A review of perceptually motivated parallel IIR audio filtering techniques

Robin Rajamäki
Aalto University School of Electrical Engineering

robin.rajamaki@aalto.fi

November 14, 2013

Abstract

Perceptually motivated filtering techniques attempt to incorporate the limitations of human hearing into the filter design process. They offer several advantages over traditional techniques, which usually operate on a linear frequency scale. The main advantage lies in the potential savings in computational cost, as lower filter orders are possible without loss of perceptual sound quality. Additionally, some of these IIR filters are easily made parallel, which allows for efficient multiprocessor implementations. This paper reviews the warped, Kautz and Bank filters. They are all well established perceptual filtering methods, of which the two latter also have a parallel structure. It was found that the Kautz and Bank filters offer the perceptually best performance at the lowest computational complexity.

Keywords — second order parallel IIR filter, warped, Kautz, Bank

1 Introduction

System linearity is a central assumption in signal processing. It is often cumbersome, if not impossible, to express and manipulate systems analytically, when linearity is not a given. As a result, e.g. filters are mostly modeled as linear-shift-invariant (LSI) systems. The human auditory system however, is highly non-linear in terms of both amplitude and frequency resolution. This presents practical problems when designing audio filters, which when using traditional design methods have a linear frequency resolution. It is a well known fact that distinguishing between two tones 5 Hz apart in frequency is a much harder task around 10 kHz than at 100 Hz. This is why the frequency resolution of the human hearing is rather considered to be logarithmic, than linear by nature. From a filter design perspective, a practical implication of this is that a disproportionately large part of the filter resolution is concentrated in the high frequency region, where the resolution of the auditory system is at its worst. Also, computational power is wasted from an optimization point of view, since increasing the resolution of at low frequencies automatically increases the resolution at high frequencies. In other words, one could reduce filter complexity by sensibly concentrating the filter's effort in the right frequency regions.

Perceptually motivated audio filtering techniques have been presented by e.g. Härmä et al. (2000), Paatero and Karjalainen (2003) and Bank (2007). So called warped filters are perhaps one of the earliest of these techniques. The idea of frequency warping is to map a desired system response into a warped (e.g. logarithmical) domain, where the filter design and filtering processes are done, whereafter the filtered signal is converted back to the original (usually linear) domain. Kautz filters are based on the same idea as warped filters, and can be considered to be a generalization of these. In addition to their perceptually relevant frequency resolution, another useful feature of Kautz filters is their parallel structure. This gives a means to efficiently implement Kautz filters on multiprocessor systems, where many parallel operations can be executed simultaneously. Kautz filters are closely related to the newest of the three filtering techniques presented in this paper, namely the parallel second order filter first introduced by Bank (2007). Bank's method builds on the theory of warped and Kautz filters, offering some practical improvements over them in terms of both computational cost and design complexity (Bank, 2008).

It is not hard to think of applications for perceptually motivated filtering techniques. First of all, they usually yield filters of lower order at the same (or better) perceptual quality than traditional design methods, which already assures almost universal applicability in audio processing. The most common applications include equalization or filtering where high filter orders are required, such as room/loudspeaker equalization (Karjalainen and Paatero (2006), Paatero and Karjalainen (2003), Härmä et al. (2000)) or instrument body modeling (Karjalainen and Paatero (2006), Härmä et al. (2000), Bank (2007)) just to name two. Other tried an tested fields of application include linear predictive audio coding, filtering of head related transfer functions (HRTFs) and frequency warped fast Fourier transform (FFT) (Härmä et al., 2000).

This paper serves as a review of perceptually motivated audio filtering methods, with an emphasis on parallel filtering (i.e. Kautz and Bank filters). The paper is structured as follows: In section 2, the theory of the three reviewed methods is presented. Before this however, a pole selection technique is briefly introduced, as the central equations of both the Kautz and the Bank filter require prior knowledge of the filter poles. After the theoretical overview, the methods are compared in MATLAB simulations in section 3. The results and implications of section 3 are then discussed in section 4. The main findings of the paper are summarized in section 5.

2 Theory

In this section the theory of warped, Kautz and Bank filters is presented. The section begins by introducing a simple pole placing technique, which is a crucial pre-step to the parallel filtering methods of section 2.3. The presented filtering techniques are divided into two subgroups according to their structure. Warped filters are considered to have a non-parallel structure, whereas Kautz and Bank filters are have a parallel structure.

2.1 Logarithmic pole placing

A simple way of modeling the non-linear frequency resolution of the human hearing is by placing filter poles non-equidistantly on the frequency axis. The resulting filter should ideally have more poles at low frequencies where the hearing has a better frequency resolution, and less at high frequencies where the resolution is worse. The logarithmic function is an often used approximation, which has this characteristic.

