LOW COMPUTATIONAL COST EQUALIZATION AND MODELING OF AUDIO SYSTEMS

PACS REFERENCE 43.60.Gk

Ramos, Germán[1]; Bank, Balázs[2]

(1) ITACA Institute, Universitat Politècnica de València

Camino de Vera s/n

Valencia – 46022

Spain

Tel: +34 963879604

e-mail: gramosp@eln.upv.es

(2) Department of Measurement and Information Systems, Budapest University of Technology and Economics

Budapest

Hungary

e-mail: bank@mit.bme.hu

ABSTRACT

The equalization and modeling of audio systems have some particular features that are not present in other science fields, including a frequency range of 10 octaves. Also, the properties of the human auditory system (e.g. logarithmic resolution both in frequency and magnitude) should be taken into account. If the design is done using conventional filter design approaches having linear frequency resolution, this results in filters requiring huge computational cost for a full-band solution. This paper presents several design methodologies that take into account the particular requirements for audio, in order to achieve solutions with better subjective quality and lower computational cost. The paper presents parametric optimization, warped filters and their combination with linear systems, and the use and design of fixed-pole parallel filter banks.

RESUMEN

La ecualización y modelado de sistemas de audio presenta ciertas particularidades no presentes en otros campos de la ciencia, como es el abarcar un rango de 10 octavas. También, la naturaleza de la percepción humana del sonido es logarítmica tanto en frecuencia como en magnitud, aspectos ambos a tener en cuenta. Si su diseño se enfoca con técnicas de diseño de filtros convencionales con una resolución en frecuencia lineal, son necesarios filtros de alto coste computacional para cubrir toda la banda audible. En este artículo se presentan nuevas metodologías de diseño que tienen más en cuenta la naturaleza del problema para obtener soluciones de mejor calidad subjetiva y menor coste computacional: optimización paramétrica, filtrado *warped* y su combinación con filtros lineales, y uso de bancos de filtros paralelos.

1. INTRODUCTION

The use of digital filters for equalization and modeling audio systems like loudspeakers, head-related-transfer-functions (HRTF), room equalizers, crossover filters, noise canceling systems, or stereo enhancement, is nowadays widely used. Fig. 1 shows the scenario of a digital filter employed for equalizing a loudspeaker $H_{lspk}(\omega)$ inside a listening room with an acoustical path between the loudspeaker and the listener modeled as $H_{room}(\omega)$. For modeling applications, the scenario is similar, but the modeling filter must approximate to the desired target.

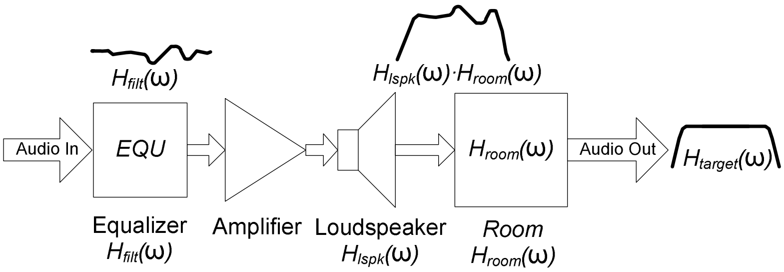


Fig. 1 – Equalization scenario

The filter $H_{filt}(\omega)$ must compensate the combined response of the loudspeaker and the room at the listening point to approximate to the desired target response $H_{target}(\omega)$. This solution is non-trivial, mainly due to $H_{room}(\omega)$ is a non-minimum phase system, and hence its inverse is unstable. As a result $H_{filt}(\omega)$ could only be approximated. For loudspeaker equalization, $H_{target}(\omega)$ must respect the natural band-pass behavior of the loudspeaker, avoiding the selection of flat full-band target responses. Several suggestions about the proper target response selection and preprocessing techniques are discussed in [1].

$$H_{filt}(\omega) = \frac{H_{target}(\omega)}{H_{lspk}(\omega) \cdot H_{room}(\omega)} \tag{1}$$

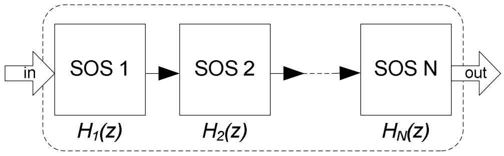
There are several classical digital filter design methods for designing $H_{filt}(\omega)$. These methods cover the design of FIR (finite impulse response) and IIR (infinite impulse response) filters. Their design could be carried out in the time-domain or in the frequency domain, but there is a common approach in these classical methodologies: they are inherently linear. This means that the problem is shown from a linear point of view, having a linear frequency resolution, and weighting the error also linearly. But in audio applications this does not happens due to:

