Parfieniuk M., Petrovsky A. Tunable non-uniform filter bank mixing cosine modulation with perceptual frequency warping by allpass transformation. - Automatic control and computer sciences, Allerton Press, Inc. — 2004. - №4. – C.44-52

Tunable non-uniform filter bank mixing cosine modulation with perceptual frequency warping by allpass transformation*)

M. Parfieniuk, A. Petrovsky

Department of Real Time Systems, Bialystok Technical University, Poland

address: ul.Wiejska 45 A, 15-351 Bialystok; fax: +48-85 7423423

e-mail: palex@it.org.by

ABSTRACT

A novel nonuniform cosine modulated filter bank is presented, mixing the concepts of warped DFT (Discrete Fourier Transform) filter bank and pseudo-QMF system (Quadrature Mirror Filter). Unequal bandwidths are achieved by allpass transformation of uniform polyphase system. The cosine modulation results in real valued channel signals, and hence simplified subband processing. The considered filter bank can well approximate popular psychoacoustic scales, and moreover dynamically change its frequency characteristics.

Key-words: time-frequency analysis, signal representations, filter banks, multirate processing, allpass transformation, nonuniform cosine modulated filter bank, perceptual frequency warping

---

*) This work was supported by Białystok Technical University under the grant W/WI/3/02

1. INTRODUCTION

Modern acoustic signal processing tends to exploit maximally the weaknesses of human auditory system. The existence of the critical bands of hearing, corresponding to the nonlinear distribution of ear sensitivity over frequencies, suggests to process signals in unequal subbands. Known approaches utilize so-called auditory filter banks designed to obtain unequal bandwidths approximating the resolution of hearing.

Many different kinds of non-uniform filter banks were proposed. One broad class of solutions is based on a direct utilization of pure uniform filter banks. Unequal bandwidths can then be obtained by merging uniform channels [1] or by cascading uniform systems to form a tree / pyramid scheme [2, 3]. Although these approaches are conceptually clear, the theory suffers from some lacks [3]. The main drawback is in a small flexibility of bandwidth shaping. Firstly, the underlying uniform decomposition fixes a certain regular grid constraining unequal bandwidths derived from it. Moreover, to approximate a more complicated non-uniformity, the grid must be disintegrated. This can be done at the cost of higher complexity, by increasing the number of the channels or attaching additional stages to the hierarchy. As the bandwidth distribution is tightly connected to the system structure, there is no way to change it, even slightly, without the modification of the entire schema.

A more novel approach is the concept of allpass transformed filter banks, first introduced in [4]. This is a direct extension to the case of filter banks, the idea of the deformation of frequency scale. This concept was already considered in the seventies [5] and became the base of the warped signal processing [6] in the nineties. The mathematical basis of warped DFT (Discrete Fourier Transform) filter bank is still under development. For example the perfect reconstruction (PR) constraints for the non-subsampled case were formalized recently [7]. Despite the imperfections of the theory, there were successful practical applications in speech enhancement [8, 9]. These exhibited one significant disadvantage of warped DFT approach. As the result of used modulation, the channel signals are complex. So the subband algorithms, such as high order adaptive filtering for echo control (see Fig. 1), dealing with complex sequences, are highly complicated.

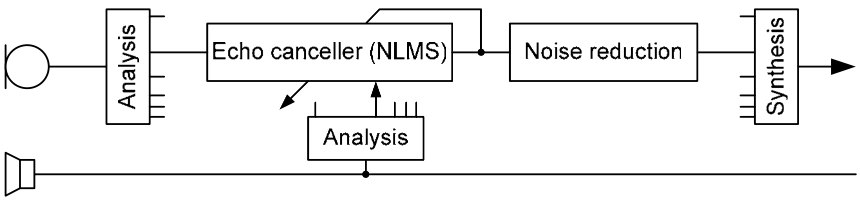


Fig. 1. Subband processing in speech enhancement system with warped filter bank

To solve this problem, this paper extends the idea of the warping to the case of well known pseudo-QMF (Quadrature Mirror Filter) filter bank. Owing to cosine modulation, the channel signals are real. The complexity of subband decomposition grows in comparison to the DFT, but that of the overall system, with advanced in-channel processing, decreases. The design can be performed easily basing on one prototype filter, the same on the analysis and synthesis sides. The most attractive properties of warped system are also kept. The allpass transformation allows the flexible shaping of bandwidth distribution, especially to approximate well the critical bands. Moreover this distribution can be changed in some range, even in working system, to adjust its characteristics to current user or environment.

