Design and Implementation of Discrete-Time Filters For Efficient Rate-Conversion Systems

Thomas A. Baran and Alan V. Oppenheim
Massachusetts Institute of Technology
Digital Signal Processing Group
77 Massachusetts Avenue, Cambridge, MA 02139
tbaran@mit.edu, avo@mit.edu

Abstract—Oversampled A-to-D converters commonly rely on sharp-cutoff, discrete-time filters that operate at fast input sampling rates. Filter design techniques for such filters typically use the length of the impulse response as an indicator of computational cost, assuming that each filter tap requires a multiplier.[5][6] This paper describes methods for designing zero-phase filters with sparse impulse responses, i.e., with many zero-valued coefficients, and presents rate-conversion structures for efficiently implementing these designs. By combining polyphase methods and delay-line folding, the structures presented require only one multiplier for each unique value in the impulse response of the filter.

I. INTRODUCTION

Techniques to reduce the computational cost of rate-conversion systems have, broadly speaking, occurred on two fronts. One such effort concentrates on improvements in flow graph structures; key results include polyphase implementations and folded delay line structures with time-symmetric impulse responses.[1][2][3][4][8] Another effort involves using filter design techniques, including the Parks-McClellan algorithm and the METEOR toolkit, as methods for choosing efficient filters from some set of permissible designs.[5][6] In this paper, we consider the problem of designing zero-phase filters for which the impulse response is sparse, i.e., where $||h[n]||_0$ is small, and propose structures that require fewer nonzero multiplications per unit time than either polyphase or folded implementations alone. The proposed techniques are capable of designing filters with smaller maximum error and greater sparsity than Parks-McClellan designs. While this paper focuses on downsampling systems, equally-efficient upsampling structures are obtained from these by invoking the Generalized Transposition Theorem[2].

II. DESIGN OF SPARSE ZERO-PHASE FILTERS

Many rate-conversion systems employ filters with linear phase; we focus on techniques for designing this class of filters. In particular, our emphasis is on methods for designing zero-phase filters with sparse impulse responses. Efficient implementations of these designs are made by omitting multipliers which correspond to the zero-valued coefficients. We

will therefore use the sparsity of a filter impulse response $h[n]$ as an indicator of its computational cost. Formally, the sparsity will be stated in terms of $||h[n]||_0$, the number of time indices at which the impulse response is nonzero.

We are concerned with designing FIR filters that have symmetric impulse responses, i.e. $h[n]$ of the form

$$h[n] = b_0\delta[n] + \sum_{m=1}^{K} b_m \left(\delta[n-m] + \delta[n+m] \right),$$

and consequently frequency responses $H(e^{j\omega})$ that are real-valued. Denoting the desired frequency response as $H_d(e^{j\omega})$, the error is

$$E(e^{j\omega}) \equiv H(e^{j\omega}) - H_d(e^{j\omega}).$$

We will also refer to an error-weighting function, denoted $W(\omega)$.

The general problem of designing the most-sparse, zero-phase, FIR filter that meets a given set of constraints in the frequency domain is computationally difficult. We therefore consider two relaxations of the problem.

A. Minimax-optimal filters with constrained coefficients

One approach to designing sparse filters is to perform an optimization in which a specified subset N_x of coefficients is explicitly constrained to have value zero. The formulation of this problem is the same as that of the minimax-optimal filter design problem, with an additional constraint for each coefficient that is preset to zero:

$$\text{minimize } \max_{\omega \in F} \left| W(\omega)E(e^{j\omega}) \right| \text{ subject to } b_n = 0 \ \forall n \in N_x. \tag{1}$$

Equation 1 can be formulated in terms of a linear program and consequently has an efficient solution. In practice, this formulation is approximate, since $E(e^{j\omega})$ can only be evaluated on a finite grid. However, the evaluation may be performed over a reasonably-dense grid without incurring prohibitive computational cost.

The optimization in Eq. 1 is typically capable of resulting in filters with smaller maximum error and smaller $||h[n]||_0$ than a Parks-McClellan design with the same desired response, for a particular choice of N_x. However, other choices of N_x may yield filters with greater maximum error and greater

$||h[n]||_0$ than comparable Parks-McClellan designs. In general, the choice of N_x can have a significant impact on the error in the resulting filter.

