K. Bielawski, A.A. Petrovsky, *Proposition of minimum bands multirate noise reduction system which exploits properties of the human auditory system and all-pass transformed filter bank*. IEEE Workshop SIGNAL PROCESSING'2001, Poznań 2001, pp. 65-70.

PROPOSITION OF MINIMUM BANDS MULTIRATE NOISE REDUCTION SYSTEM WHICH EXPLOITS PROPERTIES OF THE HUMAN AUDITORY SYSTEM AND ALL-PASS TRANSFORMED FILTER BANK†

Krzysztof Bielawski, Alexander Petrovsky

Białystok Technical University
Faculty of Computer Science, Real Time Systems Department
Wiejska 45A, 15-351 Białystok, Poland
e-mail: biel@ii.pb.bialystok.pl

Abstract: *This paper introduce the new approach to noise reduction in order to improve the speech intelligibility. The system is proposed with minimum band requirement to approximate psychoacoustic Bark scale and nonuniform filter bank constructed with the use of first order all-pass transformation. Proposed psychoacoustic weighting exploits the well known audible noise reduction rule in case of the 16-bit signal and sampling rate of 8 kHz.*

1 Introduction

The last decade shows that the noise reduction in single microphone device is still an open problem. The well known Spectral Subtraction rules [1,2] and Minimum Mean-Square Spectrum enhancement method [3,4] are just the starting point to develop the improved and quite new approach. To satisfy the listener expectation of signal quality and eliminate the annoying musical noise and unnatural sounding of the processed speech the new solution searches the inspiration in an auditory property of the human inner ear and brain processing resulting in new class of the psychoacoustic motivated noise reduction system [5-10] and improvement of the well known methods [11-15].

Mainly all solutions exploit the masking property of the human ear and calculate the excitation pattern and masking threshold based on Fourier analysis or Wavelet transform. Mentioned Fourier decomposition method provides the frequency resolution much greater than the auditory Bark scale commonly used. While the wavelet transform provides the step-linear approximation of the auditory scale. Then by simple addition of the overload decomposition calculate the psychoacoustic information in the Bark bands.

In this paper the multirate system based on filter bank with the quantity of bands equal the number of Barks for assumed sampling frequency is proposed. Solution is mentioned for the 8 kHz sampling frequency in hands-free device, which is a minimum cost noise reduction system considering the number of bands and sampling frequency but still satisfying the intelligibility improvement of the noisy speech.

2 Audible noise suppression rule

Proposed solution uses the audible noise suppression rule shown in [8]. Assuming that the speech $s(k)$ is corrupted by the additive noise $n(k)$ resulting in the noised signal

$$y(k)=s(k)+n(k), \tag{1}$$

with the same relation of the signals in the bands:

$$y_m(k)=s_m(k)+n_m(k), \qquad 0\leqslant m\leqslant M-1, \tag{2}$$

where the subband signal $y_m(k)$ is a result of bandpass filtering of the filter $h_m(k)$ and M is the quantity of the bands

$$y_m(k)=\sum_n h_m(n)\,y(k-n), \quad \text{where } 0\leqslant m\leqslant M-1. \tag{3}$$

Assuming the power estimators of the signals $y_m(k), s_m(k), n_m(k)$ in the processing frame of the length W as $P_{y,m}(k), P_{s,m}(k), P_{n,m}(k)$ to be calculated according to the Eqs. (4) and (5).

$$P_{y,m}(k_b)=\frac{1}{W}\sum_{i=0}^{W-1} P'_{y,m}(k_b W+i), \tag{4}$$

where k_b is processing block index, while the power estimation is calculated based on exponential averaging:

$$P'_{y,m}(k)=\alpha|y_m(k)|^2+(1-\alpha)P'_{y,m}(k-1). \tag{5}$$

The audible fragments of the band signals can be defined as

$$S_{x,m}(k_b)=\begin{cases} P_{s,m}(k_b), & \text{if} \quad P_{s,m}(k_b)\geqslant T_m(k_b)\\ T_m(k_b), & \text{if} \quad P_{s,m}(k_b)\leqslant T_m(k_b)\end{cases} \qquad 0\leqslant m\leqslant M-1, \tag{6}$$

$$S_{y,m}(k_b)=\begin{cases} P_{y,m}(k_b), & \text{if} \quad P_{y,m}(k_b)\geqslant T_m(k_b)\\ T_m(k_b), & \text{if} \quad P_{y,m}(k_b)\leqslant T_m(k_b)\end{cases} \qquad 0\leqslant m\leqslant M-1, \tag{7}$$

where auditory masking threshold $T_m(k_b)$ for band m and

† This work was supported by Białystok Technical University under the project W/II/2/01

block k_b is calculated according to the [16] (see Appendix A).