2. DETAILS OF CONSIDERED FILTER BANK

2.1. The construction of the analysis and synthesis stage

Let it is given a prototype *N*th order FIR (Finite Impulse Response) filter of transfer function

$$P(z)=\sum_{n=0}^{N} p(n)z^{-n}, \tag{1}$$

real coefficients $p(n)$, linear phase and cut-off frequency $\frac{\pi}{2K}$ for *K*-channel filter bank. By cosine modulation according to the formula

$$\hat{H}_k(z)=a_k b_k P\left(zW_{2K}^{(k+0.5)}\right)+a_k^* b_k^* P\left(zW_{2K}^{-(k+0.5)}\right) \tag{2}$$

common for pseudo-QMF banks [2], where

$$a_k=e^{j\theta_k} \quad b_k=W_{2K}^{(k+0.5)\frac{N}{2}} \quad W_{2K}=e^{-j\frac{\pi}{K}} \quad \theta_k=(-1)^k\frac{\pi}{4} \quad k=0,\ldots,K-1 \tag{3}$$

and superscript asterisk denotes complex conjugation, the shifted versions of prototype frequency response are obtained. Resulting uniform bandwidth distribution can then be warped by all-pass transformation. It consists in the replacing all of delays with identical first order allpass filters

$$z^{-1} \quad \Rightarrow \quad A(z) = \frac{z^{-1} - a}{1 - az^{-1}}, \qquad |a| < 1 \tag{4}$$

is real coefficient, causal and stable. The corresponding frequency response has the form

$$A(z) = e^{j\phi(\omega)} \qquad \phi(\omega) = -\omega + 2\arctan\frac{a\sin\omega}{a\cos\omega - 1}. \tag{5}$$

Its magnitude is obviously unity and the phase is the nonlinear function of frequency. The transformed analysis filters take the form

$$H_k(z) = a_k b_k P\left(A(z)^{-1} W_{2K}^{(k+0.5)}\right) + a_k^* b_k^* P\left(A(z)^{-1} W_{2K}^{-(k+0.5)}\right). \tag{6}$$

The synthesis filters are built in a similar manner to get

$$F_k(z) = a_k^* b_k P\left(A(z)^{-1} W_{2K}^{(k+0.5)}\right) + a_k b_k^* P\left(A(z)^{-1} W_{2K}^{-(k+0.5)}\right). \tag{7}$$

Notice that achieved filter banks, now involving recursive transfer function, are IIR (Infinite Impulse Response) systems. It should also be noted that the simple alteration of a value can completely change the shapes of $H_k(z)$ and $F_k(z)$. This idea will be revealed in Section 4 addressing the real time tuning.

2.2. Remarks on the implementation

The final structure of the analysis / synthesis system is shown on fig.2. Cosine modulation is done simply with the help of a double sized FFT (Fast Fourier Transform). As we utilize system only with real signals, the halves of the FFT outputs are reciprocal conjugates. So the real part of only one (post-modulated) half can simply be taken as the output of analysis bank. Any of the known fast schemes of cosine modulation can be used as an alternative. The schema contains decimators and interpolators allowing the reduction of sampling rate in the channels. The chain of allpass filters, as IIR system, must operate at original sampling frequency on both analysis and synthesis sides. The rest of the structure works with decimated signals at reduced rate. Decimation factor R is bounded by the widest bandwidth and is the same in all subbands.