A basic logarithmic frequency mapping technique is outlined by Waters and Sandler (1993). The first step consists of logarithmically sampling the frequency range of interest. Dividing the range $f_k \in [f_{lower},\ f_{upper}]$ into K points yields

$$f_k = (\frac{f_{upper}}{f_{lower}})^{\frac{k}{K-1}} f_{lower}, \tag{1}$$

where $k = 0, 1 \ldots K-1$ and $f_0 = f_{lower}$, respectively $f_{K-1} = f_{upper}$. The next step is to map the frequencies f_k to complex poles p_k, which can be done e.g. by a method presented by Bank (2007). First normalizing the pole frequencies to $\omega_k \in [0,\ \pi]$ gives

$$\omega_k = \frac{2\pi f_k}{f_s}, \tag{2}$$

where f_s is the sampling frequency. The complex conjugate pole-pairs can then be calculated using

$$p_k = r^{\frac{\omega_k}{\pi}} e^{\pm j\omega}. \tag{3}$$

$r \in [0, 1]$ is damping parameter, which reaches its maximum value at f_{upper}, i.e. $k = K-1$.

A more refined (and non-parametric) version of the method described above is also given by Bank (2010). Here, the poles are chosen in such a way that their transfer functions cross at approximately -3dB (Bank, 2010). The method requires computing a bandwidth measure for each sample point ω_k:

$$\Delta\omega_k = \begin{cases} \frac{\omega_{k-1} - \omega_{k+1}}{2}, & \text{when } k = [1 \ldots K-2] \\ \frac{\omega_1 - \omega_0}{2}, & \text{when } k = 0 \\ \frac{\omega_{K-1} - \omega_{K-2}}{2}, & \text{when } k = K-1. \end{cases} \tag{4}$$

The poles are then calculated using

$$p_k = e^{\frac{-\Delta\omega_k}{2}} e^{\pm j\omega}. \tag{5}$$

More intricate pole position techniques than the ones introduced in this section also exist. Since they are beyond the scope of this paper, the interested reader is referred to e.g. Bank and Ramos (2011) for more on the topic.

2.2 Non-parallel methods

Since the focus of this paper is on perceptually motivated parallel IIR filtering techniques, only one non-parallel technique is presented next for the sake of comparison.

Dating back to the 1960s (Härmä et al., 2000), frequency warping is a DSP technique, which essentially maps a linear scale problem onto a non-linear scale. In the case of filter design, the problem translates into finding the filter coefficients in the non-uniformly sampled, "warped" domain. Any given signal can then be filtered using the warped coefficients, thus exploiting the non-uniform frequency resolution.

2.2.1 Warped FIR filters

The warped FIR filter is designed by replacing the unit delays of an ordinary FIR by first order all-pass filters:

$$\tilde{z}^{-1} = D(z) = \frac{-\lambda + z^{-1}}{1 - \lambda z^{-1}}. \tag{6}$$

$D(z)$ is the transfer function of the first order all-pass filter, which only depends on one parameter $\lambda \in [-1, 1]$. λ affects the frequency mapping between the linear and warped domains (Fig. 1), by changing the phase distorting properties of the all-pass filter. An often used value is $\lambda = 0.78$, which increases the resolution at low frequencies and conforms reasonably well to a few well known psychoacoustic models[1]. It is worth noting that for $\lambda = 0$, the transfer function in Eq. 6 reduces back to the unit delay. (Härmä et al., 2000)

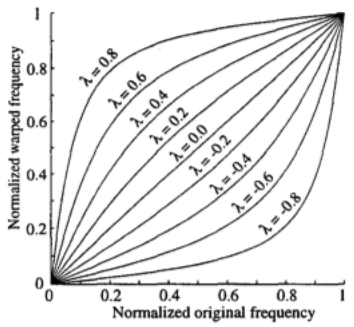

Figure 1: The frequency warping effect of the first order all-pass filter of Eq. 6. Figure by Härmä et al. (2000).

The transformation of Eq. 6 defines a new set of filter coefficients, which effectively redistribute the filter's frequency resolution as wanted. This is also called the pre-warping or analysis step, and is defined

$$\tilde{H}(\tilde{z}) = \sum_{n=0}^{\infty} h[n] z^{-n} = \sum_{n=0}^{\infty} h[n](D^{-1}(\tilde{z}))^{n} = \sum_{n=0}^{\infty} h[n] \left(\frac{\lambda + \tilde{z}^{-1}}{1 + \lambda \tilde{z}^{-1}} \right)^{n}. \tag{7}$$

The inverse transformation of $D(z)$, $D^{-1}(\tilde{z})$, is obtained by simply negating the sign of λ. (Härmä et al., 2000)

The de-warping, or synthesis step, is defined

$$H(z) = \sum_{n=0}^{\infty} \tilde{h}[n] \tilde{z}^{-n} \approx H_{WFIR}(z) = \sum_{n=0}^{Z} \tilde{h}[n](D(z))^{n} = \sum_{n=0}^{Z} \tilde{h}[n] \left(\frac{-\lambda + z^{-1}}{1 - \lambda z^{-1}} \right)^{n}. \tag{8}$$

Since the warped FIR (WFIR) filter does not actually have a finite impulse response (due to the all-pass being an IIR filter), it becomes necessary to truncate it to some finite length. The number of taps in the final truncated WFIR filter is denoted by Z in Eq. 8.