Therefore the audible noise is defined as

$$S_{n,m}(k_b) = S_{y,m}(k_b) - S_{s,m}(k_b). \tag{8}$$

Using Eqs. (6) and (7) in (8) the audible noise for bands $0 \leqslant m \leqslant M-1$, is defined as

$$S_{n,m}(k_b) = \tag{9}$$

$$\begin{cases} P_{y,m}(k_b) - P_{s,m}(k_b) & \text{if} \quad P_{y,m}(k_b) \geqslant T_m(k_b) \text{ and } P_{s,m}(k_b) \geqslant T_m(k_b) \quad (I) \\ P_{y,m}(k_b) - T_m(k_b) & \text{if} \quad P_{y,m}(k_b) \geqslant T_m(k_b) \text{ and } P_{s,m}(k_b) < T_m(k_b) \quad (II) \\ T_m(k_b) - P_{s,m}(k_b) & \text{if} \quad P_{y,m}(k_b) < T_m(k_b) \text{ and } P_{s,m}(k_b) \geqslant T_m(k_b) \quad (III) \\ 0, & \text{if} \quad P_{y,m}(k_b) < T_m(k_b) \text{ and } P_{s,m}(k_b) < T_m(k_b) \quad (IV) \end{cases}$$

As can be noticed the defined audible noise components depend on powers of the clean speech signal $P_{s,m}(k_b)$, corrupted signal $P_{y,m}(k_b)$ and auditory threshold of the clean speech $T_m(k_b)$ estimated in processing block k_b.
The suppression rule is defined as

$$S_{n,m}(k_b) \leqslant 0, \qquad 0 \leqslant m \leqslant M-1, \tag{10}$$

where the signal weighting coefficient of the method is estimated for each processing block as

$$G_m(k) = \frac{1}{\left(\dfrac{a_m(k_b)}{P_{y,m}(k_b)}\right)^{v_m} + 1}, \quad k_b \leqslant k \leqslant k_b + W \text{ and } 0 \leqslant m \leqslant M-1 \tag{11}$$

where the time-variable $a_m(k_b)$ and $v_m \in \Re^+ \leqslant 1$ are the parameters which determine the level of audible noise suppression defined by (evaluation of parameters shown in Appendix B):

$$a_m(k_b) = [T_m(k_b) + P_{n,m}(k_b)] \left[\frac{P_{n,m}(k_b)}{T_m(k_b)}\right]^{\frac{1}{v_m}}, \tag{12}$$

The coefficient $a_m(k_b)$ defines the threshold above which all noise components are suppressed. Whereas parameter v_m determines the suppression level and depends on ratio $\dfrac{P_{y,m}(k_b)}{a_m(k_b)}$.

3 Nonuniform frequency resolution polyphase filter bank

The uniform filter bank can be constructed as a set of well tuned bandpass filter, but to exploit the properties of multirate technique the polyphase filter bank has been used in proposed solution [17-21].

Designing the bandpass filter as a complex shifted version of the prototype filter $H_0(e^{j\omega})$ with following characteristic:

$$H_m(e^{j\omega}) = H_0(e^{j(\omega - 2\pi m/M)}) \text{ for } m \in \{\mathbb{N} \mid m \leqslant M-1\}. \tag{13}$$

Using the polyphase decomposition of order R where $R \leqslant M$, the permutation set of the following components will be produced:

$$p_\rho(k) = h_0(kR + \rho), \qquad \rho \in \{\mathbb{N} \mid \rho \leqslant M-1\}, \tag{14}$$

where the band signals of analysis stage [17] for input signal are:

$$X_m(k) = \sum_{\rho=0}^{M-1} e^{\frac{j 2\pi m \rho}{M}} \left[\sum_{r=-\infty}^{\infty} p_\rho\left(k - r\frac{M}{R}\right) x(rM - \rho)\right], \quad m \in \{\mathbb{N} \mid m \leqslant M-1\}, \quad r \in \mathbb{C}, \tag{15}$$

and they are synthesised [17] as follows:

$$\hat{x}(rM - \rho) = \sum_{k=-\infty}^{\infty} h_0\left(\left(r\frac{M}{R} + k\right)R - \rho\right) \left[\frac{1}{M} \sum_{m=0}^{M-1} X_m(k) e^{\frac{-j 2\pi m \rho}{M}}\right], \quad m \in \{\mathbb{N} \mid m \leqslant M-1\}, \; r \in \mathbb{C}. \tag{16}$$

The polyphase filter bank with analysis and synthesis stage is depicted in Fig. 1.