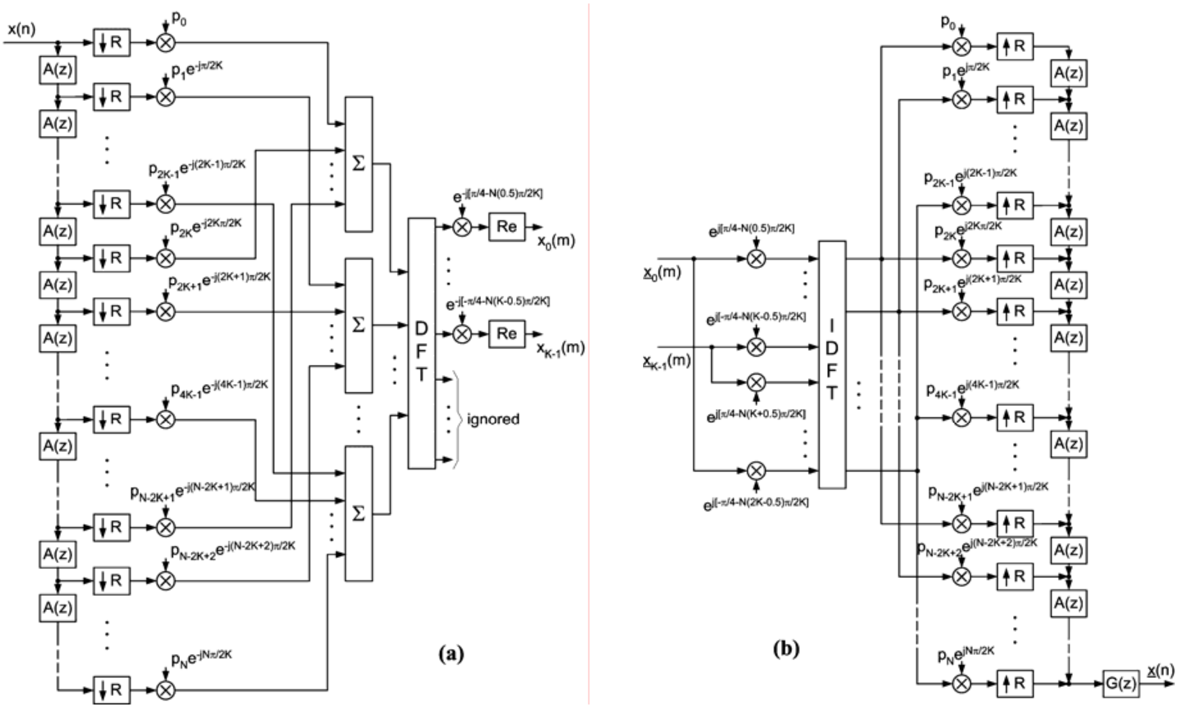


Fig. 2. Overall structure of non-uniform cosine modulated filter bank - analysis (a) and synthesis (b) parts

As the bandwidths are unequal, the channels may be subsampled differently. In this approach the decimators (each with distinct factor $R_k, k = 0...K-1$) must be placed at the output of analysis stage and the interpolators before the synthesis one. The post-filter $G(z)$ reduces nonlinear phase distortion related to all-pass chain. Its role and possible structure will be explained in the next subsection.

Computational complexity (8a) growths in comparison to that (8b) of DFT system considered in [8]. The increase is similar as comparing classical uniform DFT bank with cosine modulated one

$$\underbrace{2(L-1)}_{c_1}+\underbrace{\frac{2L}{R}}_{c_2}+\underbrace{\frac{8K\log_2 2K}{R}}_{c_3}+\underbrace{\frac{2K}{R}}_{c_4} \quad \text{(8a)} \quad \underbrace{2(L-1)}_{c_1}+\underbrace{\frac{L}{R}}_{c_2}+\underbrace{\frac{4K\log_2 K}{R}}_{c_3}, \quad \text{(8b)}$$

where c1 i– allpass chain, c2 – polyphase filtering and pre-modulation, c3 – modulation with FFT, c4 – post-modulation. But more essential complexity of advanced subband processing, like that on Fig. 2, decreases significantly (roughly 4 times – taking simply the relation between complex and real multiplication).

