INVERSE FILTER DESIGN USING MINIMAX APPROXIMATION TECHNIQUES FOR 3-D AUDIO

Harsha .I.K Rao*, V. John Mathews

Department of Electrical & Computer Engineering University of Utah Salt Lake City, UT 84112, USA

ABSTRACT

This paper presents a novel approach for implementing immersive audio rendering filters for a single listener using loudspeakers. We address the problem of crosstalk cancellation inherent in loudspeaker rendering and propose to implement the crosstalk cancellation filters using minimax finite impulse response (FIR) filters. The formulation is based on the Atal-Schroeder crosstalk canceller. The use of the optimal FIR filter design procedure ensures significant amount of separation between the direct path and the cross path. An alternative topology which requires the approximation of just one filter has also been explored using the same design principles. The minimax techniques provides superior solutions as compared to a leastsquares design and the alternate structure is shown to be robust in its performance.

1. INTRODUCTION

A 3-dimensional audio system can be used to position sounds around a listener [1], [2]. There has been a tremendous amount of interest in 3-D audio technology in the recent past. Its applications include use on multimedia desktop computers. A direct approach to implement a 3-D audio system is to use headphones. Headphones have several advantages such as excellent channel separation, invertibilty of the transmission paths from the transducers to the ears and the ability to isolate the listener from external sounds and reverberations. However, they also suffer from some disadvantages. In most applications, non-individualized head-related transfer functions (HRTFs) are used and headphones have difficulty in creating external frontal images while synthesizing the directional cues [1], [2]. That is, headphones induce errors in front and rear sound perception and are limited in their ability to externalize sound. This tends to give the listener an "inside the head" sensation. Finally, in applications such as videoconferencing, headphones might become uncomfortable to use.

A key issue that should be addressed while reproducing sounds to a listener is that the left ear signal must go to the listener's left ear only and similarly the right ear signal must only go to the listener's right ear. It is essential that there is no crosstalk, *i.e.*, the left and right channels remain separated. To mitigate crosstalk effects, we can create a set of filters which will in combination with the acoustic paths cancel the crosstalk signals at the ears [3].

In this paper, we present the method for designing the crosstalk cancellation filters using minimax techniques. The crosstalk cancellation filters are modeled using FIR filters in our work [4], [5]. Young-Cheol Park

Computer & Telecommunication Engineering Division Yonsei University Wonju, Korea

We show that an approach based on the second-order cone programming (SOCP) [5] technique provides excellent approximation to the minimax solution. The design problem is formulated as a convex optimization problem and solved using interior point solvers.

The classic Atal-Schroeder crosstalk canceler [3], [6] is shown in Fig. 1 for the case of two loudspeakers. In this figure, the left



Fig. 1. Atal-Schroeder crosstalk canceller

and right signals are p_L and p_R respectively, l_1 and l_2 are the loudspeaker signals and h_i (ipsilateral term) and h_c (contralateral term) are the head-related impulse responses (HRIRs) to the same-side and opposite-side ears, respectively. The objective is to design the filters h_1 , h_2 , h_3 and h_4 such that the crosstalk signals are cancelled, *i.e.*, none of the p_L signal is received at the right ear and similarly for the p_R signal and the left ear. Assuming that the system is acoustically symmetric so that the transfer functions from the left loudspeaker to the ears are the same as those for the right loudspeaker, it can be shown that the required filter responses in terms of the HRTFs, $H_i(\omega)$ and $H_c(\omega)$ are given by:

$$H_{1}(\omega) = \frac{H_{i}(\omega)}{H_{i}^{2}(\omega) - H_{c}^{2}(\omega)}$$

$$H_{2}(\omega) = \frac{-H_{c}(\omega)}{H_{i}^{2}(\omega) - H_{c}^{2}(\omega)}$$

$$H_{3}(\omega) = H_{2}(\omega)$$

$$H_{4}(\omega) = H_{1}(\omega)$$
(1)

If the listener is in an asymmetric position with respect to the loudspeakers, the above equations can be modified appropriately.

The frequency responses of the four crosstalk canceller filters

^{*}Email: hrao@eng.utah.edu

can be expressed in matrix form as

$$H(\omega) = \frac{1}{H_i^2(\omega) - H_c^2(\omega)} \begin{bmatrix} H_i(\omega) & -H_c(\omega) \\ -H_c(\omega) & H_i(\omega) \end{bmatrix}.$$
 (2)

This results in a slightly different realization of the crosstalk filters, as shown in Fig.2. This realization requires the design of only one filter for its implementation. This topology was suggested in [7].