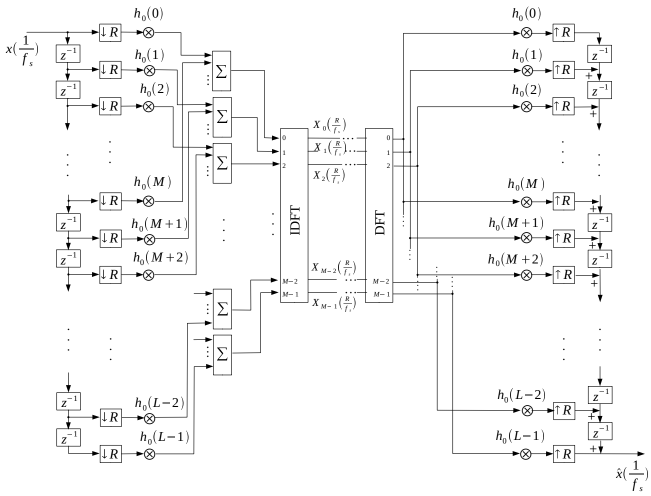

Fig. 1. Direct structure with delay chain of polyphase filter bank

The polyphase nonuniform filter bank approximating the Bark scale can be constructed with the use of the first order bilinear mapping, where the uniform Hz scale z are map to uniform Bark scale which is nonuniform when seen in Hz ζ. The bilinear transform for three point defines:

$$\frac{(\zeta - \zeta_1)(\zeta_2 - \zeta_3)}{(\zeta_2 - \zeta_1)(\zeta - \zeta_3)} = \frac{(z - z_1)(z_2 - z_3)}{(z_2 - z_2)(z - z_3)}. \tag{17}$$

In case when both scale are Z domain the following must be accomplish:

$$\begin{aligned} &z_1 = \zeta_1 = 1 \quad \text{- mapping the unit circle to itself} \\ &z_2 = \zeta_2 = -1 \quad \text{- mapping half of the sampling frequency to itself.} \end{aligned} \tag{18}$$

Using Eq. (18) in (17) the mapping formula is:

$$z = A_\alpha(\zeta) = \frac{\zeta + a}{1 + \zeta a}, \qquad a = \frac{\zeta_3 - z_3}{1 - z_3 \zeta_3}, \tag{19}$$

where a define the freedom of mapping frequency $0 \leqslant \omega \leqslant 2\pi$ related with the point $e^{j\omega}$ on unit circle to a new location $A_a(\omega)$.

From Eq. (19) it can be noticed that bilinear map has the same form as a first order all-pass filter function

$$A_a(\zeta)\equiv H_{AP}(z)=\frac{z^{-1}+a}{1+az^{-1}}, \qquad (20)$$

with all-pass $H_{AP,a}(z)$ filter phase function:

$$\phi(\omega)=2\arctan\left(\frac{1-a}{1+a}\tan\frac{\omega}{2}\right), \qquad (21)$$

which determines the the frequency warping in the first-order all-pass mapping. This technique has been used in past to scaling the cut-off frequency of the lowpass filter [22,23] or to nonuniform representation of the Fourier transform spectrum [24], and lately in audio coding [25].

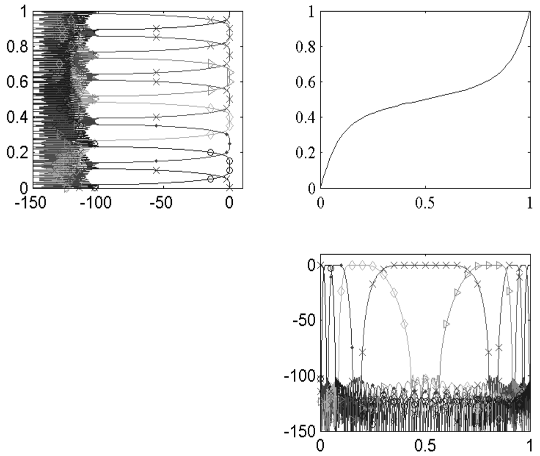

Fig. 2. All-pass transformation of polyphase filter bank

Founding on work [26] and taking the arctangent coefficient approximation of the map from Hz to Bark represented in Hz dependent of sampling frequency the following formula can be used

$$a_{Bark}=1.048\left[\frac{2}{\pi}\arctan(0.07212\frac{f_s}{1000})\right]^{\frac{1}{2}}-0.1957. \qquad (22)$$

