THE GENERALIZED LAPPED BIOORTHOGONAL TRANSFORM

Trac D. Tran\textsuperscript{1}  \quad Ricardo de Queiroz\textsuperscript{2}  \quad Truong Q. Nguyen\textsuperscript{3}

\textsuperscript{1}University of Wisconsin, ECE Department, Madison, WI, 53706
\textsuperscript{2}Xerox Corporation, Corporate Research and Technology, Webster, KY, 14580
\textsuperscript{3}Boston University, ECE Department, Boston, MA 02215
trac@saigon.ecc.wisc.edu, queiroz@wrc.xerox.com, nguyen@bu.edu

ABSTRACT

A lattice structure based on the singular value decomposition (SVD) is introduced. The lattice can also be proven to use a minimal number of delay elements and to completely span a large class of $M$-channel linear phase perfect reconstruction filter bank (LPPRFB): all analysis and synthesis filters have the same FIR length of $L = KM$, sharing the same center of symmetry. The lattice also structurally enforces both linear phase and perfect reconstruction properties. It is capable of providing fast and efficient implementation, and avoids the costly matrix inversion problem in the optimization process. From a block transform perspective, the new lattice represents a family of generalized lapped biorthogonal transform (GLBT) with arbitrary integer overlapping factor $K$. The relaxation of the orthogonal constraint allows the GLBT to have significantly different analysis and synthesis basis functions which can then be tailored appropriately to fit a particular application. Several design examples are presented along with a high-performance GLBT-based progressive image coder to demonstrate the superiority of the new lapped transforms.

1. INTRODUCTION

Linear phase perfect reconstruction filter banks have been used extensively in numerous applications, especially image processing [1]. In the two-channel case, all solutions have been found whereas there are still many open problems in $M$-channel cases. An attractive approach to the design and implementation of LPPRFB is the parameterization by lattice structures based on the factorization of the polyphase matrices $E(z)$ and $R(z)$ shown in Figure 1. The lattice structure offers fast and efficient implementation, retains both LP and PR properties regardless of coefficient quantization, and (if it is general enough) guarantees that no optimal solution will be excluded in the optimization process. Complete and minimal two-channel PR lattice structure has been reported in [2]. $M$-channel lattices have been presented for the more restricted paraunitary case [3], resulting in the generalized lapped orthogonal transform (GenLOT) [4]. No general lattice has been reported for the biorthogonal case (defined as $R(z)E(z) = z^{-K}I$). Only several particular solutions were proposed so far: Chan replaced some orthogonal matrices in [4] by cascades of invertible block diagonal matrices [5]; Malvar suggested a simple scaling of the first antisymmetric basis function of the initial block (which was chosen to be the DCT) [6].

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure1.png}
\caption{Polyphase representation of LPPRFB.}
\end{figure}

In this paper, we first introduce a general structure that propagates the linear phase property. Further imposition of the biorthogonal property on this structure results in a complete and minimal LPPRFB lattice whose invertible coefficient matrices are then parameterized by the SVD. In the lapped transform language, the structure is interpreted as a robust characterization of the GLBT. Several lapped transforms obtained from the nonlinear optimization of the lattice coefficients are presented. Finally, the superiority and the potential of the new LT family are illustrated through an image coding example. The GLBT-based embedded coder consistently outperforms the wavelet-based version SPIHT [7] by a large margin. The improvement in PSNR can be up to an astounding 2.65 dB.

Notation-wise, vectors and matrices are denoted by boldfaced characters. Special matrices have reserved symbols: I, J, 0, D represents, respectively, the identity matrix, the reversal matrix, the null matrix, and the diagonal matrix whose entry is $+1$ when the corresponding filter is symmetric and $-1$ when the corresponding filter is antisymmetric. Capital letters $M, L, K$ denote respectively the number of channels, the filter length, and the overlapping factor.

2. LATTICE STRUCTURE

2.1. General LP-propagating Structure

Consider an $M$-channel FIR LPPRFB with all analysis and synthesis filters of length $L = KM$, having the same center of symmetry ($M \times L$ GLBT). The associated polyphase matrix $E(z)$ has to satisfy the LP property [3]

$$E(z) = z^{-(K-1)} D E(z^{-1}) J. \quad (1)$$

Define the order-$(K-1+N)$ polyphase matrix $F(z) \triangleq G(z)E(z)$ where the all-zero order-$N$ $G(z)$ is the propagating structure. We are adding on or peeling off a block in the
The starting block $E_0$ has no delay element, representing LPPRFB of length $M$, and was often chosen to be DCT [4, 5, 6]. General $E_0$ satisfying Eq.(1) can be factorized as

$$E_0 = \frac{1}{\sqrt{2}} \begin{bmatrix} U_0 & U_0 \phi \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} U_0 & 0 \end{bmatrix} \begin{bmatrix} I & J \end{bmatrix}. $$

For $E_0$ to have PR, $U_0$ and $V_0$ again have to be invertible. This result should not come as a surprise. The factorization is very similar to the GenLOT's [4]: the only difference here is that $U_0$ and $V_0$ do not have to be orthogonal. Now, the difficulty arises: how do we completely characterize these $\frac{M}{2} \times \frac{M}{2}$ nonsingular matrices?

