Comparison of Loudspeaker Equalization Methods Based on DSP Techniques

MATTI KARJALAINEN, AES Member, ESA PIIRILÄ**, AES Member, ANTTI JÄRVINEN, AES Student Member, AND JYRI HUOPANIELI***, AES Member

Helsinki University of Technology, Laboratory of Acoustics and Audio Signal Processing, Espoo, Finland

Methods of loudspeaker response equalization using digital filters are compared. In addition to generally known methods and techniques a recently introduced new principle, based on warped filters, is described. Different equalization methods are compared from the points of view of equalization error both objectively and subjectively, computational efficiency, as well as robustness and precision requirements of each method. The study is limited to the linear (magnitude and phase) equalization of loudspeaker free-field responses.

0 INTRODUCTION

Digital filtering provides an attractive means to make loudspeaker performance more ideal. Equalization of the magnitude and the phase response is achievable by linear filtering, and compensation of nonlinear distortion is possible using nonlinear modeling techniques. Since loudspeakers are practically always used in reverberant spaces, the equalization of the entire path from an amplifier to a listener is important.

In this study we analyze and compare some methods of linear response equalization, the main emphasis being on the loudspeaker itself, which is considered as a linear and time-invariant (LTI) system. Thus the equalizer should be an approximation of the inverse transfer function of the loudspeaker. Nonlinear distortion will not be considered in this paper.

First an overview of digital filtering techniques for loudspeaker response equalization will be given. This overview includes the design methods of traditional FIR and IIR filters. Multirate solutions are also a potential technique, but will not be considered here. As a recently introduced new method we describe warped filter equalizers, that is, warped FIR (WFIR) and warped IIR (WIR) filters. They have the interesting property that filter design can be accomplished on psychoacoustically motivated frequency scales, such as the Bark scale or an approximated logarithmic scale.

Next some psychoacoustical aspects that are important to the topic are presented. Special attention is paid to proper frequency scales for equalizer design and the evaluation of response properties. The audibility of magnitude and phase response errors is discussed based on literature and some listening experiments that were carried out.

In the analysis and comparison of the equalization techniques we focus to the following aspects. The equalization error (magnitude and phase errors) is studied using both objective and subjective criteria. In both cases performance is evaluated in relation to filter complexity. Thus the computational efficiency of filter implementations is discussed as an essential factor, as well as the robustness and precision requirements when implementing the equalizers using typical digital signal processors.

A set of listening experiments has been carried out for subjective evaluation and comparison of different equalization techniques. The goal was to find the order of each equalizer filter type that corresponds to the JND (just noticeable difference) threshold of equalization error using various test signals. The results from listening experiments are compared to the results using objective criteria.

Finally, a general discussion is included related to many practical issues of loudspeaker response equalization, such as target response specification, equalization for a sector of free-field responses, the need for phase response correction, as well as combined loudspeaker–room response equalization.
1 OVERVIEW OF FILTER DESIGN METHODS FOR LOUDSPEAKER EQUALIZATION

Various methods for digital equalization of loudspeaker systems have been published. Digital loudspeaker equalization can be used for both magnitude and phase correction, which is unachievable with analog techniques. Nonlinear distortion can also be compensated by digital signal processing [1], [2], but will not be considered in this paper. Although digital processing can produce almost perfect results, one has to consider what needs to be equalized and what is the optimal target response. The magnitude response is audibly more significant. In fact the question of subjective importance of linear phase in loudspeaker systems remains to be answered.

A straightforward approach is to design an equalizer to correct only the on-axis response. However, nearly ideal axial equalization can even degrade off-axis and subjective performance. Therefore the desired target response in most cases should be some form of compromise between on-axis and spatially averaged responses [3].

Although the greatest benefit of equalization is generally obtained at middle frequencies, special care should be taken with low- and high-frequency regions. Especially boosting low frequencies can induce nonlinear distortion. The target response must be compensated with a proper filter defined by the loudspeaker’s characteristics or measured response preequalized prior to the filter design [4].

FIR filters have several advantages over IIR filters, such as being always stable, easier to design, and having arbitrary phase and magnitude response. Therefore FIR filters have been used widely for equalizing audio systems. On the other hand IIR filters model magnitude response more efficiently with less computational complexity. IIR filters are often used for bass equalization, which would result in excessive FIR filter lengths.

There are several general methods for designing FIR and IIR filters (for example, [5]), and they can be used for equalizing loudspeaker response. One of the simplest ways to design an FIR filter is the frequency sampling technique using the inverse Fourier transform. Other design techniques are based on either time- or frequency-domain optimization. The efficient inverse filter for both magnitude and phase correction can be optimized using the least mean squares algorithm [6]. The basic idea is to design a deconvolution filter such that the complete response of the system is an ideal impulse response.

Greenfield and Hawksford have proposed a novel IIR filter design method for loudspeaker equalization [4], [7]. The method is based on loudspeaker modeling and separating the minimum-phase and all-pass components. An advantage of this method is the efficient magnitude response equalization and optional excess-phase correction.

Other attractive approaches for digital loudspeaker enhancement are multirate signal processing [8], frequency masking technique [9], and integrating filters [10], among others.