- Audio bandwidth is relatively low (20 kHz) compared with other fields, but it covers almost 10 octaves. This does not happen in other fields.
- Wavelengths cover from millimeters to meters, having similar dimensions respect to the room size and environment objects. All the wave effects must be considered: reflection, diffraction, interference, diffusion.
- The receptor is a human being, which has more logarithmic than linear behavior, either in the frequency axis as in the magnitude axis.
- The transducers (mainly electrostatic loudspeakers) are far from being ideal, with non-uniform frequency response, non-uniform directivity pattern, and non-linearities. These last two issues are beyond the scope of this paper.

This paper collects some approaches that the authors have developed for designing low computational cost (low order) filters for audio applications. This involves taking into account the logarithmic behavior of the problem, and using specific cost functions and strategies that minimize the error from a psico-acoustic point of view, no form a classical and linear approach.

2. PARAMETRIC OPTIMIZATION OF IIR FILTERS

Ramos proposed in [1] and evolved [2] a method for designing IIR filters for loudspeaker equalization and crossover alignments. It's based on a parametric approach instead on a mathematical approach optimizing directly the numerical values of the coefficients. The digital equalization filter $H_{filt}(\omega)$ is directly designed as a second-order-section (SOS) chain (as it will be latter implemented in a DSP or microprocessor) as show in Fig. 2.



$H_{filt}(z)=H_1(z)\cdot H_2(z)\cdots H_N(z)$

Fig. 2 - Equalizer structure

Each SOS in the chain is forced to be a specific second order filter type (lowpass, highpass, shelving low and high, and peak filter) [3] all of them defined by their parameters: central frequency f_i, gain in decibels GdB_i, and quality factor Q_i. These parameters have physical meaning, being easier to interpret than the coefficients values. The coefficients' values can be calculated form their parameters using the bilinear transformation [4] or using the direct formulas found in [3] or [5]. Following the equalization example of Fig. 1, the $H_{filt}(\omega)$ must be designed minimizing the error $e(\omega)$ of Eq. 2, where $H_{lr}(\omega)$ is the combined effect of the loudspeaker and room responses $H_{lspk}(\omega)\cdot H_{room}(\omega)$ at the listening point.

$$e(\omega) = H_{target}(\omega) - H_{lr}(\omega)\cdot H_{filt}(\omega) \qquad (2)$$

As commented in the Introduction, we need to focus an audio problem using a more logarithmic than linear approach. For that, and using psico-acoustic criteria, the error will be evaluated over a discrete frequency axis logarithmically spaced (1/48th octave) $\omega_{log}=[\omega_0, \omega_1, \ldots, \omega_{N-1}]$, and the magnitudes will be evaluated in decibels, indicated as [dB]. The discrete error vector is

$$e(\omega_k)_{[dB]} = H_{target}(\omega_k)_{[dB]} - H_{lr}(\omega_k)_{[dB]} - H_{filt}(\omega_k)_{[dB]}\,. \qquad (3)$$

Our cost function to minimize is J of Eq. 4, where n_i and n_f are the initial and end indexes on ω_{log} with the initial and end frequencies where to perform the equalization. The optional vector $W(\omega_k)$ is a weighting vector that allows to emphasize, or even avoiding, the equalization over the selected frequency regions. J is a specific cost function for audio applications: works in a double logarithmic domain (frequency and magnitude), with a L1 norm (more related with human hearing) instead of a linear and L2 norm traditional cost functions.

$$J = \frac{1}{n_f - n_i + 1}\cdot\sum_{k=n_i}^{n_f}\left|W(\omega_k)\cdot e(\omega_k)_{(dB)}\right| \qquad (4)$$

The procedure of designing the whole filter $H_{filt}(\omega)$ will be explained briefly through a loudspeaker equalization example with a very irregular response (Fig. 3a). For a detailed description, refer to [1] and [2] where also the automatic cross-over alignment is commented. The design of $H_{filt}(\omega)$ is divided in two steps that are repeated for each of the SOS in the chain. The initial parameters' values for the first SOS (a peak filter) are configured as follows: the frequency f_{o1} is the mean-log frequency of the biggest error area between the target and the response (in this case, A4 of Fig. 3a); the gain G_{dB1} has the value of the error curve at that frequency; and the Q_1 is initialized to 1.5. The second step of the algorithm is a constrained