B. Minimum-multiply sparse filters

As an alternative to pre-specifying the set N_x we can consider an algorithm to obtain the most-sparse filter, i.e. the filter with the smallest $||h[n]||_0$, that meets a set of frequency-domain constraints. However, this problem is computationally difficult. We therefore consider a computationally-tractable relaxation that is based upon minimizing $||h[n]||_1$. This optimization problem is formulated as

$$\text{minimize } ||h[n]||_1 \text{ subject to } \max_{\omega \in F} \left| W(\omega) E(e^{j\omega}) \right| \leq \delta. \quad (2)$$

$W(\omega)$ and δ specify the maximum permissible frequency domain error.

Approximating $||h[n]||_0$ with $||h[n]||_1$ is a practice that has been used in the fields of compressive sampling and convex optimization.[7] In the context of filter design, minimizing $||h[n]||_1$ typically results in filters which have some coefficients that are small in value. The resulting designs are made sparse by setting these small coefficients to zero. The overall design algorithm is:

(1) Solve Eq. 2.
(2) Set $h[n] = 0$ for all n where $|h[n]| < \Gamma$.

While this process designs filters that may not, in general, meet the original frequency domain specifications, this approach empirically seems to lead to designs in which the constraints are exceeded by an acceptably small amount.

III. EFFICIENT STRUCTURES

As is well-known, a folded delay line structure requires half as many multiplications per unit time as a direct-form implementation, since each multiplier computes the value of the impulse response for two time indices. In the context of rate-conversion systems, however, the canonical polyphase structure is often a more efficient choice. In this section we present a downsampling-by-M structure that incorporates both folding and polyphase and thus is more efficient than either by itself. A complementary upsampling structure may be obtained by invoking the Generalized Transposition Theorem.[2]

The proposed structure requires M compressors and one multiplier for each unique value taken on by the impulse response of the filter. Each multiplier operates at the output sampling rate of the system (the slower sampling rate) when implemented synchronously.

We first consider a structure for a zero-phase, FIR filter that utilizes one multiplier for each unique value taken on by its impulse response $h[n]$. This structure will then be combined with a polyphase structure to implement a downsampling system.

Consider the z-transform $H(z)$ of a zero-phase, FIR filter expressed in terms of the set $\mathcal{C}$ of unique values $c_k \in \mathcal{C}$ taken on by its impulse response $h[n]$:

$$H(z) = \sum_{k=0}^{\mathcal{U}-1} c_k \sum_{p=0}^{K} \frac{d_{k,p} \left(z^p + z^{-p} \right)}{1 + \delta[p]}, \quad d_{k,p} \in \{0, 1\}. \quad (3)$$

Here, the $d_{k,p}$ select the time indices of $h[n]$ to which the value c_k applies.

In Fig. 1 we illustrate a structure that implements Eq. 3. One multiplier is required for each value c_k, and $\mathcal{U}$ multipliers are needed in total. Note that $\mathcal{U} \leq (L + 1)/2$, where L is the length of the impulse response of the filter.
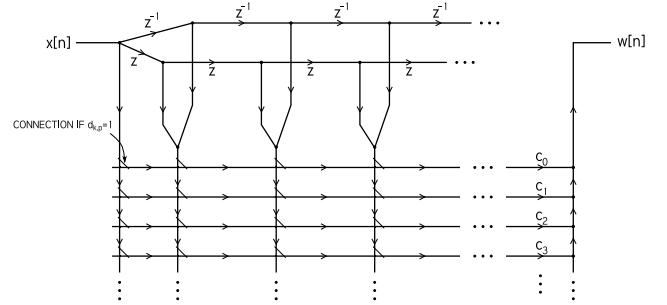

Fig. 1. Zero-phase, FIR structure requiring $|\mathcal{C}|$ multipliers. Interconnections are made for $d_{k,p}$ that have value 1.