2 WARPED FILTERS AND THEIR APPLICATION TO EQUALIZATION

Warped digital filters and their audio applications have been introduced and discussed in more detail elsewhere [11]–[15]. Only a short overview is given here.

The basic idea of warping is best illustrated using the FIR-like structure (WFIR) in Fig. 1(a). The specific WFIR structure used in the WFIR loudspeaker equalization of this paper is shown in Fig. 1(b). When the unit delays of an ordinary FIR filter are replaced with first-order all-pass sections, the resulting filter is a warped one, which can be designed on a warped frequency scale based on bilinear conformal mapping,

\[
\hat{z}^{-1} = D(z) = \frac{z^{-1} - \lambda}{1 - \lambda z^{-1}}
\]

where \(\lambda\) is a warping parameter and \(D(z)\) is a warped delay element. The group delay of \(D(z)\) is frequency dependent so that positive values of \(\lambda\) yield increased resolution for low frequencies and negative \(\lambda\) values enhance the resolution at high frequencies. This is illustrated in Fig. 2, which shows the warping by a first-order all-pass section as a function of frequency.

Warped IIR (WIIR) filters cannot be realized directly due to the delay-free propagation of the signal through
all-pass sections and recursive feedbacks, when $\lambda \neq 0$ [see Fig. 3(a)]. There exist modifications, however, that make WIIR structures realizable and practical. Fig. 3(b) shows one of them [16], [15], which is the structure used in the WIIR equalizer examples of this study.

As noted, the value of the warping parameter $\lambda$ controls the amount of warping that is desired. From the point of view of auditory perception a specific value of $\lambda$ yields a good approximation of the Bark scale [17], which traditionally is used as a psychoacoustical pitch scale. A formula to compute this value as a function of the sampling rate is given in Smith and Abel [18]. For example, for a sampling rate of 44.1 kHz this yields $\lambda = 0.7233$. Such a desirable match to the properties of the human auditory system is utilized in the following when warped filters are designed for loudspeaker response equalization.

1 The optimal value for Bark mapping has been reiterated in [19] but the original value given in [18] has been used in our study.

The effect of warped filter design on loudspeaker equalization, compared with traditional unwarped filters, can be seen in the example of Fig. 9, which will be discussed in more detail later. While with traditional filters the frequency resolution is constant on a linear frequency scale, resulting in good equalization at high frequencies but worse performance at low frequencies, a Bark scale warped filter focuses the main resolution and equalization power for middle frequencies. This desirable behavior will be further discussed later from the point of view of psychoacoustics. To some degree it is also possible to achieve a similar effect using weighting functions in traditional filter design [5], but frequency warping makes this balancing of resolution appear automatically.

One more favorable property of warped filters is their inherent robustness and noncritical precision requirements, based on the use of all-pass subsections. In particular, when the density of poles and zeros in the $z$ domain—especially corresponding to low frequencies—is high, traditional filter structures such as direct-form IIR filters become very problematic. Due to the bilinear

![Fig. 2](image-url)  
Fig. 2. Frequency warping characteristics of first-order all-pass section for different values of warping parameter $\lambda$. Frequencies are normalized to Nyquist rate.

![Fig. 3](image-url)  
Fig. 3. (a) Unrealizable direct form of WIIR filter. (b) Realizable modified implementation used in present study.
warping (rotation) of poles and zeros in the z domain such pole and zero densities are relaxed considerably. Typically direct-form IIR filters of higher order than about 20–25 cannot be implemented, even when using floating-point processors. Corresponding warped filters remain stable and realizable even with orders higher than 100 and fixed-point computation. This is the reason for using the warped structures instead of remapping them back to equivalent traditional filter structures.

As can be noticed when comparing the warped filters in Figs. 1 and 3 with traditional FIR and IIR filter structures of the same order, the warped structures are more complex and thus computationally more expensive. This is often compensated for, however, since considerable reduction in filter order is possible due to a good match to human auditory properties.

We have studied the efficiency issues of warped filters on typical digital signal processors [13]. For example, the Motorola 56000 fixed-point processor can run FIR filters one tap (order) per instruction cycle and per sample, whereas WFIR filters take three instructions. The performance ratio is 1:3 in favor of FIRs. While direct-form IIR filters take two instructions per order and per sample, WIIR filters take four instructions, so that the performance ratio is 1:2. For the TMS 320C30 floating-point processor the performance ratios are 1:4 and 1:2.5, respectively. In some applications the warping reduces the filter order by a factor of 5 or more, which means that the overall efficiency of warped filters can be better or about same as for traditional filters.

The design of warped equalizer filters can be done in a straightforward way as follows. The measured impulse response $h(n)$ of a loudspeaker to be equalized is first mapped to the warped time-domain response $\hat{h}(i)$ using the inverse mapping of Eq. (1),

$$\hat{h}(\hat{z}) = D_1^{-1}(z)H(z)$$  

(2)

as described in [15]. This warped impulse response can then, in principle, be used in any inverse filter design technique to yield an FIR or IIR structure, which has to be implemented as a corresponding warped structure. In practice, this is easiest to apply to a minimum-phase version of the loudspeaker response, which means that excess-phase response errors will not be equalized. For WIIR designs we have used successfully Prony’s method available in MATLAB [20]. Further details of warped equalizer designs will be discussed later in this paper.