Fig. 2. Single filter structure of crosstalk canceler

The desired frequency response of the single filter is given by

$$H_d(\omega) = \frac{1}{H_i^2(\omega) - H_c^2(\omega)}$$
(3)

Under ideal conditions, *i.e.*, when the exact room impulse responses are used to implement the filters in the matrix, crosstalk will be completely cancelled irrespective of the choice of the single filter that must be designed for this topology. In general, $H_d(\omega)$ has infinitelength impulse response. However, we will approximate this response using an FIR filter in this work. The most common way to design the crosstalk cancellation filters is by using the least-squares (LS) optimization technique [6]. In this paper, we compare the capabilities of the minimax design to those of the LS design for both the Atal-Schroeder structure and the single filter structure in [7]. Our results indicate that the minimax design offers higher crosstalk cancellation as compared to the LS method. Furthermore, the alternate structure of [7] appears to provide superior solutions to the crosstalk cancellation problem.

The rest of the paper is organized as follows. In Section 2, we discuss the design of crosstalk cancellation filters in the minimax sense. Experimental results comparing the design techniques and the two structures are given in Section 3. Finally, our conclusions are provided in Section 4. Throughout the paper, we have denoted vectors and matrices using bold-faced characters, \Re refers to the set of real numbers and C^q refers to a second-order cone of dimension q.

2. MINIMAX APPROXIMATION OF CROSSTALK CANCELLATION FILTERS

Given the desired crosstalk canceller frequency responses as in (1) or (3), we can consider a general FIR filter design problem. Let $H(\omega)$ represent the frequency response of an *N*-tap FIR digital filter as given by

$$H(\omega) = \sum_{n=N_1}^{N_2} h(n) e^{-j\omega n}; \ N = N_2 - N_1 + 1, \qquad (4)$$

where $\{h(n)\}_{n=N_1}^{N_2}$ represent the filter coefficients. Let $D(\omega)$ be the desired complex valued frequency response of the crosstalk filter on a compact subset *B* of the frequencies in $[-\pi,\pi)$. The minimax approximation problem consists of finding the filter coefficients that will minimize the Chebyshev norm

$$||E(\omega)|| = \max_{\omega \in B} |E(\omega)|$$
(5)

where the weighted error is given by

$$E(\omega) = W(\omega)[D(\omega) - H(\omega)].$$
(6)

In the above equation, $W(\omega)$ is a real, positive and continuous weighting function in the frequency domain. When approximating arbitrary magnitude responses, the choice of the weighting function can play a crucial role in the characteristics of the optimization algorithm. This problem can be solved using the technique given in [4], which employs the alternation theorem to iteratively find the minimax solution. While this algorithm is easy to implement and has relatively fast convergence properties, there are situations in which it may converge only on a subset of B. When this occurs, we may need to use a second stage of optimization to further refine the solution. One of the approaches to solve this problem has been incorporated in MATLAB as the second stage of the *cfirpm* function for minimax optimization [8]. The numerical implementation of the crosstalk cancellation filters using this algorithm is computationally intensive and in several of our simulations, the algorithm failed to converge.

To avoid this problem and to obtain filters optimized under the weighted Chebyshev norm, we propose the use of second-order cone programming (SOCP) to design the crosstalk cancellation filters in the minimax sense. Recently Yan, *et al.* [5] provided a unified framework for the use of the SOCP technique to design FIR filters using various measures of norms. SOCP provides us with a flexible structure to formulate the FIR filter design problem using frequency domain specifications and the optimization problem can be solved using interior point solvers such as those included in the SeDuMi toolbox of MATLAB [9]. In order to see how the SOCP approach works, we reformulate (5) and (6) in terms of the filter coefficients h(n) so as to minimize the L_{∞} norm of the error given by

$$E(\omega_k) = \max_{\omega_k \in B} (W(\omega_k) | D(\omega_k) - \mathbf{e}^T(\omega_k) \mathbf{h} |); \ k = 1, \cdots, K$$
(7)

where $\mathbf{h} = [h(0), h(1), \dots, h(N-1)]^T$, $\mathbf{e}(\omega) = [1, e^{-j\omega}, \dots, e^{-j\omega(N-1)}]^T$ and the optimization is performed over *K* frequency points. The above optimization problem can be transformed as

$$\min_{\mathbf{h}} \delta \tag{8}$$

subject to
$$W(\omega_k)|D(\omega_k) - \mathbf{e}^T(\omega_k)\mathbf{h}| \le \delta; \ k = 1, \cdots, K$$