Cascading the structure in Fig. 1 with a compressor-by-M as indicated in Fig. 2 forms an integer downsampling system where all multipliers operate at the input sampling rate. We next apply a polyphase rearrangement of Fig. 2 in which all multipliers instead operate at the output rate. The structure is developed by first separating the filter into a cascade of a single-input, multiple-output delay system and a multiple-input, single-output multiplier system. Applying the downsampling noble identity results in the polyphase implementation.
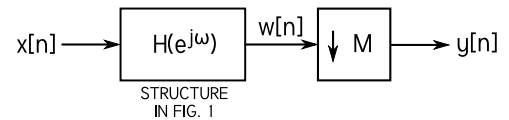

Fig. 2. Downsampling system based on Fig. 1.

The system in Fig. 1 has an input-output relationship defined by

$$w[n] = \sum_{p=0}^{K} \frac{x[n-p] + x[n+p]}{1 + \delta[p]} \sum_{k=0}^{\mathcal{U}-1} c_k d_{k,p}, \quad (4)$$

where $d_{k,p} \in \{0, 1\}$. Cascading a compressor-by-M with the system such that $y[n] = w[Mn]$ results in

$$y[n] = \sum_{p=0}^{K} \frac{x[Mn-p] + x[Mn+p]}{1 + \delta[p]} \sum_{k=0}^{\mathcal{U}-1} c_k d_{k,p}.$$

By separating $x[n]$ into the polyphase components

$$x_p[n] = x[Mn - p], \qquad (5)$$

we have

$$y[n] = \sum_{p=0}^{K} \frac{x_p[n] + x_{-p}[n]}{1 + \delta[p]} \sum_{k=0}^{\mathcal{U}-1} c_k d_{k,p}. \qquad (6)$$

The relationship between $x[n]$ and $y[n]$ is therefore represented as a cascade of a single-input, multiple-output system (Eq. 5) with a multiple-input, single-output system (Eq. 6). Let us express Eq. 5 as a cascade of a delay system and a compressor system:

$$v_p[n] = x[n - p] \qquad (7)$$
$$x_p[n] = v_p[Mn] \qquad (8)$$

Applying the downsampling noble identity to Eqns. 7 and 8 results in the following equations which relate $x[n]$ and $x_p[n]$:

$$u_{((p))_M}[n] = x[n - ((p))_M] \qquad (9)$$
$$r_{((p))_M}[n] = u_{((p))_M}[Mn] \qquad (10)$$
$$x_p[n] = r_{((p))_M}\left[n - \lfloor p/M \rfloor\right] \qquad (11)$$

Finally, changing the order of summation in Eq. 6 gives us

$$y[n] = \sum_{k=0}^{\mathcal{U}-1} c_k \sum_{p=0}^{K} \frac{x_p[n] + x_{-p}[n]}{1 + \delta[p]} d_{k,p}. \qquad (12)$$

The cascade of Eqns. 9-12, illustrated in Fig. 3, constitutes our downsampling system. The resemblance of Eq. 6 to Eq. 4 is reflected in the similarity between the interconnection structures in Figs. 1 and 3.

IV. RESULTS

To demonstrate the performance of the proposed structures and filter design techniques, we consider the following set of specifications for a $4 : 1$ downsampling lowpass filter:

- Passband edge: 0.2π
- Stopband edge: 0.25π
- Error in passband: $|E(e^{j\omega})| \le 0.01$
- Error in stopband: $|E(e^{j\omega})| \le 0.1$

These specifications were chosen to compare designs using filters of relatively low order.

The minimum-length Parks-McClellan filter that meets the design constraints is illustrated in Fig. 4. The filter in Fig. 5 was designed using the technique in Subsection II-A with a choice of N_x for which the specifications were met. This filter has a smaller zero-norm than the Parks-McClellan design.

The filter in Fig. 6 was designed using the technique in Subsection II-B, which does not require specification of N_x. The filter has the same zero-norm as the design where N_x is set (Fig. 5). While this filter does not strictly meet the design specifications, it is an acceptable design with respect to a slightly-relaxed set of tolerances.

In Table I we compare the computational cost of each of these filters in our downsampler-by-4 configuration, implemented using a folded structure, a canonical polyphase structure, and the polyphase structure presented in Section III. For all structures, we assume that multiplications by 0 are not implemented. Computational cost is stated in terms of the number of required multiplications per output sample of the system.