To obtain the nonuniform frequency resolution in polyphase filter bank just the replacement of the delay chain by the all-pass chain with propriety filters coefficient is needed

$$z^{-1}\Leftrightarrow\frac{z^{-1}+a}{1+az^{-1}}. \qquad (23)$$

Then the following passband filter characteristic of the transformed filter bank can be achieved:

$$H_m(e^{j\omega})=H_m(e^{-j\phi(\omega)})=H_0(e^{j(-\phi(\omega)-2\pi m/M)}), \qquad (24)$$

where the influence of the all-pass filter phase function is clearly visible. Fig. 2 shows the described mapping process.

The mentioned filter bank must be designed in pragmatic way to find the compromise between the overall filter bank performance and nonuniform spreading band usage.

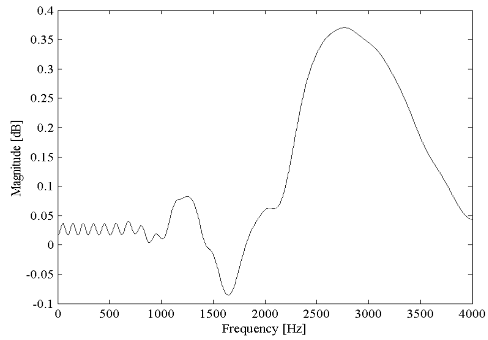

Fig. 3. Logarithmic Fourier transformation of the impulse response of the analysis-synthesis structure.

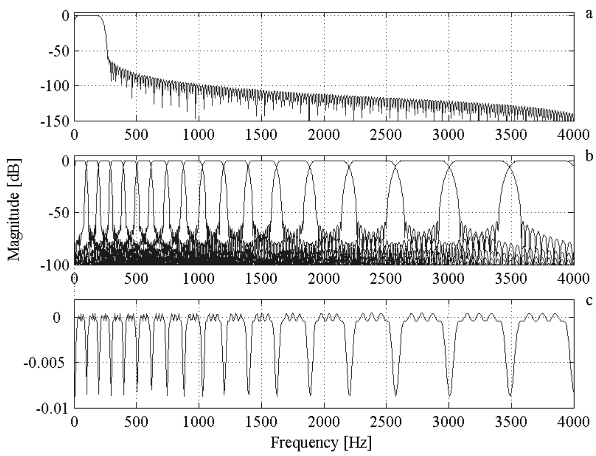

Fig. 4. Filter bank characteristic (a) prototype filer (b) analysis filter bank stage (c) sum characteristic of the analysis stage

The design example of the Bark spreaded filter bank is presented in Fig. 3. Also its magnitude characteristic after two stage processing in filter bank is depicted in fig. 4, where the mapping spread of uniform alliasing is visible for higher frequency. Such carefully constructed filter bank can achieve only the the nearly perfect reconstruction property, which results in non-audible distortion of signal.

4 Noise reduction system evaluation

The presented earlier weighting rule is based on power estimates of the clean speech signal and noise, but system setup provides only the corrupted signal so the intermediate technique is used to estimate the masking threshold of the clean speech. Fig. 5. presents the processing schema for bandpass signal.

Nonlinear noise estimate tracking method [27] with linear averaging for each frame, as in case of signal power spectrum is used. Also the rough speech enhancement method to get the estimate of the clean speech based on power spectral suppression rule [2] is used. Examples of system work results are depicted in Figs. 6 and 7.

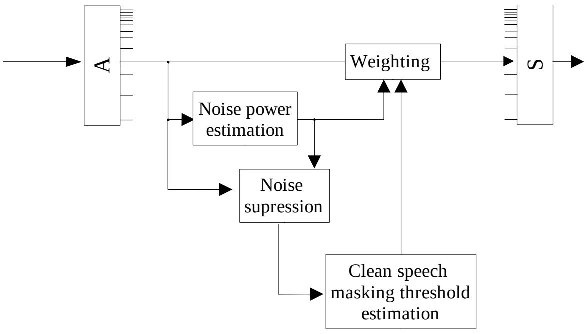

Fig. 5. In band processing schema

However, the speech enhancement quality strongly depends on noise tracking and masking threshold estimation so the careful design of estimation method must be conducted.

