DESIGN OF FILTERS FOR DISCRETE SHORT-TIME FOURIER TRANSFORM SYNTHESIS *

Z. Shpiro

Calltalk Ltd.
108 Giborey Israel St.
Tel-Aviv 67891, Israel

D. Malah

Dept. of Electrical Engineering
Technion - Israel Inst. of Tech.
Haifa 32000, Israel.

ABSTRACT

In this paper we present a method for computing the synthesis filter (window) needed in a weighted overlap-add (WOLA) scheme for the reconstruction of signals in analysis/synthesis systems used to implement the discrete short time Fourier transform (DSTFT). The method is based on an algebraic representation of the analysis synthesis process and assume that the analysis filter (window) is known, that its length N is larger than or equal to the transform size (i.e. the number of frequency bands) M, that no modification of the DSTFT is performed, and that exact signal reconstruction (unity system) is to be achieved. The last condition can be achieved only if the shift R of the sliding analysis window satisfies $1 \leq R < M$. The solution presented in this paper extends an earlier result obtained for N = M, and is of practical importance, since using N > M results in analysis filter banks with improved frequency band separation. The algebraic method presented is simple and efficient as it reduces the large dimensionality of the problem into a solution of R sets of linear equations of reduced dimensions. The solutions of these equations are the individual synthesis polyphase filters from which the synthesis filter is constructed.

I. INTRODUCTION

There has been growing interest in recent years in the application of the discrete short-time Fourier transform (DSTFT) for the analysis and synthesis of nonstationary signals such as speech [1-7]. In applications where the DSTFT is not modified, the main design goal is to achieve a unity analysis/synthesis system. Typically, the analysis filter (or window), and the number of frequency bands (i.e. the transform size - M) are selected to meet desired specifications on the analysis filter bank. Thus, the remaining issues are the selection of the decimation factor R (i.e. the sliding window shift) and the design of the synthesis filter (window). Although a necessary and sufficient condition for obtaining exact reconstruction (i.e. a unity system) by an analysis/synthesis window pair is known [1], it did not lead to an explicit design method. However, progress has been made recently with respect to this problem and the following results have been reported.

(i) If the length of the analysis window (N) is equal to the transform size (i.e. N = M), a closed form expression for the synthesis window (given the analysis window) was found [5,7], and is applicable for all values of R in the range $1 \leq R \leq M$, with a mild constraint on the analysis window. The constraint is that no analysis polyphase filter impulse response is identically zero [7]. This constraint is milder than the "no zero value" constraint assumed in [5]. Further more, this solution is also optimal in the MMSE sense, if the signal is reconstructed from a modified DSTFT sequence (i.e., the mean square absolute error between the given modified DSTFT sequence and the DSTFT of the reconstructed signal is minimized) [5,7].

(ii) If N > M the design method presented in [6] appears to be applicable to this problem as well. However, not only does this method require the solution of an eigenvalue-eigenvector problem of large dimensions (in general), it is also a singular problem. The singularity stems from the fact that we deal here with a unity system for which the minimum error is zero, and therefore infinitely many solutions exist.

Because the use of N > M is of great practical importance, as it allows the design of improved analysis filter banks by reducing the transition bandwidth of the individual filters, we aim in this paper to solve for a synthesis filter which results in a unity analysis/synthesis system, in a simpler way than in [6], by using the algebraic approach presented in [7]. Our solution is obtained for values of R satisfying $1 \leq R < M$ and is shown to coincide with the solution obtained in [7] if N = M. It is also shown, by algebraic means, that if N > M and R = M, no finite length (FIR) synthesis filter can satisfy the unity system condition. In this case the design approach in [6] can provide an approximate solution.

The paper is organized as follows: In the next section we briefly review DSTFT analysis and synthesis and its algebraic representation. In Section III we present the algebraic solution to the synthesis filter design problem and in Section IV we present a design example and summarize the paper.

* This research was supported by Technion V.P.R. Fund - Philipson Fund for Electrical Eng. Research.