2.3. The design aimed at exact reconstruction

The z-transform of the output of combined analysis / synthesis system can be expressed as

$$Y(z)=T_0(z)X(z)+\sum_{r=1}^{R-1}X\left(zW_R^r\right)T_r(z)\ , \qquad (9)$$

where

$$\begin{aligned} T_0(z)&=\frac{1}{R}\sum_{k=0}^{K-1}H_k(z)F_k(z) && \text{distortion function} \\ T_r(z)&=\frac{1}{R}\sum_{k=0}^{K-1}H_k\left(zW_R^r\right)F_k(z) && \text{aliasing transfer function} \end{aligned} \qquad (10)$$

As the aliasing terms are approximately canceled because of appropriate selection of factors (3) and high stopband attenuation of prototype [2], we can focus on $T_0(z)$. Following the derivations in [2], generalizing them to the case of warping, it can be shown that

$$T_0(z) = \frac{A(z)^N}{R}\sum_{k=0}^{K-1}\left|H_k(z)\right|^2 = e^{j\phi(\omega)N}\frac{1}{R}\sum_{k=0}^{K-1}\left|H_k(z)\right|^2 . \tag{11}$$

As in conventional pseudo QMF system, transfer function suffers from amplitude distortion which can be reduced by the proper design of prototype. Besides that, the problem with nonlinear phase arises now, caused by the existence of the all-pass chain. Two approaches can be employed here. Both are based on the idea of post-filtering, where the reconstructed signal is then passed by FIR filter designed to equalize deformed phase. The construction of such a filter can be performed in the following ways.

One, more empirical, method consists in the generating the impulse response of the overall system. The resulting sequence is then truncated and reversed in time to form the set of post-filter coefficients. For long allpass chain and for the larger absolute values of warping parameter a, the impulse response elongates, so post-filter order increases (with related computational and memory requirements). The second approach utilizes the fact [7] that the FIR transfer function of the form

$$F(z) = -az^{-d} + \left(1-a^2\right)z^{-(d-1)} + a\left(1-a^2\right)z^{-(d-2)} + \ldots + a^{d-2}\left(1-a^2\right)z^{-1} + a^{d-1} \tag{12}$$

can approximately compensate the phase distortion caused by first order allpass, owing to

$$F(z)A(z) = z^{-d} - a^{d}. \tag{13}$$

As d increases, whereas $|a|<1$, the second term decays, so the product becomes pure delay.

The major drawback is again in large delay and high computational load. The chain of N allpasses requires the very long post-filter built by the cascading of the same number of equalizing filters, what results in equivalent transfer function $F^N(z)$. Moreover, the order d of $F(z)$ reducing error to acceptable level can be very high, especially when $|a|>0.5$.

3. APPROXIMATION OF PSYCHOACOUSTIC SCALES

The design of warped filter bank approximating psychoacoustic scale requires considering two problems. First, the number of subbands must be fixed according to approximated scale and given sampling frequency. The second trouble is to adjust warping. It must be done in such a way that the band edges will comply with selected scale. This subject was extensively studied in [10]. Here, we only mention the most important facts.

The all-pass function can be interpreted mathematically as conformal bilinear mapping in the z-plane. It maps the unit circle to itself in such a way that:

- points e^{j0} (d.c. term) and $e^{j\pi}$ (Nyquist frequency) remain unchanged,
- points $e^{j\omega}, 0<\omega<\pi$, representing particular frequencies, move to new locations $e^{j\phi(\omega)}$ (still $0<\phi(\omega)<\pi$)

The mapping has only one degree of freedom, corresponding to adjustable parameter a. However this is sufficient to approximate well psychoacoustic scales for sampling frequencies 4 - 48 kHz, commonly used in audio applications.