### 3 Psychoacoustic Criteria

In an ideal case the loudspeaker is an LTI system with flat magnitude response and constant group delay to the point of interest in the acoustic field. In practice a loudspeaker normally radiates in a room where an infinite number of paths from the source to the listener is found. Thus in the first phase we must simplify the problem and study the free-field response as well as the perception by the listener in a specific direction, especially the far-field main-axis behavior.

Although nonlinear distortions from less than 1% up to much higher values at low frequencies and high amplitudes can be found, the scheme of LTI modeling and equalization of loudspeakers is found useful and will be applied in this study. Correspondingly, the perception of nonlinear distortion is not considered here.

#### 3.1 Psychoacoustically Valid Frequency Scales and Resolutions

It has been a long tradition in audio technique to plot magnitude responses using the decibel scale for the ordinate and a logarithmic frequency scale for the abscissa. This was found to describe the auditory perception better than when linear scales are used, and this is also convenient enough technically.

Digital signal processing (DSP) exhibits an inherent property to express practically everything on a linear frequency scale so that adapting to other scales needs special attention. This is due to the properties of the unit delay as a basic building block, which implies uniform time and frequency resolution. Spectrum analysis through the discrete Fourier transform shows this and, more important from the equalization point of view, filter designs follow the same rule unless special efforts are made.

In psychoacoustics it has been shown experimentally that there are yet better scales than the linear or logarithmic frequency scales and the logarithmic dB scale. Loudness in sone units [17] represents the perceived “intensity” and loudness level in phon units is a related logarithmic scale. Pitch, the perceived “height” of sound, has several competing scales. The traditional mel scale has in many technical fields been replaced by the Bark scale (or the critical-band rate scale) [17] although in practice these are very similar (1 bark = 100 mel). A strong competitor of the Bark scale is the $ERB$ (equivalent rectangular bandwidth) rate scale [21], which seems to be theoretically better motivated than the Bark scale [22].

Actually we should make a difference between frequency resolution functions and pitch scales. Fig. 4 shows the four resolution functions discussed: linear, logarithmic, Bark, and $ERB$ resolution in terms of the corresponding $Q$ value (center frequency divided by bandwidth) as a function of frequency. Linear resolution is plotted for uniform 100-Hz bandwidth and logarithmic resolution for one-third-octave bandwidth. Fig. 5 shows the corresponding rate scales versus log frequency.

As can be seen from Figs. 4 and 5, the logarithmic and the $ERB$ resolution functions are relatively close to each other. The Bark resolution is similar above 500 Hz. The constant-bandwidth resolution function related to the linear frequency scale is generally not acceptable when characterizing responses from the auditory point of view. This is unfortunate since DSP methods, excluding filter design methods, work inherently on a linear scale. Based on this theoretical discussion we may draw the conclusion that both the design of equalizer filters and the characterization of equalized responses are best represented on the $ERB$ scales, the logarithmic and the Bark

scales being useful approximations, and the linear scale being inferior.

3.2 Perception of Loudspeaker Response Errors

The two factors to be studied when considering linear distortion in loudspeakers (without room effects) are the magnitude response and the phase (or time delay) response. The detectability of linear distortion depends very much on the type of test signal, not just on the loudspeaker properties. Due to such difficulty and variability of the problem, existing knowledge is not always very consistent and the data cannot always be applied easily to specific problems of sound quality evaluation.

Narrow-band deviations from a flat magnitude response are most easily noticed when a single sinusoidal component falls into a response dip or is strongly emphasized due to a resonance. Otherwise, for broad-band signals, deviations in magnitude response are perceived on a smoothed auditory resolution scale that can be approximated (such as ERB, Bark, or one-third-octave scale). The JND deviation from ideal in A–B comparisons is about 1 dB for a critical band, although well trained listeners may perform slightly better. For some real-world signals, such as music that does not excite a problematic frequency range, the JND threshold may be considerably higher. Also, when the listener does not have a reference of flat response to compare with, even deviations of 5 dB or larger may be difficult to detect or analyze systematically.

An objective way to estimate the subjective perception of response errors is to use computational auditory models. Representative examples thereof have been proposed, for example, by Schroeder et al. [25], Karjalainen [26], Brandenburg [27], and Beerends and Stenderink [28]. These have been developed primarily for the evaluation of audio and speech coding, while the evaluation of loudspeaker systems with them has not been as successful. Since listening tests are very tedious and demanding, such model-based objective methods are highly needed, although they never can fully replace evaluation by human listening.

The second type of linear distortion, the phase or temporal distortion, is less prominent in auditory perception, although in some specific cases the human hearing may be quite sensitive to it. Phase distortion and its perception have been discussed, for example, by Preis [29], Deer et al. [30], Fincham [31], Greenfield and Hawksford [32], and Johansen and Rubak [33]. A typical way to characterize such distortion is to measure and compute the group delay from unwrapped phase as a function of frequency.

A common experience from many listening experiments shows that the level of just noticeable group-delay differences is about 2 ms. Sharp transients and impulses are among the most critical test signals, while for steady-state wide-band signals very much larger errors may remain unnoticed. In some experiments detection of differences below 1 ms have been reported. Since it is relatively difficult to generate complex phase errors without introducing magnitude errors or an uncontrolled temporal spread of signals as well, some of the early experiments may not be reliable.

