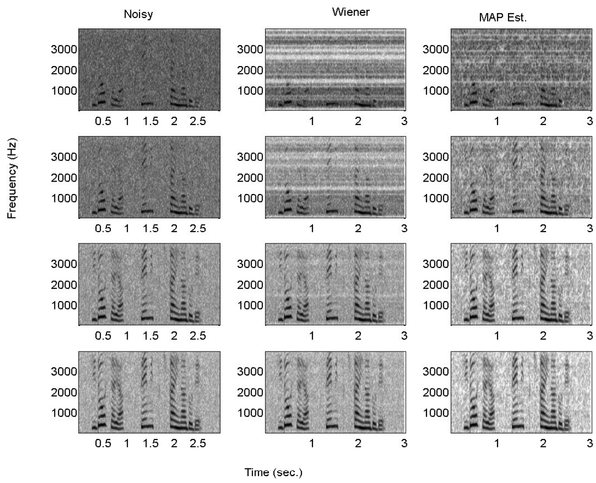


Figure 7: Spectrogram of the noisy and enhanced speech signal. The subplots from top to bottom in any column correspond to SNR conditions –5db, 5 db, 15 db and 20 db. Subplots in first column are for noised signals, subplots in second column are for enhanced signals by Wiener filtering, and that of in the third column is for the proposed algorithm.