[1]the Bark rate scale and Greenwood's frequency-to-position function, to be precise. (Härmä et al., 2000)

2.2.2 Warped IIR filters

An attempt at extending the techniques of section 2.2.1 directly to IIR filters will result in implementation complications such as delay free loops (Härmä et al., 2000). In practice, it is thus preferable to start with either an unwarped transfer function (Paatero and Karjalainen, 2003) or a pre-warped FIR. In the first case, the poles (p_i) and zeros (c_j) of the transfer function can be mapped to the warped domain using

$$\begin{cases} \tilde{p_i} = \frac{p_i+\lambda}{1+p_i\lambda} \\ \tilde{c_j} = \frac{c_j+\lambda}{1+c_j\lambda}. \end{cases} \tag{9}$$

In the second case, a pole-zero model is fitted to a warped impulse response. For example Prony's method is an effective technique, which can be used to find the IIR filter coefficients at this step (see e.g. Hayes (1996)). Obtaining the final de-warped filter response is not as straightforward as in the case of the WFIR, mainly due to the delay free loops of the direct form implementation. The derivation of a realizable WIIR structure is not relevant in the context of this paper, so the interested reader is referred to Karjalainen et al. (1997) for details.

2.3 Parallel methods

Kautz and Bank filters are closely related to warped filters, and are probably two of the most well known perceptually motivated parallel audio filtering techniques. In addition to the scalability of their parallel structure, they have been shown to offer significant improvements over warped filters in areas such as modeling accuracy (Bank, 2007) and model simplicity (Paatero and Karjalainen (2003), Bank (2008)).

2.3.1 Kautz filters

Kautz filters are another early concept in signal processing. Named after their discoverer in the 1950's, Kautz filters can be seen as an generalization of warped filters in the sense that the chain of all-pass filters no longer depends on only one parameter λ, as Eq. 6 (Paatero and Karjalainen, 2003). A Kautz filterbank consists of a set of weighted parallel and orthogonal sub-filters $G_k(z)$, i.e.

$$H_{Kautz}(z) = \sum_{k=0}^{P} w_k G_k(z), \tag{10}$$

where

$$G_k(z) = \frac{\sqrt{1-p_k p_k^*}}{1-p_k z^{-1}} \prod_{j=0}^{k-1} \frac{-p_j^* + z^{-1}}{1-p_j z^{-1}}. \tag{11}$$

Since the poles p_k and their complex conjugates p_k^* are predetermined, Eq. 10 reduces to a linear equation w.r.t. the unknown filter weights w_k. In other words, an analytic solution to the weights can be found e.g. by minimizing the mean squared error (for an example of this, see section 2.3.2). Defining

$$\begin{cases} \mathbf{G} = \left(\mathbf{g}_0[n]\ \mathbf{g}_1[n]\ \cdots\ \mathbf{g}_P[n]\right) \\ \mathbf{g}_k[n] = \left(g_k[0]\ g_k[1]\ \cdots\ g_k[N-1]\right)^T \\ \mathbf{w} = \left(w_0\ w_1\ \cdots\ w_P\right)^T, \end{cases} \tag{12}$$

the least squares solution that minimizes the error, given a target impulse response h_t of dimension $[Nx1]$, is

$$\boxed{\boldsymbol{w} = (\mathbf{G}^H\mathbf{G})^{-1}\mathbf{G}^H\mathbf{h}_t = \mathbf{G}^+\mathbf{h}_t.} \tag{13}$$

$\mathbf{g}_k[n]$ in Eq. 12 is the inverse z-transform of $\mathbf{G}_k(z)$ evaluated at indices $n = 0...N-1$, and $\mathbf{G}^+$ in Eq. 13 denotes the Moore-Penrose pseudoinvese.

In practice, it is convenient to combine the complex conjugate pole pairs of Eq. 10. This results in a computationally more efficient filter structure (Paatero and Karjalainen, 2003). Such a parallel filter bank consisting of second-order IIR sections is depicted in the left plot of Fig. 2. The right plot illustrates the sum magnitude response of the sub-filters and the ideal target response.