CH2118-8/85/0000-0537 $1.00 © 1985 IEEE

II. THE DSTFT AND ITS ALGEBRAIC REPRESENTATION

The DSTFT of a signal x(n) is defined by [1]

$$X(sR,k\Delta\Omega) \triangleq \sum_{m=-\infty}^{\infty} x(m)h(sR-m)\exp(-jk\Delta\Omega m) \quad , \quad k = 0,1,\ldots,M-1 \tag{1}$$

where h(n) is the analysis window sequence (prototype analysis filter), of length N, $\Delta\Omega = 2\pi/M$ is the frequency resolution, and R is the sliding window shift (in samples) or equivalently the decimation ratio in each channel of the analysis filter bank (if a filter bank interpretation of the DSTFT is used [1,2,4]). The synthesis, or reconstruction, is performed by using the synthesis filter f(n) to interpolate the decimated DSTFT sequence and inverse transformation (1) i.e.

$$\hat{x}(n) = \frac{1}{M}\sum_{k=o}^{M-1}\sum_{s=-\infty}^{\infty} f(n-sR)X(sR,k\Delta\Omega)\exp(jk\Delta\Omega n) \quad . \tag{2}$$

The realization of the above analysis and synthesis operations, is simplified by formulating (1) and (2) in terms of DFT's of size M on properly defined consecutive blocks of data [1-3]. The analysis can be represented in matrix form [7], as follows. Let $\underline{x}$ denote the input signal vector and $\underline{b}$ the fixed time reference DSTFT vector consisting of concatenated transform blocks. Then, the relation between $\underline{b}$ and $\underline{x}$ can be written in matrix form as [7],

$$C\underline{x} = \underline{b} \tag{3}$$

where C is a nonsquare matrix with dimensions $m \times \ell$, with $m > \ell$ (if $N \geq M$ and $R < M$). From the above description of the analysis steps, C can be decomposed into three factor matrices:

$$C = FPA \tag{4}$$

where P is a block diagonal permutation matrix (each block is $M \times M$) which performs the necessary circular rotation [2] on each aliased data block of length M, F is a block diagonal matrix with identical $M \times M$ blocks W which perform the DFT, and A is a rectangular matrix performing the windowing and time aliasing of consecutive input segments - each of length N with shifts of R samples at a time [2].

The matrix representation of the WOLA block-by-block synthesis operations [2] is given by [7]

$$D\underline{b} = \hat{\underline{x}} \tag{5}$$

where D is a non-square matrix with dimensions $\ell \times m$, $m > \ell$ (if $N \geq M$ and $R < M$). According to the above synthesis procedure D can be decomposed into three factor matrices:

$$D = S\hat{P}\hat{F} \tag{6}$$

where $\hat{F}$ performs the IDFT in the synthesis procedure, $\hat{P}$ performs the reverse circular rotation and S is a matrix describing the weighting of the periodically repeated sequence and the overlap-add operation needed to obtain the output signal. Since $\hat{P}$ and $\hat{F}$ perform the inverse operations that are performed by P and F in (4), one obtains [7]

$$P\hat{P} = I, \quad F\hat{F} = I \quad . \tag{7}$$

Hence, if no modification of the DSTFT is performed, the input-output relation of the overall analysis/synthesis system is given by

$$DC\underline{x} = SA\underline{x} = \hat{\underline{x}} \quad . \tag{8}$$

Thus, a unity system is obtained if the rectangular matrices A and S satisfy

$$SA = I \quad . \tag{9}$$

Before we proceed the following should be noted:

(i) The dimensions of the matrices A and S increase with input signal duration.

(ii) The structures of A and S here are not the same as in [7], since if $N > M$, A includes the time aliasing operation and S includes the periodic repetition operation which are not needed if $N = M$, as assumed in [7].

In the following section we present a solution for S which corresponds to WOLA synthesis and hence allows a relatively simple determination of the synthesis filter f(n).

III. SYNTHESIS FILTER DESIGN

In principle, one could attempt to solve for S in (9) (given A) by using the generalized inverse of A as was done in [7] for $N = M$. However, because A has a more complicated structure than in [7], we found it difficult to solve for S using this approach, especially considering the fact that the dimensions of S grow with the increase in input signal duration.