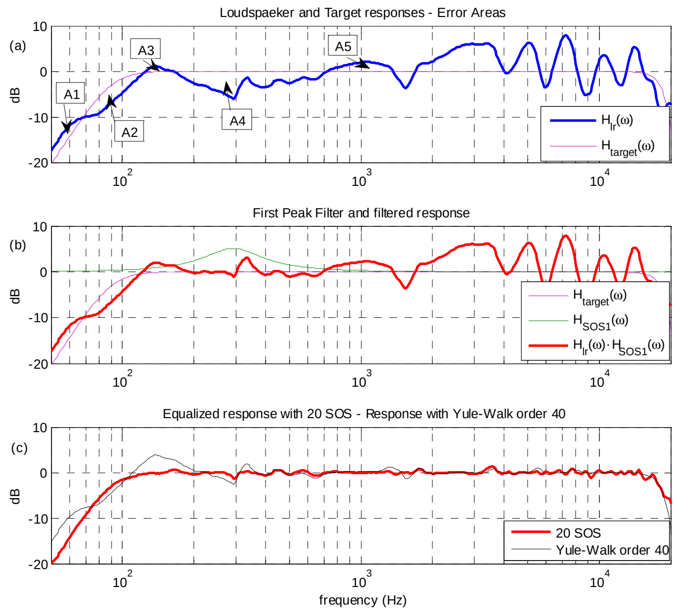


Fig. 3 – (a) Loudspeaker to be equalized, Target response and error areas. (b) Application of the first peak filter. (c) Response with 20 SOS and with Yule-Walk N=40.

heuristic optimization of the parameters looking for a better set of parameters that decreases the cost function *J*. If the new generated filter is better, it is selected as the initial filter for the next iteration, if it is worse, then it is rejected. The optimization process ends after *M* consecutive iterations without decreasing *J*. In our example, the first calculated SOS and its effect on the response is in Fig. 3b. Then the whole process is repeated but now applying the already designed SOS, till arriving at the last SOS in the chain. Finally a global post-optimization of the parameters is carried out to improve the interaction between the different SOS. They are ordered in frequency and the heuristic optimization is performed again, but now in groups of 3 frequency contiguous SOS. The achieved response with 20 SOS is in Fig. 3c that matches the target response. Normally less than 10 SOS are enough for any loudspeaker. For comparison, Fig. 3c also shows the equalized response achieved with the Yule-Walker method [6] for the same order, that has a similar behavior in mid and high frequencies, but clearly worse at low frequencies. This is due to this method makes a linear treatment of the frequency axis, and minimizes the error in this way, being difficult to find the poles for correcting the low frequencies. The same problems at low frequencies happen with other well-known filter design methods like Prony's [7] or Steiglitz-McBirde iteration [8]. In [9] a detailed comparison of methods was carried out.

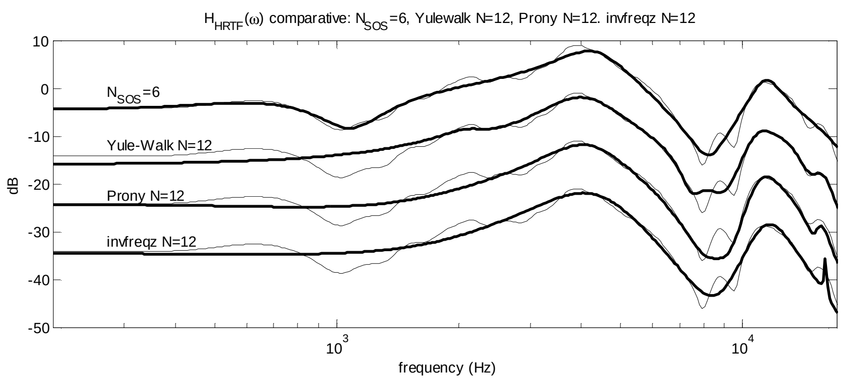


Fig. 4 - HRTF modeling with 6 SOS and comparative [10]. Responses scaled 10dB for clarity.