To test the weighting rule derived in the previous section, it was implemented in analysis/synthesis structure of polyphase all-pass transformed filter bank with band number M=36, decimation ratio $R=9$, all-pass coefficient $a=-0{,}4$, frame length $W=16$, and prototype filter of the bank length of 180 coefficients. Test was conducted to 8kHz, 16-bit wav file recorded in car cabin (speech engine off and noise at speed 100 km/h). The signals was mixed together to provide the test signal at different signal-to-noise ratio calculated at speech activity [28-29]according to average value of i frame ratio:

$$SEGSNR_n^{s,i}=10\log_{10}\left(\frac{\sum_{k=0}^{W-1} s^2(k+iW)}{\sum_{k=0}^{W-1} n^2(k+iW)}\right), \qquad (25)$$

where i is a frame index of speech activity, W - frame length, s - speech, n -noise signals.

The objective measure of noise attenuation NA was taken and also $SEGSNR_{\hat{s}-s}^{s}$ difference of speech and speech distortion after processing which has the high correlation with results from auditive tests.

$$NA=10\log_{10}\left(\frac{1}{O(K_n)}\sum_{k\in K_n}\frac{n^2(k)}{\hat{n}^2(k)}\right), \qquad (26)$$

where K_n is a set of speech pauses, $\hat{n}(k)$ attenuated noise, $O(K_n)$ number of samples in set.

$$SEGSNR_{\hat{s}-s}^{s,i}=10\log_{10}\left(\frac{\sum_{k=0}^{W-1} s^2(k+iW)}{\sum_{k=0}^{W-1}\left(\hat{s}(k+iW)-s(k+iW)\right)^2}\right), \qquad (27)$$

where $\hat{s}(k)$ is an enhanced speech the $SEGSNR_{\hat{s}-s}^{s}$ is average value over the set of speech activity from $SEGSNR_{\hat{s}-s}^{s,i}$.

5 Conclusion

In this paper, psychoacoustically motivated speech enhancement algorithm with subband approach to frequency decomposition has been presented. According to informal listening test with various speech material, presented rule offers a performance superior to conventional spectral subtraction rule used in it to get the rough clean speech estimate, and with minimum requirements needed provides close performance of other DFT based psychoacoustical methods, what has been shown in objectives tests depicted in Fig. 8.

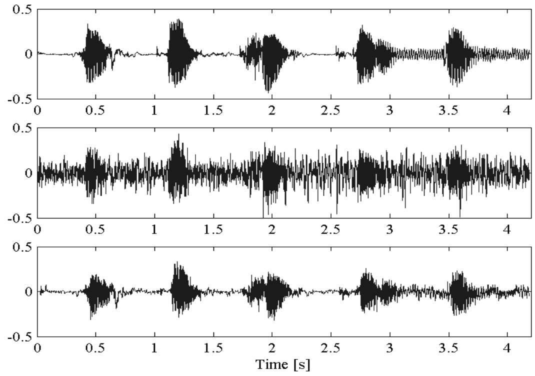

Fig. 6. Time domain plots of the signals: clean, noise corrupted and enhanced

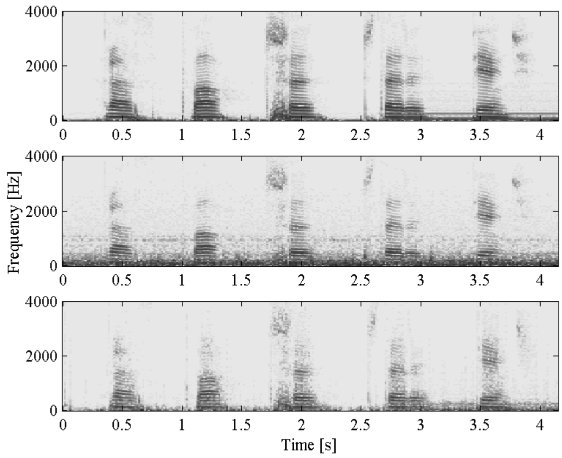

Fig. 7. Corresponding spectrograms for plots in Fig. 6

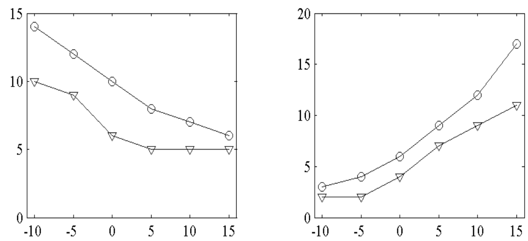

Fig. 8. Instrumental measurement data obtained from simulation of proposed rule (triangle) and the psychoacoustical most advanced rule presented in [9] results provided by author (circle)