In order to bring more intuition to the perception of phase response errors, as measured in terms of the group delay, the magnitude response, the magnitude phase response, the magnitude group delay and others, a common experience from many listening experiments shows that the level of just noticeable group-delay differences is about 2 ms. Sharp transients and impulses are among the most critical test signals, while for steady-state wide-band signals very much larger errors may remain unnoticed. In some experiments detection of differences below 1 ms have been reported. Since it is relatively difficult to generate complex phase errors without introducing magnitude errors or an uncontrolled temporal spread of signals as well, some of the early experiments may not be reliable.
delay, we carried out a small set of informal listening experiments (using headphones). We used digital all-pass filters as well as their FIR approximations to yield simple and complex group-delay distortions in an audio channel. For simple cases, such as deviations of group delay within a critical band for middle frequencies, a 2-ms JND threshold typically resulted. For more complex phase curves, such as a step up or down in group delay at a certain frequency, or slightly randomized group delay as a function of frequency, JND thresholds down to about 1 ms or slightly below were found when test signals were impulses band-limited to a 21-kHz audio range. For speech and music the JND thresholds were 3-5 ms up. Small-sized (240 by 190 by 165 mm) two-way vented-box speakers, the group-delay error will be much higher. Thus for medium to high frequencies the need for phase equalization is questionable and magnitude-only correction of response may be well motivated. Phase compensation of low frequencies by DSP may improve bass response quality, but this easily introduces a bulk delay at all frequencies.

Auditory models could also be applied to the evaluation of phase response errors or combined phase and magnitude errors, or better to say, time-frequency response errors. Most computational models used in audio applications have, however, a poor resolution for temporal fine structures (about 10 ms), which makes them useless. Better simulation of the human auditory system could help developing such sound quality tools.

4 COMPARISON OF EQUALIZER FILTER DESIGN METHODS

Three different equalization filter designs have been compared in this study, each applied to two loudspeakers. The filter types are:

1) **FIR filter.** Designed using AR-modeling technique (linear prediction algorithm (LPC) in MATLAB [20]). LPC is applied to the measured impulse response of the loudspeaker and the z polynomial obtained is used as an FIR-type inverse filter of the desired order.

2) **WFIR filter** (warped FIR filter; see Fig. 1). Designed by first inverse warping the minimum-phase version of the impulse response and then applying LPC as in 1).

3) **WIIR filter** (warped IIR filter; see Fig. 3). Designed by first inverse warping the minimum-phase version of the impulse response and then applying Prony’s method available in MATLAB so that the orders of numerator and denominator are equal. Prony’s method was found to work better than other generally available algorithms such as Yulewalk in MATLAB.

Two different loudspeakers were used in each equalizer filter design:

1) **Loudspeaker 1.** A small-sized (240 by 190 by 165 mm) two-way vented-box speaker with a 5-in (127-mm) low-frequency element, a 14-mm dome tweeter, a passive crossover network, and a built-in amplifier. This corresponds to relatively inexpensive loudspeakers typically included in home stereo systems. We have used this loudspeaker in most experiments since it has magnitude response errors and related coloration of sound, which is relatively easy to perceive.

2) **Loudspeaker 2.** A high-quality two-way vented-box studio monitor (310 by 240 by 200 mm), with a 6.5-in (165-mm) bass element, a 19-mm dome tweeter,

---

**Fig. 5.** Mapped frequency scales as functions of frequency. — logarithmic scale (constant Q); ○ — ERB rate scale; + — Bark (critical band); • — linear frequency scale.
and an active crossover network.

4.1 Analysis of Equalization Filter Performance

Figs. 6, 7, and 8 show the behavior of loudspeaker 1 when equalized using different orders of the three filter types: FIR, WFIR, and WIIR, respectively.

In Fig. 6 the FIR design shows a powerful ability to correct the magnitude response, especially at high frequencies, but the filter order must be made rather high, for example, 150–1000. Yet the bass range remains hard to make flat since the linear frequency resolution of traditional filter designs focuses most of its power on high frequencies. Filter design methods where extra weighting can be applied to low frequencies may perform better.

The WFIR design was applied to loudspeaker 1, and the resulting magnitude responses for various filter orders are shown in Fig. 7. Characteristic to it, middle and low frequencies are better equalized with much lower filter orders than with FIR filters, for example, 45–300.

Fig. 8 illustrates the case of WIIR equalizer filters. In this case the equalization power is smoothly balanced for the audio frequency range and the filter order may remain low, such as 16–100. Actually, a useful overall smoothness can be achieved with very low WIIR filter orders, such as 4–16.

Focusing on the frequency resolution, the characteristics of the remaining response ripples and the typical filter orders of the three methods can be compared even more easily in Fig. 9, where three cases with approximately the same amount of overall equalization are shown.

In the warped filter cases the warping has been accomplished according to the Bark scale. It is possible to change coefficient \( \lambda \) in Eq. (1) so that increasing the value of \( \lambda \) focuses the best equalization on lower frequencies and, correspondingly, decreasing \( \lambda \) focuses on higher frequencies. This possibility could also be utilized in digital crossover networks, where each frequency region may be designed with a separate \( \lambda \) value optimized for that band. Furthermore, cascading of subfilters with different \( \lambda \) values can be used to control the focusing of frequency resolution.