The value of transform parameter can be established every design time, due to circumstances, in optimization procedure based on preferred error measure. But, basing on numerical results from optimization in Chebyshev sense and the same function template, analytical, so-called “arctan”, formulas were proposed in [10]. This for the Bark scale has the form

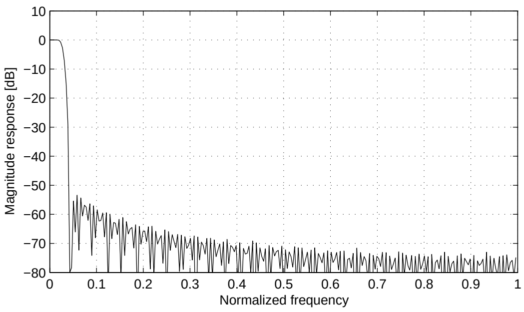
$$a_{Bark} = 0.1957 - 1.048\left[\frac{2}{\pi}\arctan\left(0.07212\frac{f_s}{1000}\right)\right]^{\frac{1}{2}} \tag{14}$$

and that for the ERB (Equivalent Rectangular Bandwidth) is

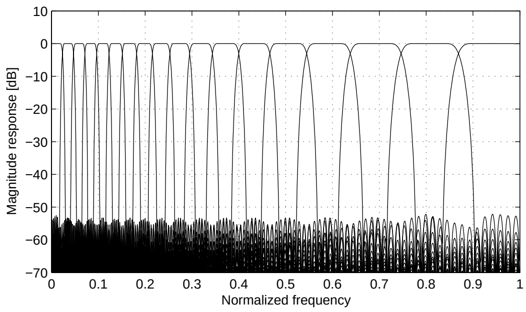
$$a_{ERB} = -0.7164\left[\frac{2}{\pi}\arctan\left(0.09669\frac{f_s}{1000}\right)\right]^{\frac{1}{2}} - 0.08667 \tag{15}$$

Both take sampling frequency in Hz as parameter.

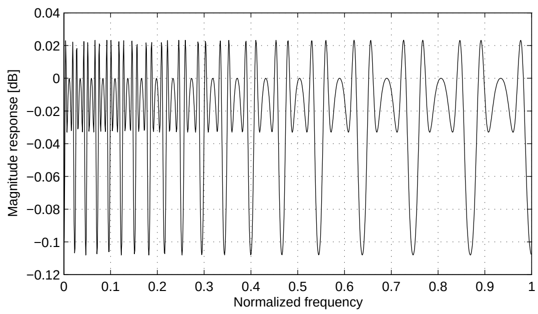
As an example, let’s consider the design of the filter bank approximating the Bark scale for system combining noise reduction and echo control [11]. The field of application is telephony where standard bandwidth is 4kHz comprising 18 critical band of hearing. Thus we need the decomposition with 18 subbands. The proper parameter of allpass transformation is given by (14) as $a_{Bark} = 0.4092$. Prototype filter has order of 288 and is designed in optimization procedure simultaneously minimizing the amplitude ripples of transfer function and maximizing the attenuation in the stopband. Its response and characteristics of resulting system for $R = 1$ are shown in Fig. 3.



(a)



(b)



(c)

Fig. 3. Magnitude responses of designed prototype (a), analysis filter bank (b) and analysis / synthesis system (c)

4. TUNING UP OF THE BANDWIDTHS

The psychoacoustic scales are only generalizations. As the perceptual abilities of different listeners significantly vary, the Barks or the ERBs can be only the better or worse approach of the actual perception of particular individual. The error can reach several dozen percents. Experiments show that the adjustment of subband decomposition to the personal preferences of listeners often benefits subjective quality. It seems attractive to give the user the possibility of the additional interactive tuning of the computer program or standalone device realizing some form of subband audio processing. The system with allpass transformation can satisfy above postulate as the simple modification of warping function coefficient can drastically change the deformation of frequency response. It is obvious that large deviation heavily affects the overall system and deeper reconfiguration is necessary. But limiting the all-pass coefficient variation to reasonable bounds, bandwidths can be tuned up without disturbing the other parts of schema.