The approach taken was to construct S from partial solutions used for the reconstruction of individual elements of $\hat{\underline{x}}$. We explain and illustrate the method via the design example described in Fig.1. The design parameters in this example are: Filter length of N = 8 samples, transform size of M = 4, and a decimation factor (window shift) of R = 2. As seen in this figure, the matrix A has the following special properties and structure:

(i) The first and last N-R columns are connected with the transient phenomena at the beginning and end due to the finite length input vector and are disregarded in the general solution.

(ii) The remaining columns appear in groups of R columns each, which repeat periodically but with a linearly progressing down shift (three such groups are noted in Fig. 1).

(iii) The non-zero elements in each column i of a group, i = 0,1,2,...,R-1 are the i-th polyphase of the analysis filter. This property is also illustrated in Fig. 1, where the analysis window samples are denoted by $\{h(0),h(1),\ldots,h(N-1)\}$, N = 8.

Now, we examine how a given element x_r in the input vector $\underline{x}$ participates in the windowing and time aliasing process described by

$$A\underline{x} = \underline{y} \quad . \tag{10}$$

We observe that the element x_r appears only in [N/R] equations which can be rewritten as a reduced system of [N/R] equations expressed by

$$A_r \underline{x}_r = \underline{y}_r \quad , \tag{11}$$

where A_r is the reduced system matrix, corresponding to element x_r, of dimensions $[N/R] \times [2(N/M)-1]$, and the vector $\underline{x}_r$ contains x_r as its center element.

Because of the periodic structure of A, noted in property (ii) above, the number of distinct matrices A_r is only R. That is, (11) can be rewritten as

$$A_\ell \underline{x}_r = \underline{y}_r, \quad \ell = r \bmod R \quad . \tag{12}$$

For the example in Fig. 1, there are therefore $R=2$ distinct matrices, A_0 and A_1, which are shown in Fig. 2 and are demonstrating (12), for $r = 6$ (which gives $\ell = 0$) and $r = 7$ (which gives $\ell = 1$).

The problem we pose now is how to solve for the single center element x_r from the set of equations in (12). The solution we consider is to find the generalized inverse $A_\ell^\dagger$ of A_ℓ, i.e.,

$$A^\dagger = (A^T A)^{-1} A^T \tag{13}$$

and finding x_r by using

$$x_r = \underline{a}_\ell^T \underline{y}_r \tag{14}$$

where $\underline{a}_\ell$ is the center row vector of $A_\ell^\dagger$, and, as explained earlier, $\ell = r \bmod R$.

A matrix S which yields a unity systems must satisfy:

$$S\underline{y} = \underline{x} \quad . \tag{15}$$

Therefore, we can construct such a matrix by properly arranging the partial solutions (for individual elements) of the form given in (14). This is done by first putting $\underline{a}_\ell^T$ in all the rows that have index i which satisfies $i \bmod R = \ell$. Now zeroes are inserted in the required location so that if the r-th row of S is multiplied by y the result is x_r. The structure of the matrix S which results from the above process, for the example in Fig. 1, is illustrated in Fig. 3. It can be verified that the resulting synthesis matrix S corresponds to the WOLA synthesis process described in Section II if the length of the synthesis filter f(n) equals to the length (N) of the analysis filter h(n).

Furthermore, since the i-th row of S is the ℓ-th ($\ell = i \bmod R$) synthesis polyphase filter (in reversed order of elements), the synthesis filter f(n) can be easily constructed. This result shows that actually one can use $\underline{a}_\ell^T$ in (14) as the ℓ-th synthesis polyphase filter (in reversed order of elements) without going through the construction of the matrix S.

It should be noted that if $N = M$ the above approach for finding the synthesis filter results in the solution obtained earlier in [7]. That this is the case can easily be seen by noting that with $N = M$, the reduced matrix A_r becomes a column vector whose elements are the analysis polyphase filter terms. Thus $(A_r^T A_r)^{-1}$ is a scalar and the synthesis polyphase filter becomes the analysis polyphase weighted by this scalar, which is the result in [7].

If $N > M$ and $R = M$, it is known [1] that no unity system exists. This is pronounced in the algebraic approach presented above by the singularity of the problem, as the number of columns of A is now larger than the number of rows.

