MINIMUM PERCEPTUAL SPECTRAL DISTANCE FIR FILTER DESIGN [1]

Shao-Po Wu[2] *and* *William Putnam*[3]

[2]Information Systems Laboratory, Stanford University
[3]Center for Computer Research in Music and Acoustics, Stanford University
Stanford, CA 94305
clive@isl.stanford.edu, putnam@ccrma.stanford.edu

ABSTRACT

This paper addresses the problem of designing finite impulse response filters which optimally approximate desired frequency responses in the sense that they minimize a perceptual audio spectral measure. This measure is based on a simplified auditory model similar to those used in the area of perceptual audio quality measurement. It is shown that this problem can be cast as a logarithmic Chebychev approximation problem, which can be solved efficiently using recent interior point methods.

1. INTRODUCTION

A *finite impulse response* (FIR) filter is defined by the input-output relation

$$y(t) = \sum_{k=0}^{n-1} h(k)u(t-k), \quad t \in \mathbb{Z}$$

where $u : \mathbb{Z} \to \mathbb{R}$ is the input signal and $y : \mathbb{Z} \to \mathbb{R}$ is the output signal. The filter *order* is n, and $h = \big(h(0), h(1), \ldots, h(n-1)\big) \in \mathbb{R}^n$ are the filter *coefficients*. The filter *frequency response* $H : \mathbb{R} \to \mathbb{C}$ is defined as

$$H(\omega) = h(0) + h(1)e^{-j\omega} + \cdots + h(n-1)e^{-j(n-1)\omega}.$$

Since H is 2π periodic and satisfies $H(-\omega) = \mathbf{conj}\big(H(\omega)\big)$, it suffices to specify it over the interval $\omega \in [0, \pi]$.

In this paper we consider designing an FIR filter that best approximates a desired frequency response *perceptually*, *i.e.*, given a test signal, the audible difference between the filtered output and the desired output is minimized. We will use a *perceptual spectral distance* (PSD) as an objective method of evaluating the audible difference between the resultant signals. The model used is similar to those used in the area of perceptual audio quality measurement; see [10], [6], [4] and [2]. Specifically, the filter design problem is posed as an optimization problem in which the PSD is minimized. Furthermore, this problem can be expressed as a logarithmic Chebychev approximation problem.

[1]Research supported in part by AFOSR (under F49620-95-1-0318), NSF (under ECS-9222391 and EEC-9420565), and MURI (under F49620-95-1-0525).

2. PERCEPTUAL MODEL AND PSD

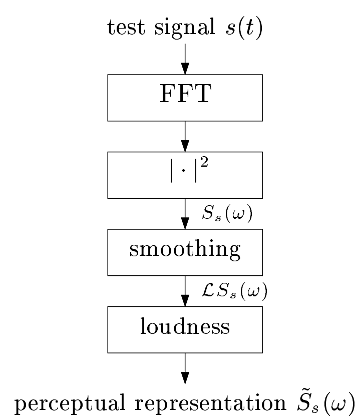

Figure 1: Simplified auditory model used to determine perceptual spectral distance

Auditory models have been used to model the perceptual representations and audible differences between signals. A simplified perceptual model is used which consists of the following operations. Given a signal $s(t)$, $t \in \mathbb{Z}$, the power spectral density, $S_s(\omega)$ is computed by taking the squared magnitude of the Fourier transform of the signal. It is then smoothed by a frequency dependent smoothing function f_ω:

$$\mathcal{L}S_s(\omega) = \int_0^\pi S_s(\mu) f_\omega(\mu) d\mu, \tag{1}$$

This models the critical bands of the auditory system [10], and is denoted by $\mathcal{L}$. The last stage of the model compresses the smoothed power spectrum on a dB scale to model loudness [6]. Combining these operations results in the perceptual representation of $s(t)$:

$$\tilde{S}_s(\omega) = \log \mathcal{L}S_s(\omega) = \log \int_0^\pi S_s(\mu) f_\omega(\mu) d\mu.$$

The perceptual spectral distance between two signals $s_1(t)$ and $s_2(t)$ over the frequency range $\Omega \in [0, \pi]$ is defined as:

$$\mathbf{psd}_\Omega(s_1, s_2) \triangleq \max_{\omega \in \Omega} |\tilde{S}_{s_1}(\omega) - \tilde{S}_{s_2}(\omega)|,$$

which gives the maximum error between the perceptual representations of the two signals. In this manner, the PSD can be used as a measure of the *perceptual difference* between two signals.