We solve the K second-order cone constraints for **h**, using SeDuMi, by redefining the conic problem as

max
$$\mathbf{b}^T \mathbf{y}$$
, subject to $\mathbf{c} - \mathbf{A}^T \mathbf{y} \in Q_{cone}$ (9)

where $\mathbf{y} = [\delta, \mathbf{h}^T]^T$, $\mathbf{b} = [-1, \mathbf{0}_{1 \times N}]^T$, such that $\delta = -\mathbf{b}^T \mathbf{y}$,

$$\mathbf{c} = \begin{bmatrix} 0 \\ W(\omega_k)D(\omega_k) \end{bmatrix}, \mathbf{A} = \begin{bmatrix} -1 & \mathbf{0}_{1 \times N} \\ 0 & W(\omega_k)\mathbf{e}^T(\omega_k) \end{bmatrix}$$

and Q_{cone} is a symmetric cone of dimension q that can be defined as

$$Q_{cone} \equiv \left(\begin{bmatrix} x_1 \\ \mathbf{x_2} \end{bmatrix} : x_1 \in \Re, \mathbf{x_2} \in C^{q-1}, x_1 \ge \|\mathbf{x_2}\| \right) \,.$$

Before proceeding, we define the terms equalization and cancellation in the context of crosstalk cancellation. Equalization is a measure of the signal coming the direct path and has an ideal value of 0 dB. Cancellation is a measure of the signal arriving along the cross path and should ideally be $-\infty$ dB.

We implemented the crosstalk cancellation filters using the two structures discussed in Section 1. The first structure basically designs the two filters having frequency responses as given in (1) and the second structure designs the filter having frequency response as given in (3).

2.1. Selection of the weighting function

In all our designs, we chose the weighting function as

$$W(\omega) = \frac{PR(\omega)}{|D(\omega)|^p} + \frac{Q(1 - R(\omega))}{|D(\omega)|^q}, \qquad (10)$$

where

$$R(\omega) = \begin{cases} 1; & |D(\omega)| \ge \lambda \\ 0; & |D(\omega)| < \lambda \end{cases}$$

P and *Q* are arbitrary positive constants that were chosen to be unity in our work and $D(\omega)$ is the desired frequency response of the filter being designed. This choice was based on the perspective gained from extensive experimentation. The basic idea is to give relatively more weight to the frequency response if its magnitude is less than a threshold λ (we choose λ =1 in our work). The parameters *p* and *q* were chosen such that $2.5 \le p \le 4.5$, $1.5 \le q \le 3.5$ and p > q.

3. EXPERIMENTAL RESULTS

A number of simulations were conducted to evaluate the performance of the crosstalk canceller, designed using the minimax algorithms for both the conventional structure and the single filter structure. The crosstalk cancellers were implemented using 256-tap FIR filters. The impulse responses used in our simulations had 128 coefficients and were taken from the set of HRTFs measured using a Knowles Electronic Manikin for Acoustic Research (KEMAR) [10]. The systems were evaluated for different angles of elevations and azimuth and the results are tabulated in Table 1. The complex Chebyshev algorithm only gives the best approximate over a subset of the frequencies in all cases. The percentage of sample frequencies in B for which the algorithm converged to the "minimax solution" is denoted as "% of B". As mentioned before, SOCP technique can be used as an alternative and it can design filters having a minimum norm over the entire set B. The minimax filters were compared with the filters designed using a weighted least-squares (WLS) optimization criterion and the performance was evaluated through a measure of the mean crosstalk cancellation offered in B. The weighted error plots for the two filters for angle of elevation = 70° , angle of azimuth = 30° and the corresponding plots for the transfer functions of the direct path and the cross path are shown in Fig. 3 and Fig. 4 respectively. The plots and the results in Table 1 indicates that the maximum errors of the individual filters from the minimax design is considerably lower than those obtained through a WLS design. Also, the average cancellation across the entire frequency grid offered by the minimax filters is lower than that offered by the WLS filters. We note that none of the algorithms performed well at very low frequencies. However, this may not be a problem in many applications since the frequencies of interest are higher than the range over which the crosstalk canceller performs poorly. An important parameter involved in the design process is the modeling delay. On extensive

experimentation, satisfactory results were obtained for delays in the range 60-80 samples.



Fig. 3. Weighted errors of the crosstalk cancellation filters - elevation= 70° , azimuth= 30° (minimax filters - dashed lines, WLS filters - solid lines)



Fig. 4. Direct path and cross path transfer functions - elevation= 70° ,azimuth= 30° (minimax design - thick lines, WLS design - thin lines)

The results for the single filter topology are tabulated in Table 2. The evaluation criterion in this case was the measure of the mean deviation of the achieved equalization from unity, *i.e.*, mean($||T(\omega)|-1|$). The direct path transfer functions comparing the minimax design with the WLS design is displayed in Fig. 5. Though this filter was comparatively more difficult to approximate using an FIR model due to its inherent all-pole structure, it can be seen that in general, the minimax system offers better equalization as compared to a crosstalk canceller system implemented using the WLS filters.