REFERENCES

[1] M. Bellanger. *Digital Processing of Signals: Theory and Practice*. John Wiley & Sons, Chichester, 1984.
[2] T. A. Classen and W. F. Mecklenbräuker. On the transposition of linear time-varying discrete-time networks and its application to multirate digital systems. *Philips J. Res.*, 23:78–102, 1978.
[3] R. E. Crochiere and L. R. Rabiner. *Multirate Digital Signal Processing*. Prentice-Hall, Englewood Cliffs, NJ, 1983.
[4] A. V. Oppenheim, R. W. Schafer, and J. R. Buck. *Discrete-Time Signal Processing*. Prentice-Hall, Inc., Englewood Cliffs, NJ, 1999.
[5] T. W. Parks and J. H. McClellan. Chebyshev approximation for nonrecursive digital filters with linear phase. *IEEE Transactions on Circuit Theory*, pages 189–194, Mar. 1972.
[6] K. Steiglitz, T. W. Parks, and J. F. Kaiser. METEOR: a constraint-based FIR filter design program. *IEEE Transactions on Signal Processing*, 40(8):1901–1909, Aug. 1992.
[7] J. A. Tropp. Just relax: Convex programming methods for identifying sparse signals. *IEEE Transactions on Information Theory*, 51(3):1030–1051, Mar. 2006.
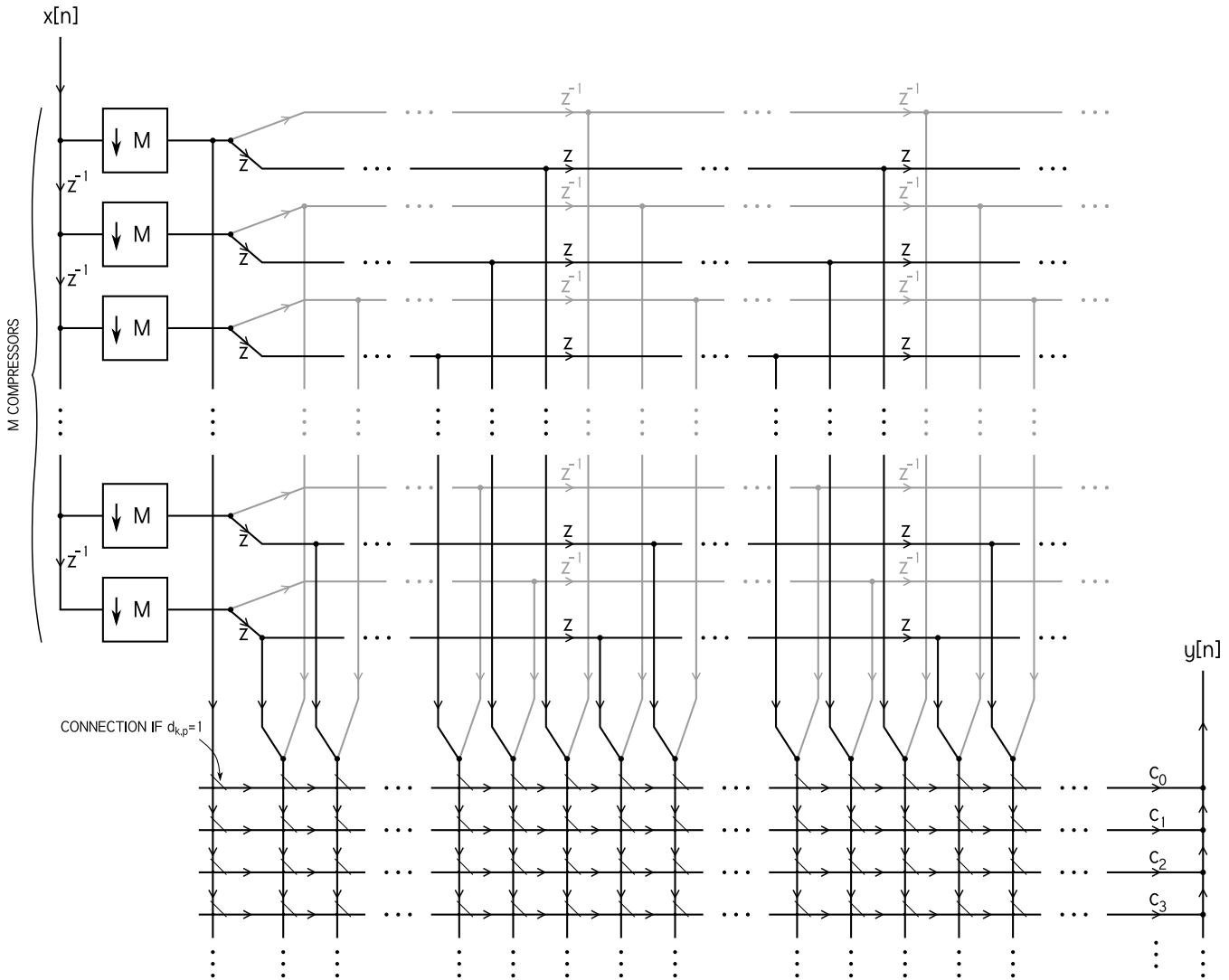[8] P. P. Vaidyanathan. *Multirate Systems and Filter Banks*. Prentice-Hall, Inc., Englewood Cliffs, NJ, 1993.