The consequence of axial response equalization for off-axis behavior is illustrated in Fig. 10 (original directivity curves), Fig. 11 (FIR equalization of order 105), and Fig. 12 (WIIR equalization of order 24). As can be noticed from the comparison, the high-frequency equalization power of FIR filters does not help much with flattening the off-axis response. Actually, a less ideal axial response with WIIR filters means a slightly smoother off-axis high-frequency response.

Figs. 13, 14, and 15 describe the impulse response behavior of loudspeaker 1 without equalization, with FIR of order 105, and WIIR of order 24, respectively. FIR equalization, due to better high-frequency smoothing, also best reduces fast ringing. Since all three filter designs are basically magnitude-only or minimum-phase based techniques, little or no group-delay equalization has resulted. The need for such equalization is discussed in Section 7.

As another case of loudspeaker equalization we studied the high-quality studio monitor, loudspeaker 2. Fig. 16 illustrates its magnitude response—originally and with various orders of WIIR equalizers. As can be noticed, already relatively low-order filters (16–40) improve the response. The original response is, however, so good that it is somewhat questionable whether the improvement is really needed. Probably much more important would be to equalize the combined response properties of the loudspeaker and the room.

Fig. 6. Equalization of loudspeaker 1 with FIR filters of different orders and designed using LPC algorithm.
4.2 Equalization Errors as Spectral Distance Measure

There is a need for having a simple numerical measure of the equalization quality that is meaningful also from the perceptual point of view. Instead of using more complex auditory modeling schemes, we applied a spectral distance measure in the following way:

- The equalized impulse response is first fast Fourier transformed to power spectrum, resampled (by interpolation) uniformly on a logarithmic frequency scale, smoothed with about 0.2 octave resolution, and converted to the decibel scale.
- The difference between the spectrum to be analyzed and a reference spectrum is computed for the passband region of equalization. The reference spectrum may be simply the average level of the spectrum to be analyzed or some other reference. In our case it was the listening reference in our listening experiments, described in Section 5.

4 This resolution value was specified somewhat arbitrarily to be not too far from their ERB resolution; see Section 3.

Fig. 7. Equalization of loudspeaker 1 with WFIR filters of different orders and designed using warped LPC algorithm.

Fig. 8. Equalization of loudspeaker 1 with WIR filters of different orders and designed using warped Prony's method.
A root-mean-square value of the difference spectrum is computed, and this is used as an objective spectral distance measure to characterize the perceptual difference between the magnitude responses or a deviation from a flat response. Notice that the values of the spectral distance measure used in our study are not calibrated to be compared directly with any perceptual difference measures.

Fig. 17 plots the spectral distance measures as a function of the filter order for the three equalizer filter types used in our study—FIR, WFIR, and WIIR. The reference for the distance computation was the highest order equalized response in order to make the results compatible with the setup used in our listening experiments in Section 5.

There is one observation in the curves of Fig. 17 that needs special attention. The spectral distance measure of FIR equalizers does not decrease monotonically with increasing filter order. This is due to problems with the FIR equalizer design at low frequencies whereby FIR filter orders of 60–150 may not be comparable to...
other data.

Fig. 17 implies that, measured in terms of filter order for a given level of spectral distance, WIIR filters are most compact, followed by WFIR filters, and FIRs need the highest order. If we do the evaluation using the relative computational efficiencies, as given for the two digital signal processors mentioned, the WFIR structure turns out to be the least efficient. For good equalization (low spectral distance) or minor coarse equalization the WIIRs are best, and for a medium level of equalization the FIRs are best. (The problem with the FIR design algorithm disturbs the comparison.)

5 LISTENING EXPERIMENTS

The final judgment of a sound reproduction system must be based on human auditory perception. Thus the final evaluation of loudspeaker equalization techniques should be based on listening experiments. For this reason we carried out a series of such tests in order to compare the performance of various equalizer filter designs. The subjects listened to the loudspeaker under study in an anechoic chamber, comparing various equalized versions of the loudspeaker with different signal stimuli. The equalized signals were prefilted off-line. Data

![Figure 11](image)

Fig. 11. Magnitude responses of loudspeaker 1 in four direction angles after FIR equalization of order 105.

![Figure 12](image)

Fig. 12. Magnitude responses of loudspeaker 1 in four direction angles after WIIR equalization of order 24.
KARJALAINEN ET AL.

PAPERS

were played back using an Apple Macintosh host computer and the QuickSig signal processing environment [36], including a DSP extension based on a National Instruments NB-DSP2300 DSP card with a Texas Instruments TMS320C30 signal processor and high-quality 16-bit analog-to-digital/digital-to-analog converters based on an NB-A2100 board.

A total of nine test subjects participated in the listening experiment, six male and three female, with ages ranging between 21 and 35 years. In the final test only four male and three female subjects were used since the others did not show consistency in their results. The hearing of all test subjects was tested using standard audiometry [37]. None of the subjects had reportable hearing loss that could affect the test results.