IV. DESIGN EXAMPLE AND SUMMARY

An example of the results obtained using the design method described in the previous section is shown in Fig. 4. In this example, the analysis window (Fig. 4a) is a sin(x)/x window of length $N = 1024$, the transform size is $M = 128$, and the decimation factor is $R = 32$. The resulting synthesis window is shown in Fig. 4b. This pair of analysis/synthesis filters was used in an analysis/synthesis system and indeed was found to yield a unity system.

In summary, a simple and efficient method for the construction of a synthesis filter which assures a unity analysis/synthesis system (if the DSTFT is not modified) for a given analysis filter having length N which is larger than the transform size M was presented. The solution is applicable for $1 \leq R < M$ since for $R = M$ no finite length synthesis filter can satisfy the unity system condition if $N > M$. The solution involves the inversion (generalized inverse) of R matrices of reduced dimensions $([N/R] \times [2(N/M)-1])$, from the center row of which the synthesis polyphase filters are formed and used to construct the synthesis filter (window).

We conclude by noting that if the DSTFT is modified, the above synthesis filter is not necessarily optimal (in the MMSE sense). However, if $N = M$ the above result coincides with the result in [7] which is optimal. The problem of solving for a synthesis filter, which satisfies the unity system condition for $N > M$, and is optimal if the DSTFT is modified, is now under study.

REFERENCES

[1] M.R. Portnoff, "Time Frequency Representation of Digital Signals and Systems Based on Short Time Fourier Analysis", IEEE Trans. on ASSP, Vol. ASSP-28, No. 1, pp. 55-69, February 1980.

[2] R.E. Crochiere, "A Weighted Overlap-Add Method of Short Time Fourier Analysis/Synthesis", IEEE Trans. on ASSP, Vol. ASSP-28, No. 1, pp. 99-102, February 1980.

[3] D. Malah and J.L. Flanagan, "Frequency Scaling of Speech Signals by Transform Techniques", Bell System Technical Journal, Vol. 60, No. 9, pp. 2107-2156, November 1981.

[4] R.E. Crochiere and L.R. Rabiner, *Multirate Digital Signal Processing*, Prentice Hall, NJ, 1983.

[5] D.W. Griffin and J.S. Lim, "Signal Estimation from Modified Short Time Fourier Transform", IEEE Trans. on ASSP, Vol. ASSP-32, No. 2, pp. 236-243, April 1984.

[6] V.K. Jain and R.E. Crochiere, "A Novel Approach to the Design of Analysis/Synthesis Filter Banks", IEEE ICASSP-83, Boston, pp. 228-231.

[7] Z. Shpiro and D. Malah, "An Algebraic Approach to Discrete Short-Time Fourier Transform Analysis and Synthesis", IEEE ICASSP-84, pp. 2.3.1 - 2.3.4, pp. 804-807.

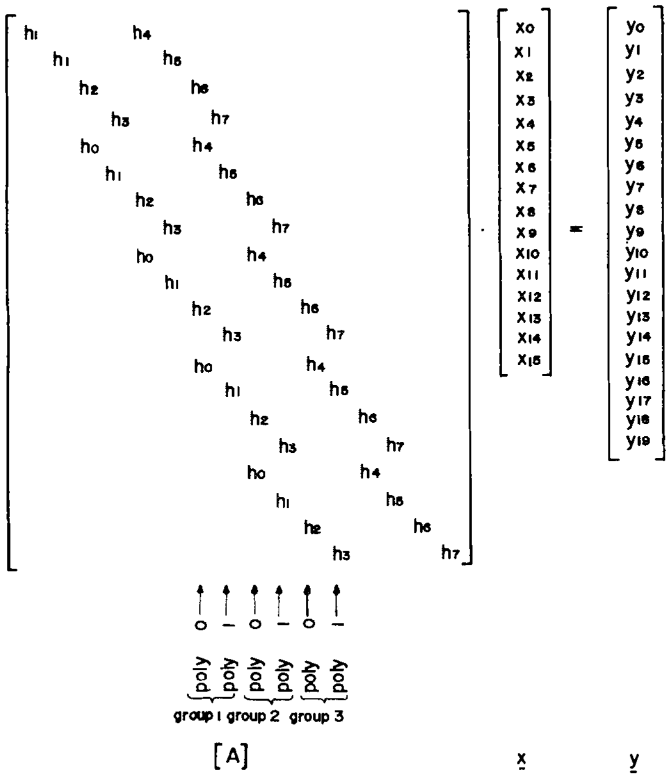

Fig. 1: DSTFT analysis - windowing and aliasing.