To conclude our analysis of the minimax crosstalk canceller filters, we studied the robustness of the two crosstalk cancellation systems experimentally. The room acoustics was simulated with the HRTF measurements from [10]. Two statistically independent and uncorrelated sequences were applied as loudspeaker inputs and the elements of the room transfer function matrix were identified using a least-mean-squares (LMS) algorithm. The inverse filters were then designed using this estimated set of room impulse responses. The performance of the filters as shown in Fig. 6 indicates that the single filter design is more robust as it offers better crosstalk cancellation and the "overlapping" near dc frequencies is absent in most cases.

Table 1. Performance metrics for a conventional crosstalk canceller design

(Elevation,	Filter	Delay	Chebyshev approximation			SOCP	WLS	Mean minimax	Mean WLS
Azimuth)	type		$\mid \delta \mid$	max.error	% of <i>B</i>	max. error	max. error	crosspath (dB)	crosspath (dB)
$(70^{\circ}, 30^{\circ})$	H_1	80	0.0046	0.0144	93%	0.0046	0.3650	-123.7076	-116.7222
	H_2	79	0.0162	0.0440	95%	0.0162	0.1161		
$(70^{\circ}, 45^{\circ})$	H_1	59	0.0062	0.0082	99 %	0.0062	0.0340	-117.2189	-113.3327
	H_2	60	0.0222	0.0514	85%	0.0225	0.1235		

Table 2. Performance metrics for a single filter crosstalk canceller design

(Elevation,	Filter	Delay	Chebyshev approximation			SOCP	WLS	Mean minimax	Mean WLS
Azimuth)	type		$ \delta $	max. error	% of <i>B</i>	max. error	max. error	$(\mathbf{T}(\omega) -1)$	$(T(\omega) -1)$
$(70^{\circ}, 45^{\circ})$	Η	80	0.0352	0.1113	98%	0.0353	0.8795	0.0252	0.0604
$(0^{o}, 10^{o})$	Н	80	0.1431	0.3941	99 %	0.1432	0.1888	0.1023	0.1937



Fig. 5. Direct path transfer functions for the single filter structure - elevation= 70° , azimuth= 45° (minimax design on top, WLS design below)

4. CONCLUSIONS

We have proposed a novel method to design the crosstalk cancellation filters for a conventional system as well as a single filter topology using minimax techniques. While the complex Chebyshev approximation is a fast algorithm and it achieves near optimal solutions, a more accurate and refined solution can be obtained through a SOCP formulation. We also showed that the single filter structure is more robust to errors in estimation of the room impulse responses as it can offer better cancellation as compared to the conventional design. Based on the results of our evaluations, we believe that the method of this paper offers an attractive alternative to crosstalk canceller design techniques available in the literature.

5. REFERENCES

- [1] W.G. Gardner, *3-D Audio Using Loudspeakers*, Kluwer Academic, Norwell, MA, 1997.
- [2] C. Kyriakakis, "Fundamental and technological limitations of immersive audio systems," *Proc. of the IEEE*, vol. 86, no. 5, pp. 941–951, May 1998.
- [3] B.S. Atal and M.R. Schroeder, "Apparent sound source translator," US Patent 3236949, Feb 1966.



Fig. 6. Direct path and cross path transfer functions evaluated using estimated room impulse responses for elevation= 70° ,azimuth= 45° (conventional design - thin lines, single filter design - thick lines)

- [4] L.J. Karam and J.H McClellan, "Complex Chebyshev approximation for FIR filter design," *IEEE Trans. on Circuits and Systems-II*, vol. 42, no. 3, pp. 207–216, Mar 1995.
- [5] S.F. Yan and Y. Ma, "A unified framework for designing FIR filters with arbitrary magnitude and phase response," *Digital Signal Processing*, vol. 14, pp. 510–522, Dec 2004.
- [6] D.B Ward, "Joint least squares optimization for robust acoustic crosstalk cancellation," *IEEE Trans. on Speech and Audio Processing*, vol. 8, no. 2, pp. 211–15, Feb 2000.
- [7] C. Bourget and T. Aboulnasr, "Inverse filtering of room impulse response for binaural recording playback through loudspeakers," *Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing*, vol. III, pp. 301–304, April 1994.
- [8] L.J. Karam and J.H. McClellan, "Chebyshev digital FIR filter design," *Signal Processing*, vol. 76, no. 1, pp. 17–36, 1999.
- [9] J.F. Strum., "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optimization Methods Software*, vol. 11-12, pp. 625–653, 1999.
- [10] W.G. Gardner, "HRTF measurements of a KEMAR dummyhead microphone," *MIT Media Lab Perctual Computing: Technical Report*, , no. 280, May 1994.