In the listening experiment we used an adaptive 2AFC (two alternatives forced choice) bracketing method, similar to audiometric tests but adapted to our purpose. In each trial, two test stimuli were presented with a 0.5-s interval between the samples. The first test signal was always a reference signal which corresponded to practically ideal equalization of the loudspeaker. The second signal varied according to adaptation, descending and ascending five times the order of the equalization filter,

![Image of impulse response] Fig. 13. Original impulse response of loudspeaker 1.

![Image of impulse response] Fig. 14. Impulse response of loudspeaker 1 after FIR equalization of order 105.
bracketing the JND between reference and test signal. This was repeated three times for each filter type using a pink noise test signal (1 s), a speech sample, and a music sample as excitation signals. Peak A-weighted sound pressure levels were 68–70 dB. The average duration of a test session was about 50 min, and the subjects had several short pauses during the session.

5 Music for Archimedes, CD B&O 101 (1992), track 5: 4.0–5.5 s.
6 Best of Sade, CD 01-477793-10 (1994), track 3: 1 min 19 s–1 min 24.6 s.

The test persons were given written and oral instructions. They were also familiarized with a test sequence that demonstrated both distinguishable and undistinguishable test signal pairs. The final test subject was seated in the anechoic chamber and a computer keyboard was placed in front of the listener. Each person was individually familiarized and instructed to respond by pressing key 1 if the signals were perceived the same and key 2 if the signals were perceived different. The space key could be used to repeat a signal pair. The experiment was carried out automatically. Results were gathered automatically by a program written for the

Fig. 15. Impulse response of loudspeaker 1 after WIIR equalization of order 24.

Fig. 16. Magnitude responses of loudspeaker 2 originally and with WIIR equalization of filter orders, 4, 16, 30, and 40.
QuickSig environment. The resulting data were transferred into MATLAB, where the analysis was performed.

5.1 Listening Test Results—Loudspeaker 1

Fig. 18 summarizes the results of the listening tests for loudspeaker 1 equalization experiments. Fig. 18(a) shows cumulative percent curves of inaudibility of differences between the reference and the test signals, as computed from the subjects' judgments, for the three types of filters (FIR, WFIR, and WIIR) and the pink noise excitation. The ordinate is the filter order. Fig. 18(b) summarizes the distribution limits, the median values, and the lower and upper quartiles (25 and 75% levels).

The results show that the distributions of the test subjects' results are relatively broad, which is the consequence of an inhomogeneous listener panel. Some subjects were experienced analytic listeners while most did not have prior experience in a listening panel. A longer training prior to the final experiment could have made the test results more systematic.

Fig. 18 shows that, if the order of the filter as such is a criterion for selecting the equalization scheme, the WIIR filters are found most compact for the purpose. FIR and WFIR filters do not show substantial differences. A useful criterion to select the order of the equalization filter could be the upper quartile (75%) level of the subject's reactions. This means that for FIR filters an order of 80–90 is enough, for WFIRs 75, and for WIIRs 35 might be sufficient, based on results of listening to pink noise. The other types of test signals, speech and music, were found to be less critical. Thus pink noise could be used as the ultimate test signal (for magnitude errors) unless a more critical test signal is found.

To analyze further the listening test results of Fig. 18 we must compare the equalization filters from the point of view of the computational expense of each filter type. The efficiency naturally depends on the hardware and software environment at hand. If the performance data of the Motorola 56000 or the TMS 32OC30, as discussed in the warped filter section, are used, it turns out that the FIR equalizers perform best, WFIRs are 2.5–3.5 times slower, and WIIRs are about 50% slower. If the WIIR order is low enough so that it can be expanded to unwarped direct-form IIR [38], then it is definitely the most efficient solution.

We can compare the results of spectral distance measures and listening tests in order to see whether these correlate. The filter order values of quartile points from Fig. 18 are marked in Fig. 17 for such a comparison. If objective and subjective results were in good agreement, a quartile point (such as 75%) for the three filter types should correspond to the same value of spectral distance. This is quite true for the comparison of WFIR and WIIR filters, although the medians do not match as well, but the 75% and 25% quartiles for the FIR filter are severely out of line. Partly this may be due to the low-frequency response problems discussed earlier in the context of the spectral distance measure.

From Fig. 17 it is difficult to argue why FIR filters performed relatively well in the listening test. One possibility is that the spectral distance measure is not a good indicator of perceived equalization quality. Another explanation might be that the problems with FIR filter design made the corresponding listening experiment unreliable, which may also be the reason for the wide distribution of the subjects' responses. Further studies are needed to answer this question.

![Fig. 17. Characterization of equalization quality using spectral distance measure as a function of filter order for three filter types: FIR, WFIR, and WIIR. Reference response was taken from highest order filter of same type. Marked points correspond to quartile results of listening tests (see Fig. 18).](image-url)
5.2 Listening Test Results—Loudspeaker 2

In another small set of listening tests we studied how equalizations of loudspeaker 2 can be perceived. As can be concluded from Fig. 16, the deviations from the flat magnitude response are within ±1.5 dB between the cut-off frequencies, which is just slightly above the JND threshold. This was confirmed in listening experiments in an anechoic chamber with the pink noise test signal.