Appendix A

Using linear averaging to the power estimate calculated according to the (5) we get the estimate of the power spectrum of the band's signal:

$$P_{y,m}(k_b)=\frac{1}{W}\sum_{i=1}^{W}P_{y,m}(k-i), \tag{28}$$

The spreading function defining the response of the basilar membrane on signal depending of the distance and frequency of the stimuli is defined in dB as:

$$SF_{m_i}(m_i,m_j)=a+\frac{v-u}{2}(m_i-m_j+c)-\frac{v+u}{2}\sqrt{d+(m_i-m_j+c)^2} \tag{29}$$

where the model dependent parameters are shown in tab. 1. The total power spectrum in band is calculated by convolution of the spreading function and power spectrum of the band m for each processing block k_b :

$$D_m(k_b)=P_{y,m}(k_b)*10^{\frac{SF_m}{10}} \tag{30}$$

Tab. 1. Model dependent parameters of the spreading fun.

Model	v	u	d	c	a
Schroeder	25 dB/Bark	10dB/Bark	1	4,74	15,81
Bark	30 dB/Bark	25 dB/Bark	0,3	0,05	15

The masking threshold for stimuli $P_{y,m}(k_b)$ is approximated by

$$MT_m(k_b)=10^{\log_{10}D_m(k_b)-\frac{O_m(k_b)}{10}} \tag{31}$$

where offset is done by:

$$\begin{aligned}&\alpha_{ton}(m)=-0.275(m)-15.025 \quad [\text{dB}]\\ &\alpha_{szum}(m)=-9.0 \quad [\text{dB}] \qquad\qquad 0\leqslant m\leqslant 17\\ &O_m(k_b)=t_m(k_b)\cdot\alpha_{ton}(m)+(1-t_m(k_b))\cdot\alpha_{szum}(m) \quad [\text{dB}]\end{aligned} \tag{32}$$

and the tonlike or noiselike nature of the signal is determined by the statistical characteristic of the power spectrum and mathematically given by:

$$t_m(k_b)=min\left(\frac{SFM_m(k_b)}{SFM_{max}},1\right),0\leqslant m\leqslant 17 \tag{33}$$

where $SFM_{max}=-60\,\text{dB}$

$$SFM_m(k_b)=10\log_{10}\left(\frac{\left[\prod_{i=1}^{W}|y_m(W\cdot k_b+i)|^2\right]^{\frac{1}{W}}}{\frac{1}{W}\sum_{i=1}^{i=W}|y_m(W\cdot k_b+i)|^2}\right),\quad k_b\in\mathbb{N} \tag{34}$$

[dB]

The example of the proposed masking threshold calculation is depicted in Fig. 8.

Appendix B

In single microphone system only the noise corrupted signal is available. Therefore, all power estimates must be based on this signal. In this situation only branch I and II of (9) are engaged in noise reduction process. Assuming:

$$P_{y,m}(k_b)=P_{s,m}(k_b)+P_{n,m}(k_b)\ , \tag{35}$$

$$P_{\hat{s},m}(k_b)=G_m(k)P_{y,m}(k_b)\ , \tag{36}$$

and taking into account Eq. (9), the enhancement rule is stated as:

$$\begin{cases}G_m(k_b)P_{y,m}(k_b)-P_{s,m}(k_b)\leqslant 0 & if \quad P_{y,m}(k_b)\geqslant T_m(k_b)\text{and } P_{s,m}(k_b)\geqslant T_m(k_b)\ (I)\\ G_m(k_b)P_{y,m}(k_b)-T_m(k_b)\leqslant 0 & if \quad P_{y,m}(k_b)\geqslant T_m(k_b)\text{and } P_{s,m}(k_b)\leqslant T_m(k_b)\ (II)\end{cases}. \tag{37}$$