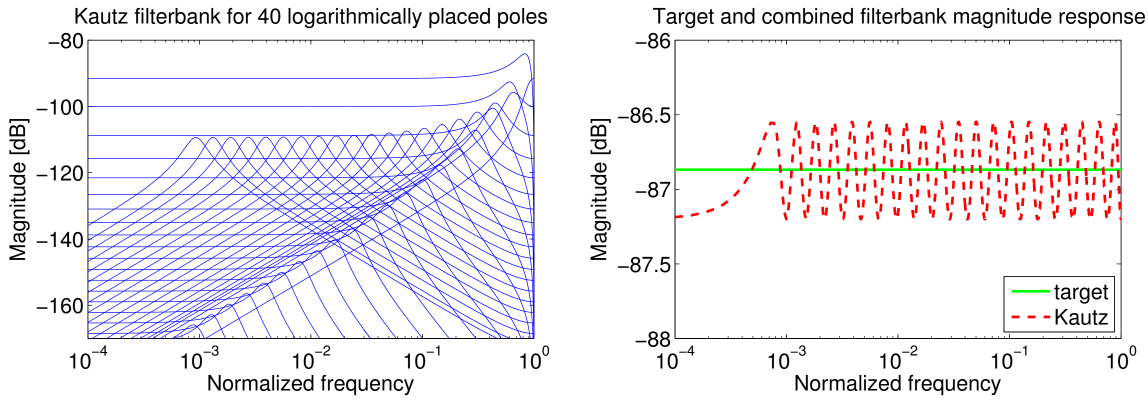

Figure 2: Kautz filterbank. Left: magnitude responses of the sub-filters. Right: target and sum magnitude responses of the filter bank.

2.3.2 Bank filters

An efficient least squares approach to parallel IIR filter design was first presented by Bank (2007). The model is based on the fact that if no duplicate poles exist, any LSI transfer function $H(z)$ can be written in the form

$$H(z) = \frac{B(z)}{A(z)} = \sum_{k=1}^{P} \frac{c_k}{1 - p_k z^{-1}} + \sum_{i=0}^{Z} b_i z^{-i}. \tag{14}$$

P is the number of poles p_k, c_k are the weights of the parallel, first order IIR sections and b_i the coefficients of the FIR section of order Z. Combining the (usually) complex first order

IIR sections with their complex conjugate pairs, or alternatively, combining two real valued poles, yields a computationally more efficient second order parallel IIR structure. If the pairs are consecutive indices of k, their combined term becomes

$$\begin{aligned}\frac{c_k}{1-p_k z^{-1}}+\frac{c_{k+1}}{1-p_{k+1}z^{-1}} &= \frac{c_k(1-p_{k+1}z^{-1})+c_{k+1}(1-p_k z^{-1})}{(1-p_k z^{-1})(1-p_{k+1}z^{-1})}\\ &= \frac{(c_k+c_{k+1})-(c_k p_{k+1}+c_{k+1}p_k)z^{-1}}{1-(p_{k+1}+p_{k+1})z^{-1}+p_k p_{k+1}z^{-2}} = \frac{d_{x,0}+d_{x,1}z^{-1}}{1+a_{x,0}z^{-1}+a_{x,1}z^{-2}}. \quad (15)\end{aligned}$$

In case of complex conjugate pairs, the denominator in Eq. 15 simplifies to

$$1-(p_{k+1}+p_{k+1})z^{-1}+p_k p_{k+1}z^{-2} = 1-(p_k+p_k^*)z^{-1}+p_k p_k^* z^{-2} = 1-2*\mathbb{R}\{p_k\}z^{-1}+|p_k|^2 z^{-2}. \quad (16)$$

Assuming X pairable poles which are reordered after one the other, using Eq. 15 the expression of the IIR section in Eq. 14 can be rewritten as

$$\sum_{k=1}^{P}\frac{c_k}{1-p_k z^{-1}} = \sum_{x=1}^{X}\frac{d_{x,0}+d_{x,1}z^{-1}}{1+a_{x,0}z^{-1}+a_{x,1}z^{-2}}+\sum_{k=2X+1}^{P}\frac{c_k}{1-p_k z^{-1}}. \quad (17)$$

Defining the IIR transfer functions and their time-domain impulse responses

$$\begin{cases}\frac{1}{1+a_{x,0}z^{-1}+a_{x,1}z^{-2}} = U_{x,0}(z) & \Longleftrightarrow \mathcal{Z}^{-1}\{U_{x,0}(z)\} = u_{x,0}[n]\\ \frac{z^{-1}}{1+a_{x,0}z^{-1}+a_{x,1}z^{-2}} = U_{x,1}(z) & \Longleftrightarrow \mathcal{Z}^{-1}\{U_{x,1}(z)\} = u_{x,1}[n] = u_{x,0}[n-1]\\ \frac{1}{1-p_k z^{-1}} = V_k(z) & \Longleftrightarrow \mathcal{Z}^{-1}\{V_k(z)\} = v_k[n],\end{cases} \quad (18)$$

enables expressing Eq. 14 in the time-domain using Eq. 17 and 18:

$$h[n] = \mathcal{Z}^{-1}\{H(z)\} = \sum_{x=1}^{X} d_{x,0}u_{x,0}[n]+d_{x,1}u_{x,0}[n-1]+\sum_{k=2X+1}^{P} c_k v_k[n]+\sum_{i=0}^{Z} b_i\delta[n-i]. \quad (19)$$