Fig. 3. Polyphase implementation of downsampling system in Fig. 2. M compressors and $|\mathcal{C}|$ multipliers are required. The folded delay lines can be implemented in two layers: one for delay elements (gray) and another implementing advances (black).

Design technique	Illustration of designed filter	Cost, folded structure	Cost, polyphase structure	Cost, presented structure
Parks-McClellan	Figure 4	108	53	27
Constrained coefficients (Subsection II-A)	Figure 5	84	41	21
Minimize $\|h[n]\|_1$, zero small coefficients (Subsection II-B)	Figure 6	84	41	21

TABLE I
COMPUTATIONAL COST OF DOWNSAMPLER-BY-4, IN TERMS OF THE NUMBER OF REQUIRED MULTIPLICATIONS PER OUTPUT SAMPLE.
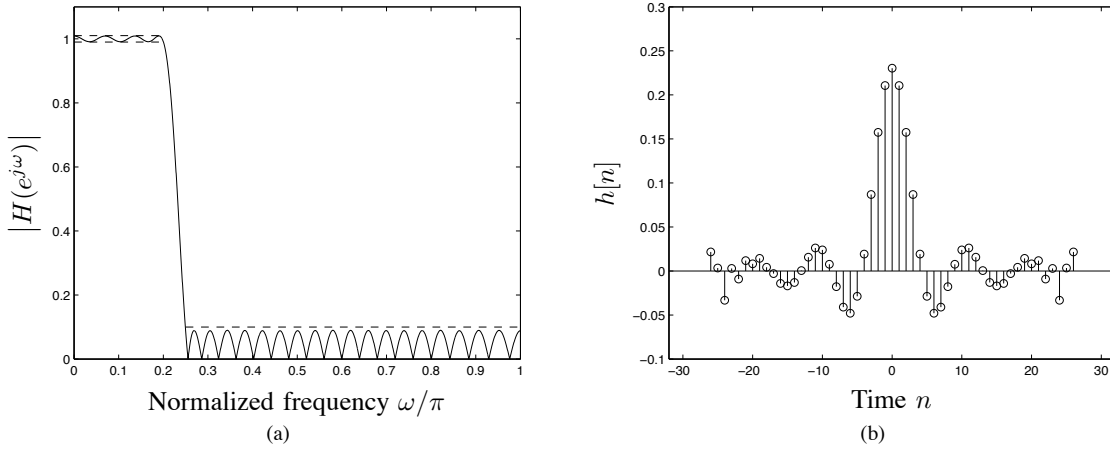
Fig. 4.   Magnitude response (a) and impulse response (b) of minimum-length Parks-McClellan filter that meets the design constraints. Dotted lines in (a) represent design specifications. $K = 26$, $||h[n]||_0 = 53$.
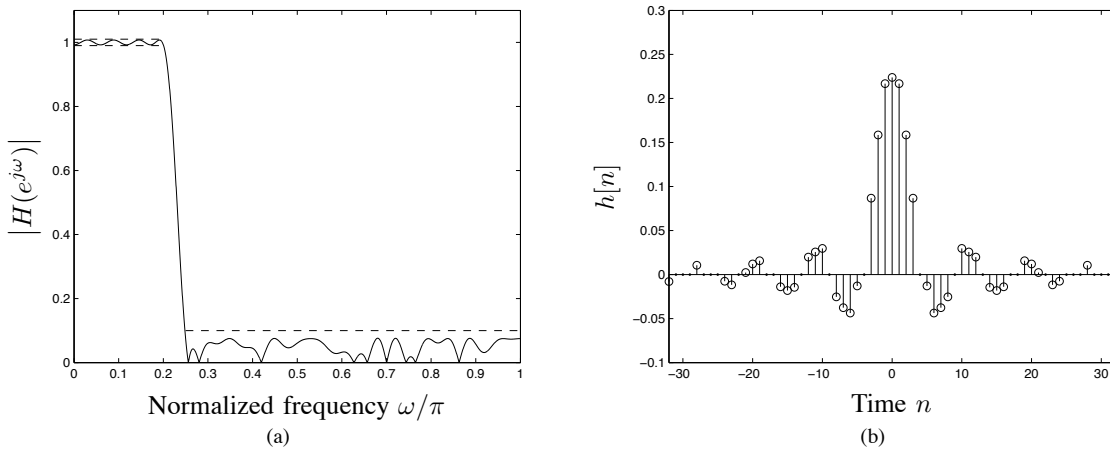