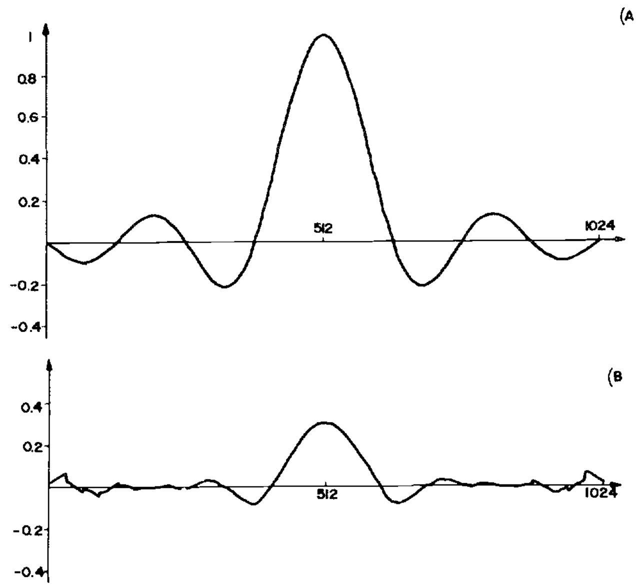

Fig. 4: Design example - (a) analysis window; (b) synthesis window (M = 128, R = 32).

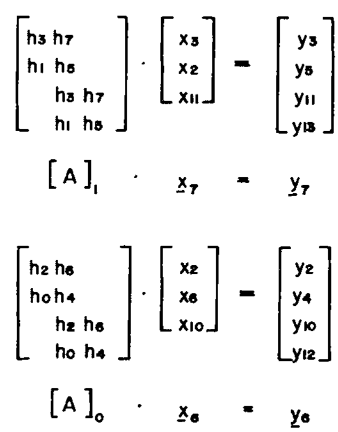

Fig. 2: Reduced analysis.

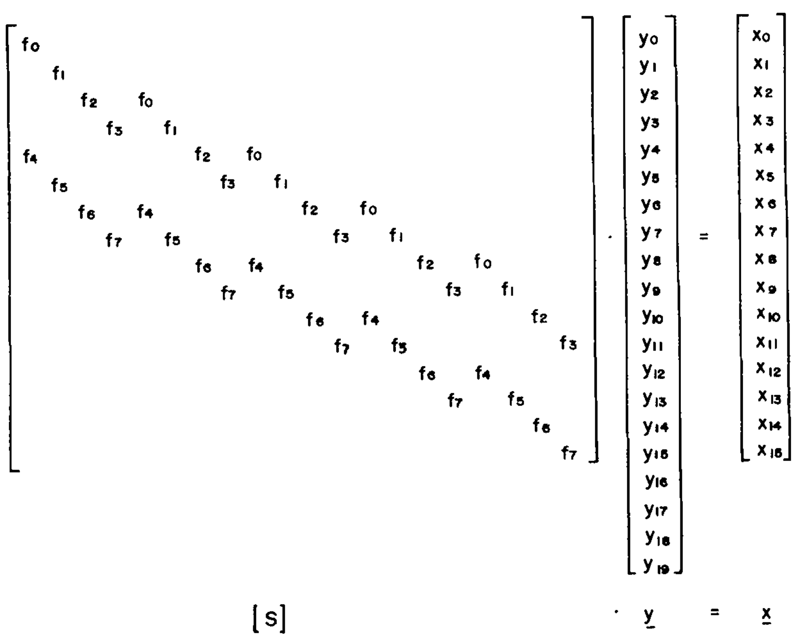

Fig. 3: DSTFT synthesis - periodic repetition and WOLA.