Eq. 19 can be written in matrix form when considering N samples of $h[n]$, i.e. $n = 0...N-1$:

$$\mathbf{h} = \mathbf{M}\boldsymbol{\theta}. \quad (20)$$

In Eq. 20, $\mathbf{M}$ denotes the modeling signal matrix

$$\mathbf{M} = \left(\mathbf{u}_{1,0}[n]\ \mathbf{u}_{1,1}[n]\ \cdots\ \mathbf{u}_{X,0}[n]\ \mathbf{u}_{X,1}[n]\ \mathbf{v}_{2X+1}[n]\ \cdots\ \mathbf{v}_P[n]\ \boldsymbol{\delta}[n]\ \cdots\ \boldsymbol{\delta}[n-Z]\right), \quad (21)$$

where

$$\begin{cases}\mathbf{u}_{x,0}[n] = \left(u_{x,0}[0]\ u_{x,0}[1]\ \cdots\ u_{x,0}[N-1]\right)^T\\ \mathbf{u}_{x,1}[n] = \mathbf{u}_{x,0}[n-1] = \left(u_{x,0}[-1]\ u_{x,0}[0]\ \cdots\ u_{x,0}[N-2]\right)^T\\ \mathbf{v}_k[n] = \left(v_k[0]\ v_k[1]\ \cdots\ v_k[N-1]\right)^T\\ \boldsymbol{\delta}[n] = \left(1\ 0\ \cdots\ 0\right)^T.\end{cases} \quad (22)$$

Respectively, $\boldsymbol{\theta}$ is the parameter vector containing all of the system's numerator coefficients

$$\boldsymbol{\theta} = \left(d_{1,0}\ d_{1,1}\ \cdots\ d_{X,0}\ d_{X,1}\ c_{2X+1}\ \cdots\ c_P\ b_0\ \cdots\ b_Z\right)^T. \quad (23)$$

Eq. 20 enables writing the mean squared error between any target impulse response $h_t[n]$, and its estimate $\hat{h}[n]$, in a compact form:

$$\begin{aligned}\epsilon = \frac{1}{N}\sum_{n=0}^{N-1}(h_t[n]-\hat{h}[n])^2 = \frac{1}{N}||\mathbf{h}_t - \mathbf{M}\boldsymbol{\theta}||^2 \\ = \frac{1}{N}(\mathbf{h}_t - \mathbf{M}\boldsymbol{\theta})^H(\mathbf{h}_t - \mathbf{M}\boldsymbol{\theta}) = \frac{1}{N}(\mathbf{h}_t^H - \boldsymbol{\theta}^H\mathbf{M}^H)(\mathbf{h}_t - \mathbf{M}\boldsymbol{\theta}) \\ = \frac{1}{N}(\mathbf{h}_t^H\mathbf{h}_t - \mathbf{h}_t^H\mathbf{M}\boldsymbol{\theta} - \boldsymbol{\theta}^H\mathbf{M}^H\mathbf{h}_t + \boldsymbol{\theta}^H\mathbf{M}^H\mathbf{M}\boldsymbol{\theta}).\end{aligned} \quad (24)$$

The parameter vector $\boldsymbol{\theta}$ which minimizes the mean squared error, is found by differentiating Eq. 24 with respect to $\boldsymbol{\theta}$ and setting the derivative to zero. If a solution exists, ϵ must clearly obtain its minimum for that $\boldsymbol{\theta}$.

$$\begin{aligned}\frac{\partial\epsilon}{\partial\boldsymbol{\theta}} = \frac{1}{N}(-\mathbf{h}_t^H\mathbf{M} - (\mathbf{M}^H\mathbf{h}_t)^H + 2\boldsymbol{\theta}^H\mathbf{M}^H\mathbf{M}) = \frac{1}{N}(-2\mathbf{h}_t^H\mathbf{M} + 2\boldsymbol{\theta}^H\mathbf{M}^H\mathbf{M}) = 0 \\ \Rightarrow \boldsymbol{\theta}^H\mathbf{M}^H\mathbf{M} = \mathbf{h}_t^H\mathbf{M}^H \Rightarrow (\boldsymbol{\theta}^H\mathbf{M}^H\mathbf{M})^H = (\mathbf{h}_t^H\mathbf{M}^H)^H \Rightarrow \mathbf{M}^H\mathbf{M}\boldsymbol{\theta} = \mathbf{M}^H\mathbf{h}_t.\end{aligned} \quad (25)$$

Multiplying the final expression in Eq. 25 from the left by the inverse of the square matrix $\mathbf{M}^H\mathbf{M}$, yields the well known least squares solution

$$\boxed{\boldsymbol{\theta} = (\mathbf{M}^H\mathbf{M})^{-1}\mathbf{M}^H\mathbf{h}_t = \mathbf{M}^+\mathbf{h}_t. \quad (26)}$$

Note again that a necessary prerequisite for being able to solve Eq. 14 using a linear system of equations, is the prior assumption that all poles of the filter are known. It is only in this case that Eq. 19 reduces to a linear problem with regard to the unknown filter coefficients.