3. MINIMUM PSD DESIGN

3.1. Problem formulation

Given a test signal $s(t)$ and a desired frequency response $D(\omega)$ (with impulse response $d(t)$), we would like to design an n-tap FIR filter with coefficients $h \in \mathbb{R}^n$ that minimize the perceptual spectral distance between the desired output and the filtered output over Ω, *i.e.*,

$$\text{minimize} \quad \mathbf{psd}_\Omega \left((s * d)(t), (s * h)(t) \right), \tag{2}$$

where $*$ denotes convolution.

Unfortunately (2) is not a convex optimization problem in h and is intractable in general. However, it can be reformulated as a convex problem via proper change of variables shown as follows.

Define the *autocorrelation coefficients* associated with h as:

$$r(t) = \sum_{k=-n+1}^{n-1} h(k) h(k+t), \quad t \in \mathbb{Z}. \tag{3}$$

Since $r(t) = r(-t)$, and $r(t) = 0$ for $t \geq n$, it is sufficient to specify the correlation coefficients for $t = 0, \ldots, n-1$: $r = \big(r(0), \ldots, r(n-1)\big) \in \mathbb{R}^n$. The Fourier transform of $r(t)$:

$$R(\omega) = \sum_{k=-n+1}^{n-1} r(k) e^{-jk\omega} = |H(\omega)|^2 \tag{4}$$

is the magnitude squared of $H(\omega)$.

The power spectral densities of $(s*h)(t)$ and $(s*d)(t)$ are given by $R(\omega)S_s(\omega)$ and $|D(\omega)|^2 S_s(\omega)$. Using r as the design variable, and applying equations (1) and (4), (2) can be reformulated as

$$\begin{array}{ll} \text{minimize} & \\ & \left|\log \mathcal{L}R(\omega)S_s(\omega) - \log \mathcal{L}|D(\omega)|^2 S_s(\omega)\right| \\ \text{subject to} & R(\omega) \geq 0, \quad \omega \in [0, \pi], \end{array} \tag{5}$$

By the spectral factorization theorem, the extra constraint $R(\omega) \geq 0$ provides a necessary and sufficient condition for r to be a valid autocorrelation function corresponding to a $h \in \mathbb{R}^n$. [8, p.231]. Since $\mathcal{L}R(\omega)S_s(\omega)$ is a linear function of r, the problem (5) turns out to be a logarithmic Chebychev approximation problem on the smoothed spectral density. Moreover, it is a convex optimization problem.

Without loss of generality, assume the test signal $s(t)$ is white noise with $S_s(\omega) = 1$, $\omega \in [0, 2\pi)$. It is shown in [11] and [12] that (5) can be cast as

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \dfrac{\mathcal{L}|D(\omega)|^2}{t} \leq \mathcal{L}R(\omega) \leq t\mathcal{L}|D(\omega)|^2, \quad \omega \in \Omega \\ & R(\omega) \geq 0, \quad \omega \in [0, \pi], \end{array} \tag{6}$$

where $t \in \mathbb{R}$ and $r \in \mathbb{R}^n$ are the variables.

Given a solution of (6), there exists at least one $h \in \mathbb{R}^n$ that satisfies (3). Such an h can be obtained via spectral factorization of r. Many algorithms are available to perform this factorization. For a specific example, see [8].