An example of parametric modeling of HRTF sets with a simple interpolation has been recently proposed by Ramos in [10]. Fig. 4 shows in thin line the HRTF response to be modeled. Using only 6 SOS, the achieved parametric model follows the original over the whole spectrum. The modeled reposes with Prony's method, Yule-walker, and Levi [11] (invfreqz) of the same filter order are also displayed. They work similar, or even slightly better, at high frequencies, but they are not able to model properly (to find the poles) at low frequencies. In any case, the perceived quality of the proposed model is better for the same filter order.

3. WARPED AND LINEAR FILTER COMBINATIONS

One of the consequences of the linear treatment of the frequency axis is that, in general, an excellent equalization or modeling is achieved at high frequencies. In contrast, the resolution at low frequencies is too low, or very high order filters are needed. If using a FIR filter of order *N*, the frequency resolution is Δf_{FIR} of Eq. 5, where f_s is the sampling frequency used in the system.

$$\Delta f_{FIR} = \frac{f_s}{N} \qquad \Delta f_{WARP} = \Delta f_{FIR} \cdot \frac{1+\lambda^2 - 2\cdot\lambda\cdot cos(2\cdot\pi\cdot f / f_s)}{1-\lambda^2} \tag{5}$$

To face the lack of resolution at low frequencies, there are several techniques that do not use a linear frequency axis, and even, they do a more logarithmic-like treatment. The use of warped filters is one of these options. See [12] where Karjalainen made a detailed description of them in audio applications. When using warped filters, the frequency axis resolution is modified to Δf_{WARP} of Eq. 5 that is a function of the warping parameter λ, allowing increasing the resolution at low frequencies (for positive λ values) losing resolution at high frequencies (where is enough). The main drawback of using warped filters is that their implementations need more operations, normally 3 times more than a conventional linear filter.

Ramos proposed a smart combination of warped and linear filters for audio applications [13]. This combination takes the best part of both worlds. On the one hand, it uses short linear filters for excellent results at high frequencies where there is resolution enough even for low order filters. On the other hand, it selects the proper λ value to tune and maximize the frequency resolution at low frequencies, and achieves good results with low order warped filters. For the total computational cost employed, the achieved equalization or modeling is excellent over the whole frequency axis.