Solving (37) for $a_m(k_b)$

$$\begin{cases}a_m(k)\geqslant P_{y,m}(k_b)\left[\frac{P_{y,m}(k_b)}{P_{s,m}(k_b)}-1\right]^{\frac{1}{v_m}}, & if \quad P_{y,m}(k_b)\geqslant T_m(k_b)\text{and } P_{s,m}(k_b)\geqslant T_m(k_b)\ (I)\\ a_m(k)\geqslant P_{y,m}(k_b)\left[\frac{P_{y,m}(k_b)}{T_m(k_b)}-1\right]^{\frac{1}{v_m}}, & if \quad P_{y,m}(k_b)\geqslant T_m(k_b)\text{and } P_{s,m}(k_b)\leqslant T_m(k_b)\ (II)\end{cases} \tag{38}$$

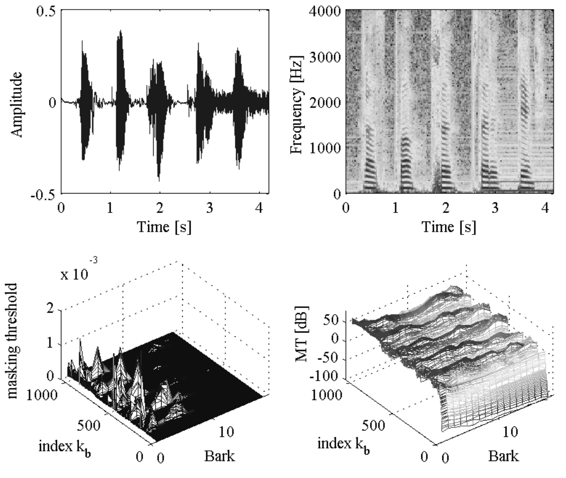

Fig. 9. An example of estimation masking threshold, upper the male voice counting "1-5" time plot and spectrogram, below the calculated masking threshold left – linear scale and adequately scaled dB representation-right

Using (35) and skipping the inequality in order to take the minimum estimate in place of $P_{s,m}(k_b)$ from I of (38) the $T_m(k_b)$ is taken, where $T_m(k_b)\leqslant P_{s,m}(k_b)$ for $0<v_m\leqslant 1$ the

$$[T_m(k_b)+P_{n,m}(k_b)]\left[\frac{P_{n,m}(k_b)}{T_m(k_b)}\right]^{\frac{1}{v_m}} \geqslant [P_{s,m}(k_b)+P_{n,m}(k_b)]\left[\frac{P_{n,m}(k_b)}{P_{s,m}(k_b)}\right]^{\frac{1}{v_m}} \tag{39}$$

and this fulfils the branch I of (38).

From equation II of (35) it is clear that $a_m(k_b)$ is proportional to $P_{s,m}(k_b)$ therefore replacement by $T_m(k_b)$ also fulfils condition of branch II (38).

The grater estimation of the coefficient $a_m(k_b)$ in equation

(11) gives the over estimated psychoacoustic weighting rule for constant v_m, so the scaling can be done by the value of v_m.

References

[1] Boll S.F., *Suppression of acoustic noise in speech using spectral subtraction*, IEEE Transactions on Acoustic Speech and Signal Processing, vol. 27, no 2, 1979, pp. 113-120

[2] Berouti M., Schwartz R., Makhoul J., *Enhancement of speech corrupted by acoustic noise*, In Proc. International Conf. on Acoustic, Speech and Signal Processing, April, 1979, pp. 208-211

[3] Ephraim Y., Malah D., *Speech enhancement using a minimum mean-square error short-time spectral amplitude estimator*, IEEE Transactions on Acoustic Speech and Signal Processing, vol. 32, no 6, 1984, pp.1109-1121

[4] Ephraim Y., Malah D., *Speech enhancement using a minimum mean-square log-spectral amplitude estimator*, IEEE Transactions on Acoustic Speech and Signal Processing, vol. 33, no 2, 1985, pp. 443-445

[5] Jebara S.B., Benazza-Benyahia A., Khelifa A.B., *Reduction of musical noise generated by spectral subtraction by combining wavelet packet transform and Wiener filering*, in X European Signal Processing Conference Proc., 2000, CD

[6] Gulzow T., Engelsberg A., Heute U., *Comparison of a discrete wavelet transformation and nonuniform polyphase filterbank applied to spectral-subtraction speech enhancement*, Signal Processing, vol. 64, 1998, pp.5-19