3.2. A quadratic programming approach

A good engineering approximation of (6) is to discretize the problem by sampling the frequency domain, *i.e.*, imposing the constraints on a finite subset of $[0, \pi]$, ω_i, $i = 0, \ldots, m$. Additionally, sampling can be done on a non-uniform grid to model the warped nature of the bark frequency scale [2]. A rule of thumb in choosing m is $m \approx 15n$ [1].

Doing this, the problem becomes

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \dfrac{\mathcal{L}|D(\omega_i)|^2}{t} \leq \mathcal{L}R(\omega_i) \leq t\mathcal{L}|D(\omega_i)|^2, \; \omega_i \in \Omega \\ & R(\omega_i) \geq 0, \quad i = 0, \ldots, m. \end{array} \tag{7}$$

Since $\mathcal{L}R(\omega_i)$ is a linear function in r for each i, (7) is a quadratic programming problem and can be solved very efficiently using the interior-point methods described in [7] and [11].

Because of discretization, the non-negativity constraint cannot be imposed for all $\omega \in [0,\pi]$, thus one cannot ensure the existence of h such that (4) holds. Several sub-optimal techniques have been proposed for this case. For example, offset $R(\omega)$ obtained via solving (6) to make it non-negative [5][9]. Another approach is to replace the non-negativity constraint by a more strict one, *i.e.*, $R(\omega_i) \geq \epsilon > 0, \quad i = 0, \ldots, m$.

3.3. A semidefinite programming approach

The constraint $R(\omega) \geq 0, \omega \in [0,\pi]$, can be imposed as a matrix inequality *exactly* at the cost of introducing some auxiliary variables using the following theorem [3, CH2.7.2].

Theorem 1 (positive-real) *Given a discrete-time linear system (A,B,C,D), A stable, and (A,B,C) minimal. The transfer function*

$$G(z) = C(zI - A)^{-1}B + D$$

satisfies

$$G(e^{j\omega}) + G(e^{j\omega})^* \geq 0 \quad \textit{for all} \quad \omega \in [0, 2\pi)$$

if and only if there exists real symmetric matrix P such that the matrix inequality

$$\begin{bmatrix} P - A^TPA & C^T - A^TPB \\ C - B^TPA & D + D^T - B^TPB \end{bmatrix} \geq 0 \tag{8}$$

is satisfied.

To apply Theorem 1, we would like to find (A,B,C,D) in terms of r such that

$$\begin{aligned} &C(zI-A)^{-1}B + D \\ &\quad = \frac{1}{2}r(0) + r(1)z^{-1} + \cdots + r(n-1)z^{-(n-1)}. \end{aligned} \tag{9}$$

An obvious choice is the controllability canonical form:

$$A = \begin{bmatrix} 0 & 0 & & \cdots & 0 \\ 1 & 0 & & \cdots & 0 \\ & 1 & & & \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & & & 1 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$C = \begin{bmatrix} r(1) & r(2) & \cdots & r(n-1) \end{bmatrix} \quad D = \frac{1}{2}r(0).$$

The realization is not unique, for example, $(T^{-1}AT, T^{-1}B, CT, D)$ realizes the same transfer function for non-singular state transformation T.

It can be easily checked that the (A,B,C,D) given satisfies all the hypotheses of Theorem 1. Therefore the existence of r and P which satisfy the matrix inequality (8) is necessary and sufficient for the non-negativity of $R(\omega)$, $\omega \in [0, 2\pi)$. We can pose (6) as the semidefinite programming problem (SDP, see [3] and [11]) in $r \in \mathbb{R}^n$, $t \in \mathbb{R}$, and $P = P^T \in \mathbb{R}^{(n-1)\times(n-1)}$:

$$\begin{aligned} &\text{minimize} \quad t \\ &\text{subject to} \quad \frac{\mathcal{L}|D(\omega_i)|^2}{t} \leq R(\omega_i) \leq t\mathcal{L}|D(\omega_i)|^2, \ \omega_i \in \Omega \\ &\qquad \begin{bmatrix} P - A^TPA & C^T - A^TPB \\ C - B^TPA & D + D^T - B^TPB \end{bmatrix} \geq 0, \end{aligned} \tag{10}$$

and its solution is guaranteed to be factorizable.