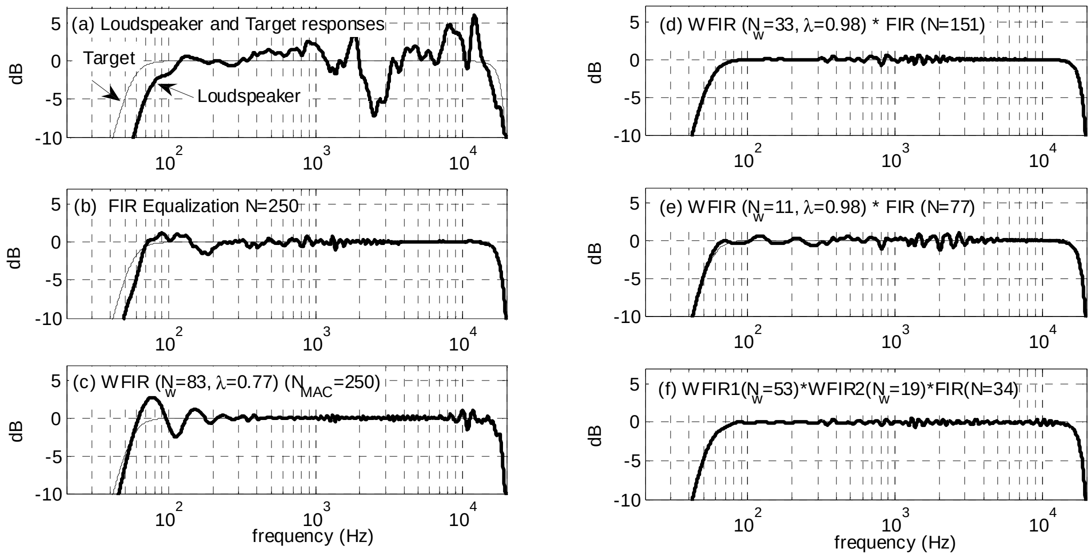


Fig. 5 - Combination of Warped FIR and FIR Equalization examples

Fig. 5 shows several examples for loudspeaker equalization. The loudspeaker and the selected target responses are in Fig. 5a. In all cases, the same computational cost of 250 MACS has been selected for comparison, using a penalty factor of 3 for the warped implementations. Fig. 5b displays the equalization achieved with a FIR filter of N=250, where it is clear how the high frequencies are excellent but the error at low frequencies is considerable due to the lack of resolution. When using only a warped FIR (WFIR) filter of N_w=83 (equivalent to 250 MACS) with λ=0.776 (that maximizes the frequency resolution at 1.7kHz with fs=44.1kHz), the result is Fig. 5c, with better results at low-mid frequencies, but still with problems at the lower and higher frequency regions. The proposed combination [13] is in Fig. 5d. First it applies WFIR of N_w=33 with λ=0.98 for having the maximum frequency resolution at 140 Hz. This WFIR filter achieves an excellent equalization at low frequencies up to 600 Hz. Then a conventional linear FIR filter is applied with N=151 that only needs to equalize above 600 Hz, and achieves it without problem. The result, with the same computational cost is considerably better than the one obtained with only a FIR or only a WFIR filter. Fig. 5e shows the same filter structure of Fig. 5d but now with only 100 MACS using N_W=11 and N=77. The result relies within ±1.5 dB in the whole frequency band with only 100 MACS of computational cost. Finally Fig. 5f show the equalization result obtained with the multistage WFIR and FIR combination proposed in [14] with the objective of having the frequency resolution even more uniform over the spectrum. In this case, first a WFIR with N_W=53 and λ=0.71 is used for the mid frequencies. Second another WFIR with N_W=19 and λ=0.98 is applied for the lower frequencies, and finally, a FIR of N=34 is used for the high frequencies. The result is similar than Fig. 5d but now the error is more uniform over the spectrum. Similar conclusions are obtained in case of using IIR filters and warped IIR (WIIR) combinations [14].

4. PARALLEL FILTER BANK

The parallel implementation of filters is common in different fields [3]. The author Balázs Bank proposed the use of the parallel filter bank with an optional FIR part for audio equalization, with a predefined and perceptually motivated fixed pole set [15], see Fig. 6 and Eq. 6. In the simplest case, the poles are logarithmically spaced in frequency, leaving only the zeros of the system as the free parameters.

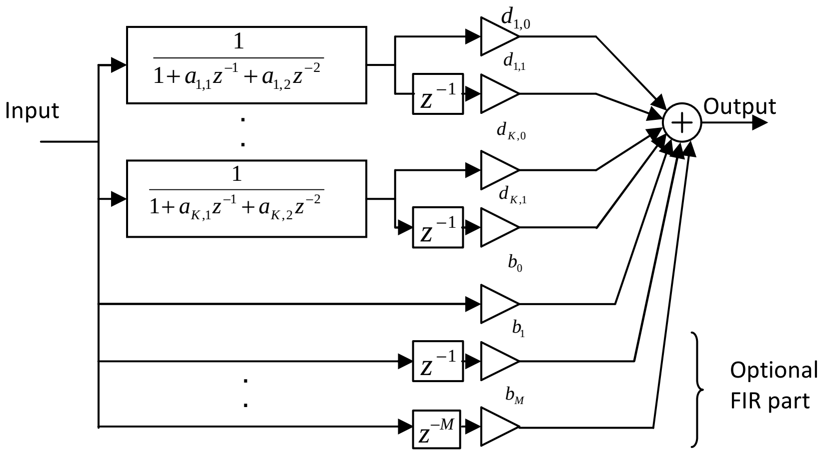


Fig. 6 - Parallel Filter [15]

$$H(z^{-1}) = \sum_{k=1}^{K} \frac{d_{k,0} + d_{k,1} \cdot z^{-1}}{1 + a_{k,1} \cdot z^{-1} + a_{k,2} \cdot z^{-2}} + \sum_{m=1}^{M} b_m \cdot z^{-m} \qquad (6)$$

Once the order of the filter (number of parallel sections K and order of the optional FIR part M) is defined and the poles (and so the coefficients $a_{k,1}$ and $a_{k,2}$) are fixed (e.g., by following a logarithmic scale), all the zeros (defined by $d_{k,0}$, $d_{k,1}$, and b_m) are easily obtained by the linear least squares equations, either in the time domain as explained in [15], or in the frequency domain [16]. The main benefit of this approach is that it is very straightforward to introduce some poles at low frequencies and spread the resolution logarithmically over the whole frequency band as needed. When using linear methods for finding the poles like Prony's or Yule-Walker, as shown in section 2, it is difficult to get good results at low frequencies (introduce some poles) with low order filters (Fig. 3c and 4). The parallel filter approach can be also used as an alternative to complex-smoothing and has been demonstrated to be mathematically equivalent to Kautz filters [17], but requiring lower computational cost.