[7] Dreiseitel P., Puder H., *Speech Enhancement For Mobile Telephony Based on Non-Uniformly Spaced Frequency Resolution*, in Proc. IX European Signal Processing Conference, 1998, pp. 965-968

[8] Tsoukalas D.E., Mourjopoulos J.N., Kokkinakis G., *Improving the intelligibillity of noisy speech using an audible noise suppresion technique*, in 5th European Conference on Speech Communication and Technology Proc.,1997, pp. 1415-1418

[9] Gustafsson S., Jax P., Vary P., A *Novel Psychoacoustically Motivated Audio Enhancement Algorithm Preserving Background Noise Characteristic,* In Proc. International Conf. on Acoustic, Speech and Signal Processing, 1998, CD

[10] Czyżewski A., Królikowski R., *Noise reduction Algorithms employing an intelligent inference engine for multimedia applications*, IEEE Workshop on Multimedia Signal Processing, Dec. 1998, CD

[11] Haulick T., Linhard K., Schrogmeier P., *Residual Noise Suppression Using Psychoacoustic Criteria*, in 5th European Conference on Speech Communication and Technology Proc., 1997, pp. 1395-1398

[12] Cappe O., *Elimination of the musical noise phenomenon with the Ephraim and Malah noise suppressor*, IEEE Transactions on Acoustic Speech and Signal Processing, vol. 2, 1994, pp. 345-349

[13] Beaugeant C., Scalart P., *Noise reduction using perceptual spectral change*, in 6th European Conference on Speech Communication and Technology Proc.,Sept. 1999, pp. 2543-2546

[14] Quatieri T.F., Baxter R.A, *Noise reduction Based on spectral change*, IEEE ASSP Workshop on Application of Signal processing to Audio and Acoustics, 1997, CD

[15] Virag N., *Single Channel Speech Enhancement Based on Masking Properties of the Human Auditory System*, IEEE Transactions on Speech and Audio Processing, vol. 7, no 2. 1999, pp. 126-137

[16] Johnston J. D., *Transform Coding of Audio Signals Using Perceptual Noise Criteria*, IEEE Journal on Selected Areas in Communications, vol. 6, no 2, 1988, pp. 314-323

[17] Crochiere R.E., Rabiner L.R., *Multirate Digital Signal Processing*, Prentice Hall, 1983

[18] Strang G., Nguyen T., *Wavelets and Filter Banks*, Wellesley-Cambridge Press, 1996

[19] Vaidyanathan P.P., *Multirate Systems and Filter Banks*, PTR Prentice Hall, 1993

[20] Vary P., Heute U., Hess W., *Digitale Sprachsignalverarbeitung*, B.G. Teubner Stuttgart, 1998

[21] Akansu A.N., Haddad R.A., *Multiresolution signal Decomposition: Transforms, Subbands, and Wavelets,* Academic Press, INC., 1992

[22] Constantinides A.G., *Spectral transformations for digital filters*, Proc. IEE, vol. 117, 1970, pp. 1585-1590

[23] Oppenheim A.V., Schafer R.W., *Cyfrowe przetwarzanie sygnałów*, WKiŁ, 1979

[24] Oppenheim A., Johnson D., Steiglitz K., *Computation of Spectra with Unequal Resolution Using Fast Fourier Transform*, Proceedings of the IEEE, vol. 59, no 2, 1971, pp. 299-301

[25] Strube H.W., *Linear prediction on warped frequency scale*, J. Acoustical Soc. of America, vol.68, no 4, 1980, pp. 1071-1076

[26] Smith III J.O. , Abel J.S., *Bark and ERB Bilinear Transforms*, IEEE Transactions on Speech and Audio Processing, vol. 7, no 6, 1999, pp. 697-708

[27] Doblinger G., *Computationally Efficient Speech Enhancement by Spectral Minima Tracking in Subbands*, in 4th European Conference on Speech Communication and Technology Proc., Madrid, Spain,18-21 Sept. 1995, pp. 1513-1516

[28] Quackenbush S.R.,et al. *Objective measures of speech Quality* Printice Hall, Engelwood Clifs, New Jersey, 1988

[29] Heute U., *Objektive Sprachqualitatsmessungen: Vergleichende Ubersicht und ein neuer Ansatz*, In Proc. 8 Aachener Kolloquium Signaltheorie, VDE-Verlag, Berlin, 1994, pp.21-28