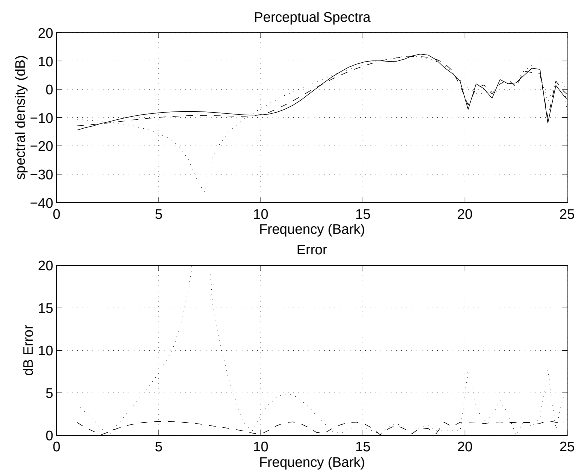

Figure 2: Perceptual spectral representation and approximation error

4. EXAMPLE

Figure 2 depicts the results of the design technique described in Section 3. In this example, a 50-tap impulse response is approximated by one of length 20. The solid line is the desired power spectrum, the dashed line is the minimum PSD design, and the dotted line is the design obtained using the Remez exchange algorithm. The second plot in Figure 2 shows the error present in each of the approximations. In this case, the perceptual optimization technique results in a worst case error less than 1.75 dB.

5. CONCLUDING REMARKS

The technique presented in this paper is for designing an FIR filter which best approximates a desired response in a perceptual sense. In the formulation, the infinity norm is chosen to measure the distance between the perceptual representations of signals. In the case where the error is below a threshold of perceptibility, it is the appropriate choice since it makes maximum use of the available degrees of freedom. The technique can also be applied to design the minimum order FIR filter subject to an upper bound on the PSD. This can be done by bisecting the interval of possible filter orders, and solving a feasibility problem at each iteration until the minimum order is found.

References

[1] A. S. Alkhairy, K. F. Christian, and J. S. Lim. Design of FIR filters by complex Chebyshev approximation. *IEEE Trans. Signal Processing*, 41:559–572, 1993.

[2] J. G. Beerends and J. A. Stemerdink. A perceptual audio quality measure based on a psychoacoustic sound representation. *J. Audio Eng. Soc.*, 40(12):963–978, 1992.

[3] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory*, volume 15 of *Studies in Applied Mathematics*. SIAM, Philadelphia, PA, June 1994.

[4] K. Brandenburg and T. Sporer. NMR and masking flag: Evaluation of quality using perceptual criteria. In *Proceedings 11th Intl. Conference on Audio Test and Measurement*, pages 169–179, 1992.

[5] J. O. S. III. *Techniques for Digital Filter Design and System Identification with Application to the Violin*. PhD thesis, Stanford, California, 1983.

[6] M. Karajalainen. Sound quality measurements of audio systems based on models of auditory perception. In *Proceedings of IEEE ICASSP*, pages 19.9.1–19.9.4, 1984.

[7] Y. Nesterov and A. Nemirovsky. *Interior-point polynomial methods in convex programming*, volume 13 of *Studies in Applied Mathematics*. SIAM, Philadelphia, PA, 1994.

[8] A. Papoulis. *Signal Analysis*. McGraw-Hill, New York, 1977.

[9] T. Parks and C. Burrus. *Digital Filter Design*. J. Wiley, New York, 1987.

[10] M. R. Schroeder, B. S. Atal, and J. L. Hall. Optimizing digital speech coders by exploiting masking properties of the human ear. *Journal of the Acoustical Society of America*, 66(6):1647–1652, 1979.

[11] L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM Review*, 38(1):49–95, Mar. 1996.

[12] S.-P. Wu, S. Boyd, and L. Vandenberghe. FIR filter design via spectral factorization and semidefinite programming. In *Proc. IEEE Conf. on Decision and Control*, 1996. To appear.