Fig. 5.   Magnitude response (a) and impulse response (b) of filter where some coefficients are constrained to zero (Subsection II-A). This choice of zeroed coefficients results in a filter that meets the design constraints. Dotted lines in (a) represent design specifications. Solid dots in (b) represent values of $h[n]$ that were constrained to have value zero. $K = 32$, $||h[n]||_0 = 41$.
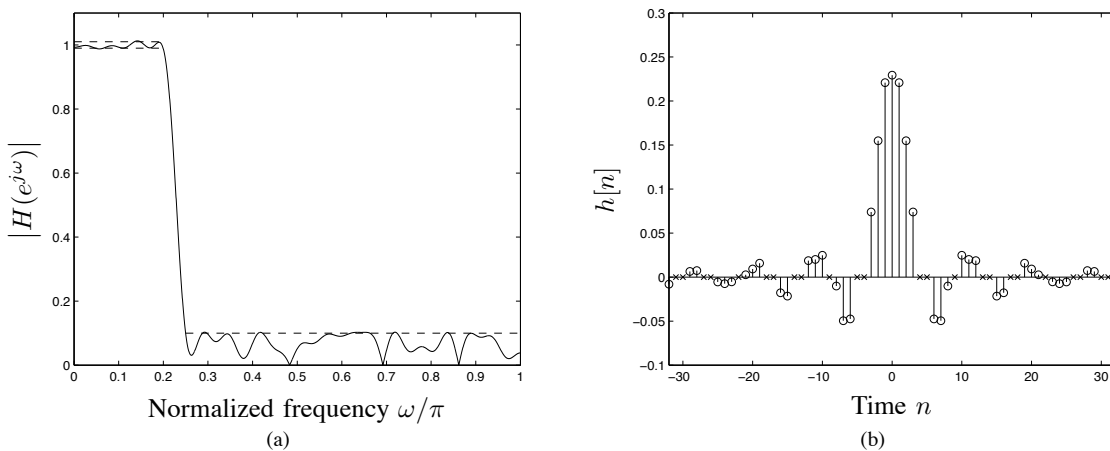

Fig. 6.   Magnitude response (a) and impulse response (b) of filter where $||h[n]||_1$ is minimized, subject to the design constraints, and where coefficients c_k with $|c_k| < 0.002$ are zeroed (Subsection II-B). Dotted lines in (a) represent design specifications. Symbols $\times$ in (b) represent values of $h[n]$ that were zeroed. $K = 32$, $||h[n]||_0 = 41$.