Fig. 3 shows the parallel second-order subfilters of a Bank filter (without the FIR section of Eq. 14). The figure is in essence the same as Fig. 2, but for a higher filter order. In fact, the two figures are indentical when given the same filter order and complex conjugate pole pairs, which illustrates the similarity of Kautz and Bank filters.

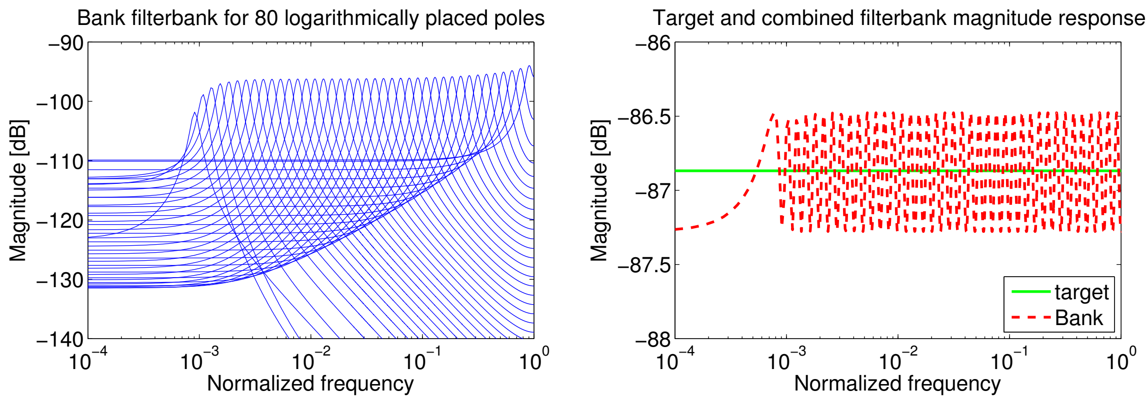

Figure 3: Bank filterbank. Left: magnitude responses of the sub-filters. Right: target and sum magnitude responses of the filter bank.

3 Simulation

In this section the performance of the presented algorithms is evaluated in two MATLAB simulations. The first simulation compares how well the algorithms conform to a given specification and the second simulation shows how the fitting error changes as a function of the filter order. Impulse responses measured at Pori concert hall in Finland by Merimaa et al. (2005) were used as data for the simulations.

Fig. 4 displays the magnitude responses of the 12^{th} octave smoothed target impulse response and its approximation given by the six algorithms. The Bank (without the FIR section) and Kautz filters use 40 logarithmically spaced poles (black crosses in the upper plot of Fig. 4), which are calculated using Eq. 5. The poles and zeros of the warped filters are determined by their fit to the pre-warped target impulse response. The reference FIR and IIR filters in the lower part of Fig. 4 have a linear frequency resolution (IIR poles marked as black crosses in the lower plot of Fig. 4), but the same model order as the perceptually motivated filters. Table 1 shows the root mean squared errors (RMSE) between the target and model magnitude responses between 20-20 000 Hz. The first row shows the error measured using a linear sampling of the frequency axis, whereas a logarithmic spacing was used for the second row. Fig. 5 is an extension of Table 1 and shows the logarithmic frequency scale RMSE as a function of model order.