For other samples (a speech and a music sample) it was very difficult to say whether any differences could be detected between the original and any high-order equalized response. We can conclude that equalization helps, but it may be questionable if it is really needed in a normal case. Equalization is more useful, however, for the combined response of a loudspeaker and a listening room.

6 PRACTICAL ISSUES AND DISCUSSION

There are many special questions and details that are important when designing digital filters for loudspeaker response equalization. In this section we will discuss some of them as well as further topics for research and development.

6.1 Perception of Phase Equalization

The sensitivity of the hearing system to phase errors was discussed shortly in Section 3. We did not include any equalization methods in our listening tests which do accurate phase correction. Instead we carried out informal experiments to gain some insight into the importance of detailed phase equalization.

We compared some equalizations using signal pairs, where one was with and another without phase equalizations realized by FIR techniques. In the case of loudspeaker 1, using impulselike test signals, we noticed that it is difficult to compensate for the relatively large group delay around the lower cut-off frequency without creating other problems. The reason for this is that additional bulk delay at other frequencies results. Related to this, a kind of preecho or prehiss effect of ripple prior to the main response, which smears the transient response, is very easily obtained. Even when the design is careful with regard to the axial response, off-axis responses may look worse.

Even a preecho of less than 5 ms in duration and much below the peak response level may be noticeable. Thus the use of phase equalization should be carefully weighed against its possible negative side effects. For good loudspeakers where the group-delay distortion is below 1–2 ms (except for low frequencies) the need of such equalization is questionable. For the lowest frequencies the direct sound is not perceived independent of the room resonance modes so that the group-delay response of the whole system is the decisive factor.

6.2 Target Response Design

One important filter design problem is to define a good target response. An ideally flat magnitude response, from zero to Nyquist frequency, cannot be achieved, and attempting it may lead to totally useless results. The most important aspect here is the low-frequency rolloff. For vented-box loudspeakers the relatively sharp fourth-order high-pass characteristics introduce increased group delay in this frequency range. This could be counteracted either by compensating the high pass by its inverse, thus widening the bandwidth downward, or by using FIR techniques in order to delay higher frequencies to match the low-frequency delay.

The first technique is out of the question since such a boost can very easily overload the bass driver unit and increase nonlinear distortion. A good strategy for choosing a target response is to follow the natural rolloff...
below the cut-off frequency. For closed-box loudspeakers the situation is somewhat different. A boost of bass frequencies and widening the band downward may be attractive if the bass driver is good enough from the point of view of nonlinearities. The selection of the target response is always a decision where all important quality factors must be balanced.

The second technique can be implemented with FIR filters of high enough order using methods that realize both magnitude and phase equalization [6]. The resulting excess bulk delay may be problematic in some delay-critical applications and it may also introduce preecho or prehis problems, which was already discussed briefly. Otherwise this is an attractive technique if very flat good phase characteristics are required.

Problems similar to those at low frequencies appear also in the high-frequency region. Too much boosting of any range or individual frequency will easily degrade some other performance measure. The narrow-band resonances and antiresonances as well as the high directivity and direction dependence at high frequencies make optimization of the equalizer design difficult.

There is another aspect of difficulty in target response selection. In methods where the final equalization filter is of the IIR type the problem of stability and numerical accuracy is a critical one. For example, in our study when we designed the WIIR equalizers using the warped Prony’s method we had to set the target function very carefully in order to avoid poles outside or very close to the unit circle.

6.3 Off-Axis Behavior, Sector Equalization, and Room Response

In this paper we considered equalization only in free-field conditions and almost entirely in one direction, such as the main axis—the most natural choice for a single-direction equalization. The off-axis behavior was checked for loudspeaker 1 in Section 4.

The next step, still in free-field conditions, is to optimize the equalization for a given sector of listening angles or for a given weighting function to define the preference of different angles. For magnitude-only equalization of a sector a natural choice is to compute an averaged or angle-weighted power response, which is used for the equalizer design with any inverse filter design method of interest. If both magnitude and phase equalization is needed, a proper error measure, such as the least mean square difference to an ideal impulse, in the time domain and averaged over the sector of interest, could be tried. This seems to be a difficult problem, which is even more difficult when we notice that magnitude and phase properties should be weighted separately.

The final goal is to equalize loudspeakers in common listening environments, such as rooms that add reflections and reverberation [39], [40]. Now the equalization will be dependent on the loudspeaker, the room properties and the listener position. The question is how to balance between direct sound, rapid reflections that add coloration, and later reflections or reverberation. This depends on the time–frequency resolution pattern and the importance of its components from the point of view of auditory perception. A proper time window, such as some decaying tail, is one alternative to weight different delayed components in the loudspeaker–room response. We do not discuss this complex question in this paper. However, it is important to remember that for loudspeakers with a fixed (uncontrollable) directivity pattern the equalizer filter design itself remains the same as long as an optimal or desired target response is known.

7 SUMMARY AND CONCLUSIONS

In this paper we discussed the problem of equalizing loudspeaker free-field responses by means of digital signal processing. After a short overview of known methods which use digital filtering for equalization we presented a new technique where warped digital filters are used to match the frequency resolution of the human auditory system. The advantages of warped filters (WFIRs and WIIRs) in comparison with traditional filter structures are a reduced filter order and increased numerical robustness. A disadvantage is the higher computational complexity of the warped structures, depending on which hardware or software environment is used.