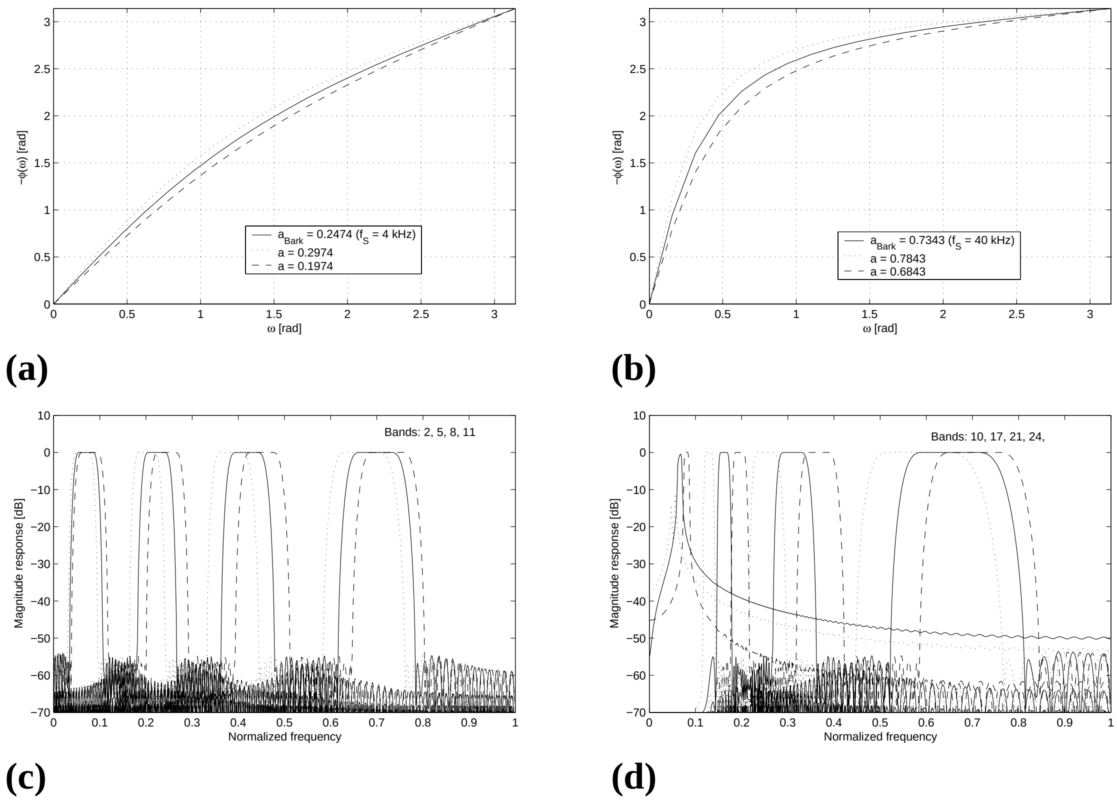


Fig. 4. Deviation of all-pass phase response and bandwidths corresponding to changes in warping parameter

From the experiments with the subband speech enhancement system mentioned previously, it was concluded that acceptable change of a is ±0.05. In this case the deviation of impulse response is moderate, so FIR post-filter can remain unchanged, nevertheless giving good phase equalization. The corresponding shifts in the center frequencies of band are of 10 - 100% their bandwidths, which simultaneously squeeze / spread of 5 - 10%. This is depicted in Fig. 4 for two ranges of a value – the one closer to zero and the second near to unity for comparison.

5. CONCLUSIONS AND FUTURE WORKS

The foundations of the novel structure of nonuniform cosine modulated filter bank have been drawn. It seems attractive for audio applications because of rather easy and flexible design oriented to the approximation of psychoacoustic scales. We don't judge which of the above solutions is superior - we prefer more pragmatic approach tightly connecting the design decisions with the practical use of filter bank and other circumstances and tradeoffs related to its implementation. It should be noted here that in many applications exact reconstruction is not obligatory. Small distortions related to the filter bank doing subband decomposition can often be neglected as the errors of the main in-channel processing (e.g. estimation) are more significant.