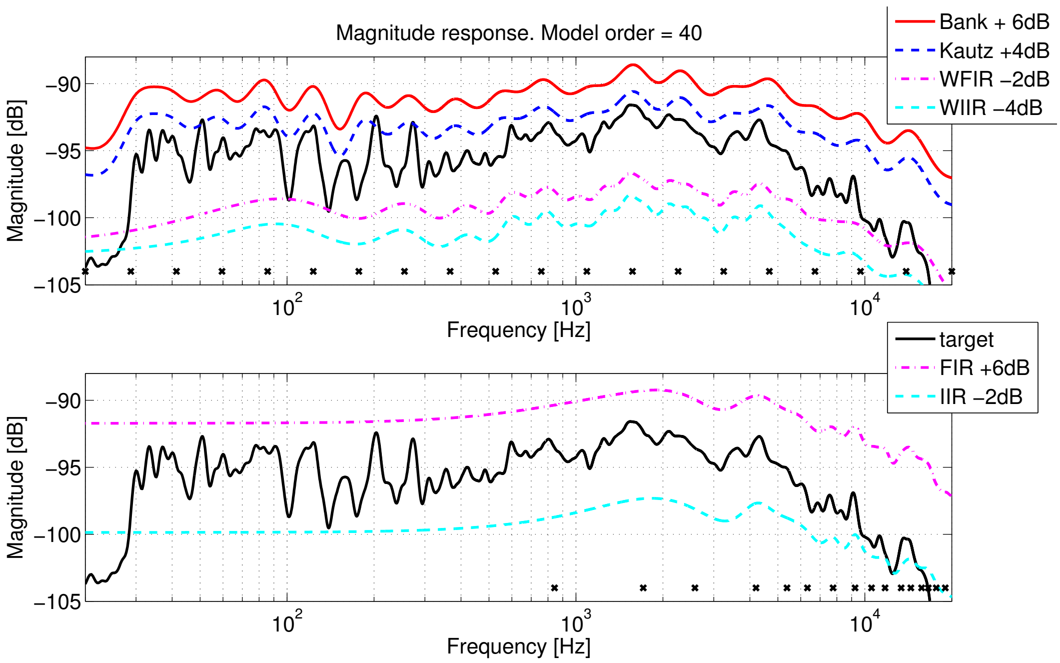

Figure 4: Conformance to target specification of the 40th order filters. Upper plot, bottom to top: WFIR, WIIR, target, Kautz, Bank. Lower plot, bottom to top: linear scale FIR, target and linear scale IIR. The black crosses mark the pole frequencies of the Bank and Kautz filters (upper plot) and the linear IIR filter (lower plot).

Table 1: Magnitude spectrum RMSE [dB] (20-20k Hz, model order 40) using linear and logarithmic frequency scale weighting.

	Bank	**Kautz**	**WFIR**	**WIIR**	**FIR**	**IIR**
lin	3.1	3.1	3.1	2.9	3.1	3.3
log	2.4	2.4	2.6	2.6	2.8	2.9

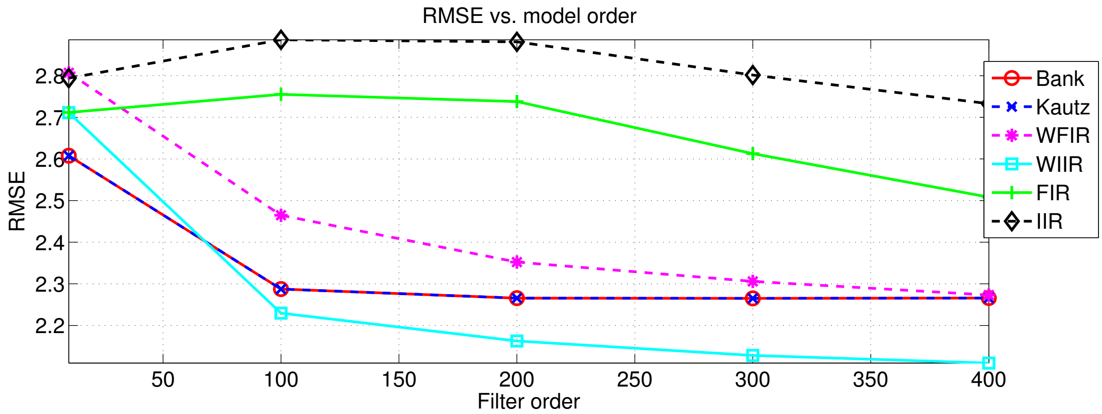

Figure 5: Magnitude spectrum RMSE [dB] as a function of filter order, using a logarithmic frequency scale weighting between 20-20000 Hz).

4 Discussion

4.1 Comparing the methods

Comparing the two plots in Fig. 4 illustrates the main difference between perceptually motivated and traditional linear-scale filtering techniques. Assuming that the logarithmic frequency scale approximates the frequency resolution of the human ear better, it is clear from the Fig. 4 alone that the perceptually motivated techniques outperform their linear scale counterparts. The price to pay for this is decreased resolution at high frequencies, which is better for the FIR and IIR filter of the lower plot. Overall, from Table 1 one can see that neither filter type is generally better in the linear scale (the first row), as the RMSEs of all filters are practically the same. However, a difference between the two can be seen when using a logarithmic frequency weighting. This essentially gives the low frequencies a higher weight in the error, which as expected results in a lower RMSE for the perceptual techniques.

Fig. 5 shows the RMSE of the discussed algorithms as a function of the filter order. Clearly, the WIIR offers the best fit for high model orders. It is important to keep in mind though, that the logarithmically placed poles used by the Kautz and Bank filters are not used by the WIIR, whose poles are determined by Prony's method on the pre-warped target impulse response. In fact, at least Bank filters are known to outperform WIIRs when using the same pole set (Bank, 2007). One can however also see from Fig. 5 that both the Bank and the Kautz filters offer the second best fit to the target magnitude response over a wide range of filter orders. The main advantage of these two methods compared to the WIIR is that they offer a smaller computational complexity. Both methods also benefit from having a parallel structure. In fact, the theoretical savings are directly proportional to the number of parallel filters, which for second order parallel sections can be up to $\frac{model\ order}{2}$. The fact that the warped filters cannot match the resolution or the computation times of parallel filters, renders them somewhat obsolete in comparison. In addition, the pre-warping process was found to be a very CPU intensive task, which further decreases the appeal of frequency warping techniques. As a result, they are included in this study mainly due to being conceptually fairly simple to understand and due to their ability to generate perceptually relevant filter poles.