A performance analysis of three filter types—FIR, WFIR, and WIIR—was carried out. We studied the equalization power of each filter type as a function of filter order, applied to free-field responses of two different loudspeakers. Equalized magnitude responses, impulse responses, and directional behavior were illustrated and compared. In addition to visual inspection we used a spectral distance measure to obtain a single number characterizing the equalization quality.

As another approach to compare the performance of the equalization filters we carried out a set of subjective listening tests. The order of each filter type corresponding to the JND threshold of the equalization error was searched for.

Based on the objective analysis method we can conclude that the warped filters have lower order than equivalent FIR filters. If the implementation is based on typical signal processors, however, the performance evaluation shows that FIRs and WIIRs are the most efficient realizations. Comparison between them was somewhat difficult due to a problem in our FIR filter design algorithm. FIRs are simple and absolutely stable. WIIRs have well balanced frequency resolution and best smoothness at middle frequencies. For very low-order coarse equalization the WIIR structure is most attractive.

The results of the subjective listening tests partially supported the objective analysis, but not entirely. Again the warped filters have lower order than the FIR filters for a specified accuracy. In terms of computational efficiency the FIR filters performed in listening tests better than expected, being slightly more efficient than the WIIR filters. Some correlation between the objective and subjective results can be found, but it is fair to state that either a better objective measure is needed or more careful listening experiments should be carried out.

As a final conclusion, the results of our study may
serve as guidelines and insight for loudspeaker equalizer filter design, although no simple rules can be given. Many other filter design alternatives than the three of the present study are available. Many practical factors should be taken into account, such as the type (magnitude versus phase) of correction needed, the frequency distribution of the response error to be equalized, balancing between linear and nonlinear distortion, and the type of signal processing hardware available.

In order to gain further insight into the problem, a more extensive and careful evaluation, both objective and subjective, is needed. One important direction is to develop computational “auditory models” that have good correspondence to human auditory perception. Another necessary extension to the problem domain of the present study is to analyze and model the entire signal path from the loudspeaker through the room acoustics to the listener.

8 REFERENCES


---

### THE AUTHORS

Matti Karjalainen was born in Hankasalmi, Finland, in 1946. He received the M.Sc. and Dr.Tech. degrees in electrical engineering from the Tampere University of Technology in 1970 and 1978, respectively. His doctoral thesis dealt with speech synthesis by rule in Finnish. In 1980 he became associate professor and from 1986 has been full professor in acoustics in the faculty of Electrical Engineering at the Helsinki University of Technology. His research activities cover speech synthesis, analysis, and recognition, auditory modeling and spatial hearing, DSP hardware, software, and programming environments, as well as different areas of audio and acoustics, including musical acoustics and modeling of musical instruments. Professor Karjalainen is a member of IEEE, ASA, EAA, AES, ICMA, ESCA, and several Finnish scientific and engineering societies. E-mail: Matti.Karjalainen@hut.fi

Esa Piirilä was born in Espoo, Finland, in 1972. Currently he is completing his M.Sc. degree in electrical engineering at the Helsinki University of Technology. He is also employed by Sample Rate Systems, Tampere, Finland, where his duties include development of algorithms for use in home theater DSP modules. His main interests are signal processing algorithms and real-time audio applications. He is a member of AES and Acoustical Society of Finland. E-mail: esa.piirila@samplerate.fi

Antti Jarvinen was born in Helsinki, Finland, in 1970. He has been studying at the Helsinki University of Technology, Department of Electrical and Communications Engineering, since 1989. He is currently finishing his master thesis in listening room design and modeling. Mr. Jarvinen has worked in the Laboratory of Acoustics and Audio Signal Processing as a research assistant since 1995. E-mail: antti.jarvinen@hut.fi

Jyrri Huopaniemi was born in Helsinki, Finland, in 1968. He studied acoustics and audio signal processing, multimedia and information technology at the Helsinki University of Technology (HUT), and received his M.Sc. and Licentiate of Technology degrees in electrical engineering 1995 and 1997, respectively. He will be defending his doctoral thesis in the spring of 1999 on the topic of virtual acoustics and 3-D sound. Mr. Huopaniemi worked as research scientist at the Laboratory of Acoustics and Audio Signal Processing of the HUT from 1993 until 1998. In 1998 he was a visiting scholar at the Center for Computer Research in Music and Acoustics (CCRMA) at Stanford University. He is currently with Nokia Research Center’s Speech and Audio Systems Laboratory in Helsinki, Finland. Mr. Huopaniemi’s research activities include 3-D sound and auralization, virtual audiovisual environments, digital audio signal processing, psychoacoustics, room acoustics and musical acoustics. He is author or co-author of over 30 technical papers published in international journals and conferences, and has been actively involved in the MPEG-4 standardization work.

---

Mr. Huopaniemi is a member of the AES and secretary of the AES 16th International Conference on \textit{Spatial Sound Reproduction} to be held in Rovaniemi, Finland, in April 1999. He was secretary and committee member of the AES Finnish Section in 1995–1997. He is also a member of the IEEE, the International Computer Music Association, and the Acoustical Society of Finland. E-mail: Jyri.Huopaniemi@hut.fi