The reality of channel signals gives computational profits, with the respect of warped DFT banks. The different aspects of the idea need further improvement and investigation. The case of subsampling requires more detailed consideration. Efficient phase equalization also is up-to-date problem. The existing subband processing algorithms are oriented rather on regular bandwidth distribution with critical sampling – deeper adaptation to the possibilities (and limitations) offered by the warped filter bank is needed. In spite of all these shortcomings, proposed solution is promising and worth interesting.

REFERENCES

[1] F. Argenti, E. Del Re, “Non-Uniform Filter Banks Based on a Multi-Prototype Cosine Modulation”, Proc. IEEE International Conference on Acoustic, Speech and Signal Processing, ICASSP’96, Vol. 3, Atlanta, Georgia, 7 - 9 May 1996, pp. 1511 - 1514

[2] P. P. Vaidyanathan, “Multirate Systems and Filter Banks”, PTR Prentice Hall, Englewood Cliffs, New Jersey, 1993

[3] S. Akkarakaran, P.P. Vaidyanathan, „New results and open problems on nonuniform filter-banks”, Proc. IEEE International Conference on Acoustic, Speech and Signal Processing, ICASSP’99, Phoenix, AZ, USA, 15 - 19 May 1999, pp. 1501 - 1504

[4] M. Kappellan, B. Strauss, P. Vary, „Flexible nonuniform filterbanks using allpass transformation of multiple order“, Proc. European Signal Processing Conf., EUSIPCO’96, Vol. 3, Trieste, Italy, September 1996, pp. 1745 - 1748

[5] A Oppenheim, D. Johnson, K. Steiglitz, “Computation of spectra with unequal resolution using Fast Fourier Transform”, Proc. of IEEE, Vol. 59, No 2, 1971, pp. 299 - 301

[6] A. Härma, M. Karjalainen, L. Savioja, V. Välimäki, U.K. Laine, A.J. Huopaniemi, “Frequency-Warped Signal Processing for Audio Applications”, J. Audio Eng. Soc., Vol. 48, No 11, November 2000, pp. 1011 - 1030

[7] E. Galijasevic, J. Kliewer, “Non-uniform near-perfect-reconstruction oversampled DFT filter banks based on allpass-transforms”, Proc. 9th IEEE DSP Workshop (DSP 2000), Hunt, TX, USA, October 2000

[8] T. Guzlow, A. Engelsberg, U. Heute, “Comparison of a discrete wavelet transformation and a nonuniform polyphase filterbank applied to spectral-subtraction speech enhancement”, Signal Processing, Vol. 64, pp. 5 - 19

[9] K. Bielawski, A.A. Petrovsky, “Speech enhancement system for hands-free telephone based on the psychoacoustically motivated filter bank with allpass frequency transformation”, Proc. 6th European Conf. on Speech Communication and Technology, EUROSPEECH’99, Budapest, Hungary, September 1999, pp. 2555 - 2558

[10] J.O. Smith, J.S. Abel, “Bark and ERB Bilinear Transforms”, IEEE Trans. of Speech and Audio Processing, Vol. 7, No. 6, 1999, pp. 697 - 708

[11] A. Petrovsky, K. Bielawski, “Auditory model based speech enhancement system for hands-free devices”, Proc. EUSIPCO 2002, Vol. I , Toulouse, France, 2002, pp. 487 – 490

Address of authors:

Marek Parfieniuk, assistant, is with the Department of Real Time Systems, Faculty of Computer Science, Bialystok Technical University, 45A Wiejska Str., 15-351, Bialystok, Poland.

Alexander Petrovsky, professor, Dr.Sc., is with the Department of Computer Engineering, Faculty of Computer Aided Design, Belarusian State University of Informatics and Radioelectronics, Belarus and with the Department of Real Time Systems, Faculty of Computer Science, Bialystok Technical University, 45A Wiejska Str., 15-351, Bialystok, Poland.

fax: +48-85 7423423

e-mail: palex@it.org.by