It has been estimated that perceptually motivated filtering techniques can require four to five times less operations than traditional IIR filtering techniques (Bank (2007), Karjalainen et al. (1997)). Furthermore, Karjalainen et al. (1997) have argued that numerical issues prevent the design of direct form IIR filters of as high order as perceptual filters. These facts, along with the possibility of a non-uniform frequency resolution, should ultimately justify the use of perceptual, and in particular parallel filters in audio applications.

4.2 Future work

Since this paper only acts as an introduction to some perceptually motivated (parallel) filtering methods, many relevant and interesting topics naturally remain to be covered. Perhaps the most important issue for future work would be to implement efficient versions of some of the presented algorithms on real-world DSPs or GPUs. This would give a realistic insight into the computational costs of the methods and their applicability to real-time systems. Although the computational cost issue has to some extent been addressed by Karjalainen et al. (1997) and Bank (2008), no one seems to have taken the final step from simulation and theory to practice in any of these publications.

It would also be interesting to confirm the perceptual superiority of the presented techniques in listening tests. So far, it is just assumed that the presented perceptual models (e.g logarithmic pole placing) sound better. Obviously, these models make sense and should work. Some reports of informal listening test exist (Bank, 2007), but there still seems to be room for a more valid large-scale listening test.

5 Conclusions

Perceptually motivated filtering techniques attempt to model the frequency resolution of the human ear, which is known to be non-linear by nature. As a result, these filters concentrate their efforts at low frequencies at the expense of high frequencies. Consequently, lower model orders are sufficient for perceptually motivated filters in comparison to linear scale FIR or IIR filters.

In this paper, three central perceptually motivated filtering techniques were introduced. These were the warped, Kautz and Bank techniques. The two latter can also be divided into parallel second order IIR sections, which in theory enables efficient filter realizations.

All presented methods were implemented and simulated in MATLAB. The methods were compared on two metrics - conformance to a given magnitude specification and theoretical filtering time. The WIIR filter was found to have the best performance in the first task, as it provided the lowest RMSE on a logarithmic frequency scale over a wide range of model orders. This is most likely due to it not using the same (simpler) pole selection technique as the Kautz and Bank filters, which had the second lowest fitting error. In the second task, warped filtering was found to be the computationally most demanding task, whereas Kautz and Bank filters were the most efficient. Theory suggests that parallel IIR filter structures could reach even faster computational speeds than linear scale filters, provided they are implemented using the right architecture (i.e. parallel processors). Also, perceptual (and in particular parallel) IIR filters are known to remain stable at higher filter orders than traditional (direct form) IIR filters (Karjalainen et al. (1997), Bank (2007)).

6 References

B. Bank. Direct design of parallel second-order filters for instrument body modeling. In *Proc. International Computer Music Conference (ICMC 2007), Copenhagen, Denmark*, pages 458–465, 2007.

B. Bank. Perceptually motivated audio equalization using fixed-pole parallel second-order filters. *Signal Processing Letters, IEEE*, 15:477–480, 2008.

B. Bank. Audio equalization with fixed-pole parallel filters: An efficient alternative to complex smoothing. In *Audio Engineering Society Convention 128*. Audio Engineering Society, 2010.

B. Bank. Loudspeaker and room response equalization using parallel filters: comparison of pole positioning strategies. In *Audio Engineering Society Conference: 51st International Conference: Loudspeakers and Headphones*. Audio Engineering Society, 2013.

B. Bank and G. Ramos. Improved pole positioning for parallel filters based on spectral smoothing and multiband warping. *Signal Processing Letters, IEEE*, 18(5):299–302, 2011.

A. Härmä, M. Karjalainen, L. Savioja, V. Välimäki, U. K. Laine, and J. Huopaniemi. Frequency-warped signal processing for audio applications. *Journal of the Audio Engineering Society*, 48(11):1011–1031, 2000.

M. Hayes. *Statistical digital signal processing and modeling*. John Wiley & Sons, 1996. ISBN 9780471594314.

M. Karjalainen and T. Paatero. Equalization of loudspeaker and room responses using Kautz filters: Direct least squares design. *EURASIP Journal on Advances in Signal Processing*, 2007, 2006.

M. Karjalainen, A. Harma, and U. K. Laine. Realizable warped IIR filters and their properties. In *1997 IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP-97.*, volume 3, pages 2205–2208. IEEE, 1997.

J. Merimaa, T. Peltonen, T. Lokki, and A. O. C. Engineers. Concert hall impulse responses Pori, Finland: Reference. *Available online at:* `http://www.acoustics.hut.fi/projects/poririrs` *[Last viewed 28-Oct-2013]*, 2005.

T. Paatero and M. Karjalainen. Kautz filters and generalized frequency resolution: Theory and audio applications. *Journal of the Audio Engineering Society*, 51(1/2):27–44, 2003.

M. Waters and M. B. Sandler. Least squares IIR filter design on a logarithmic frequency scale. In *IEEE International Symposium on Circuits and Systems, ISCAS'93*, pages 635–638. IEEE, 1993.