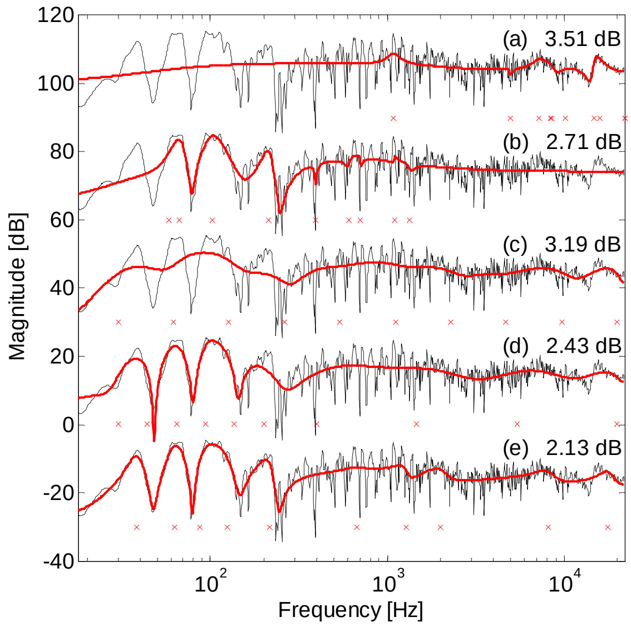


Fig. 7 - Parallel filter implementations of order 20 with different pole-positioning techniques. Graphs are scaled by 20 dB for clarity. The crosses indicate the poles frequencies.

An example of modeling a loudspeaker measured in a room is shown in Fig. 7, which has a particularly difficult response to model at low frequencies. Only the IIR part of the parallel filter is used in these examples, where filters of order 20 (10 sections) with different pole positioning techniques have been employed. Fig. 7a shows the measured response by thin black line, and the response obtained using a 20th order IIR filter designed by the Steiglitz-McBride algorithm. It can be seen that only the high-frequency part of the target is well modeled, due to the linear frequency resolution of the method. Fig. 7b displays a parallel filter response using a warped IIR design based pole set. A warping parameter λ=0.95 was used to help modeling the low frequencies, but still some peaks and notches are not identified and the high frequencies are not considered. Using a logarithmic pole positioning, Fig. 7c, a smoothed model [17] is obtained in the whole frequency range, but losing the details needed in the range below 200 Hz. In Fig. 7d, a stepwise logarithmic pole positioning with higher resolution (more pole density) at low frequencies gets a better result because in this particular example, more poles are necessary at low frequencies where the response to be modeled has more peaks and notches. Finally Fig. 7e displays the results obtained with the idea proposed recently by the authors at [18]. It combines the benefits of a multi-stage warped design for improving the frequency resolution with the simple implementation of the parallel bank (Fig. 6), requiring less computational cost than using a WIIR filter implementation as in [14]. This last model leads to the best modeling results since it is able to look for the proper frequencies and Q factors of the poles in the whole audio range. The steps of the design procedure are as follows (for details, refer to [18]):

1. Split the frequency axis in two or more bands. In most cases, two bands are enough.
2. Distribute the total number of poles between the bands as needed.
3. For each band, force the outer band part of the response to be constant. This will avoid the later processes to use poles outside their band.
4. Use a λ value for each band that maximizes the frequency resolution at the mid-log frequency of its band (Eq. 5).
5. In the warped domain, use any linear IIR filter design method like Prony's or Steiglitz-McBbride to find the warped-poles.
6. Then, for each band un-warp the poles [18] to get them to the original frequencies.
7. Finally, unite all the de-warped poles from the different bands, and find the zeros by the linear least squares fit as before [15].

5. CONCLUSIONS

It is clear that audio-related transfer functions to be equalized or modeled require logarithmic frequency resolution as opposed the linear one used by most traditional filter design methods. Otherwise, high order filters (and high computational cost) would be needed to properly model or equalize a system with a similar perceived quality at the whole frequency range. This paper has covered several audio-specific strategies for obtaining low-order filters developed by the authors, including the use of a parametric optimization; combination of linear and warped filters; and a parallel filter bank implementation where the poles are pre-defined with different strategies. In all cases significantly better perceived results are obtained for the same filter order compared to the ones obtained with conventional linear frequency resolution approaches.

6. ACKNOWLEDGEMENTS

The Generalitat Valenciana has supported this work under the project GV/2013/134. Part of this work has been also founded by the Spanish Ministry of Economy and Competitiveness supported under the projects TEC2012-37945-C02-01/02. The work of Balázs Bank has been supported by the Bolyai Scholarship of the Hungarian Academy of Sciences.

7. REFERENCES

[1] G. Ramos and J. J. Lopez, "Filter design method for loudspeaker equalization based on IIR parametric filters," J. Audio Eng. Soc. 54, 1162–1178 (2006)
[2] G. Ramos, P. Tomás, "Improvements on Automatic Parametric Equalization and Cross-Over Alignment of Audio Systems", presented at the 124th Convention of the Audio Engineering Society, (2009)
[3] U. Zölzer, Digital Audio Signal Processing, Ed. John Wiley & Sons, 2nd Ed. (2008)
[4] A. V. Oppenheim, R. W. Schafer, Discrete-Time Signal Processing. Ed. Prentice Hall (1989)
[5] R. Bristow-Johnson, "Cookbook formulae for audio EQ biquad filter coefficients", on-line publication: http://www.musicdsp.org/files/Audio-EQ-Cookbook.txt
[6] B. Friedlander, B. Porat, "The Modified Yule-Walker Method of ARMA Spectral Estimation," IEEE Transactions on Aerospace Electronic Systems, AES-20, no. 2, pp. 158-173 (1984)
[7] Parks, T., C.S. Burrus, Digital Filter Design, Wiley & Sons (1987)
[8] K. Steiglitz and L. E. McBride, "A Technique for the Identification of Linear Systems", IEEE Transactions on Automatic Control, vol. AC-10, no.4, pp. 461-464, (1965)
[9] G. Ramos, J. J López, "Parametric Loudspeaker Equalization. Results and Comparison with other Methods", presented at the 122nd Convention of the Audio Engineering Society, preprint 7027 (2007)
[10] G. Ramos, M. Cobos, "Parametric head-related transfer function modeling and interpolation for cost-efficient binaural sound applications", J. of Acoustical Society of America, vol 134, no. 3, pp. 1735-1738 (2013)
[11] E. C. Levi, "Complex-Curve Fitting," IRE Trans. on Automatic Control, vol. AC-4, pp. 37-44 (1959)
[12] M. Karjalainen, E. Piirilä, A. Järvinen, and J. Huopaniemi. "Comparison of Loudspeaker Equalization Methods Based on DSP Techniques", J. Audio Eng. Soc. vol. 47, no. 1/2, pp. 14-31 (1999)
[13] G. Ramos et al,, "Cascaded warped-FIR and FIR filter structures for loudspeaker equalization with low computational cost requirements", Digital Signal Processing, vol. 19, no. 3, pp. 393-409, Ed. ELSEVIER (2009)
[14] G. Ramos, Jose J. Lopez, Basilio Pueo, "Combination of warped and linear filter structures for loudspeaker equalization", presented at the 124th Convention of the Audio Engineering Society, preprint 7402 (2008)
[15] B. Bank, "Perceptually Motivated Audio Equalization Using Fixed-Pole Parallel Second-Order Filters," IEEE Signal Processing Letters, vol, 15, pp. 477-480 (2008)
[16] B. Bank, "Logarithmic Frequency Scale Parallel Filter Design With Complex and Magnitude-Only Specifications" IEEE Signal Processing Letters, vol. 18, no. 2, pp. 138-141 (2011)
[17] B. Bank, "Audio Equalization with Fixed-Pole Parallel Filters: An Efficient Alternative to Complex Smoothing," The Journal of the Audio Engineering Society, Vol. 61, Iss. 1/2, pp. 39-49 (2013)
[18] B. Bank, G. Ramos, "Improved Pole Positioning for Parallel Filters Based on Spectral Smoothing and Multi-Band Warping", IEEE Signal Processing Letters, vol. 18, no. 5, pp. 299-302 (2011)