Recall that every invertible matrix has an SVD: $U_1 = U_0 \Gamma_1 U_0$, where $U_0$ and $U_1$ are orthogonal matrices, and $\Gamma_1$ is a diagonal matrix with positive elements [1]. Thus, $\Phi_1$ can be further factorized as

$$\Phi_1 = \begin{bmatrix} U_1 & 0 \\ 0 & V_1 \end{bmatrix} \begin{bmatrix} \Gamma_1 & 0 \\ 0 & \Delta_1 \end{bmatrix} \begin{bmatrix} U_1 & 0 \\ 0 & V_1 \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & V_1 \end{bmatrix}. $$

The orthogonal matrices $U_0$, $U_1$, $V_0$, and $V_1$ are parameterized by $\frac{M(M-2)}{2}$ rotations each. The diagonal matrices $\Gamma_1$ and $\Delta_1$ are characterized by $\frac{M}{2}$ positive parameters each. The complete lattice structure is shown in Figure 2 (drawn for $M = 8$). The most general $M \times L$ GLBT can be parameterized by $\frac{ML}{2}$ parameters as expected from LP systems. However, the SVD parameterization in Eq.(5) has three advantages: (i) exact reconstruction is guaranteed structurally under a mild condition – as long as none of the diagonal coefficients is quantized to zero; (ii) SVD representation avoids the costly matrix inversion problem in the optimization process; (iii) it is much simpler to prevent the matrices from being singular or near-singular.

It is also very easy to verify that all previously reported LPPRFB’s lattice structures are special cases of the new GLBT lattice. For examples, the GLT design example in [5] has $M = 8$, $K = 2$, $U_0$ and $V_0$ from the DCT. $U_0 = V_0 = \Gamma_0 = \Delta_0 = \Delta_1 = I$, and $V_1$ parameterized as a cascade of block diagonal matrices. The LBT in [6] has $M = 8$, $K = 2$, $U_0$ and $V_0$ from the DCT. $U_0 = V_0 = \Gamma_0 = \Delta_0 = \Delta_1 = I$, $\Delta_0 = diag(\sqrt{2} 1 1 1)$, and $U_1$, $V_1$ to be orthogonal. When orthogonality is imposed, we get back GenLOT [4]. When $M = 2$, the lattice turns into a simplified form of Type-A system lattice in [2]. Formal proofs of the lattice’s completeness and minimalit y will be presented in [8].

3. DESIGN EXAMPLES AND GLBT APPLICATION IN IMAGE CODING

Figure 3 presents several GLBT design examples obtained from nonlinear optimization of the new lattice coefficients with various cost functions. The analysis banks are on top. Designs in Figure 3(a), (b), and (d) are DCT-based. While increasing the GLBT length does not improve the coding gain much, it helps in the case of stopband attenuation (where longer filters are desired) as testified in Figure 3(c).

One of the most popular applications of the GLBT is image compression. Overlapping analysis filters reduce inter-block redundancy, providing higher coding efficiency of the transform coefficients, while overlapping synthesis filters whose ends decay asymptotically to zero eliminate blocking artifacts. Each bank can now be designed appropriately.
Figure 2. Lattice structure for the GLBT.

Figure 3. GLBT design examples (a) \( M = 8 \) \( L = 16 \) optimized for coding gain, DC attenuation and mirror frequency attenuation (b) \( M = 8 \) \( L = 32 \) optimized for coding gain and DC attenuation (c) \( M = 8 \) \( L = 32 \) optimized for analysis stopband attenuation (d) \( M = 16 \) \( L = 32 \) optimized for coding gain, DC and mirror frequency attenuation.

For the case of \( M = 8 \) \( L = 16 \), GLBT optimized for pure coding gain can attain 9.63 dB. However, we trade off 0.01 dB of coding gain in Figure 3(a) for high attenuation at DC, near-DC, and mirror frequencies to ensure high visual quality in the reconstructed images as well. The \( 16 \times 32 \) GLBT in Figure 3(d) achieves an impressive coding gain of 9.96 dB. When the new transforms are incorporated into the block-transform progressive coding framework described in [9], the resulting GLBT-based embedded coders provide unrivaled objective and subjective performance as indicated in Table 1 and Figure 4(b)-(d). For a smooth image like Lena which the wavelet transform can sufficiently decorrelates, the best wavelet-based embedded coder SPIHT [7] provides a comparable performance. However, for a highly-textured image like Barbara, \( 16 \times 32 \) GLBT coder can provide a PSNR gain of around 2.5 dB over a wide range of bit rates. The visual reconstructed image quality is also superior: texture is beautifully preserved, blocking is completely eliminated, and ringing is barely noticeable. Comparing to the optimal \( 8 \times 40 \) GenLOT in [9], the \( 8 \times 16 \) GLBT in Figure 3(a) already offers comparable performance at much lower computational complexity. More objective and subjective evaluation of GLBT-based progressive coders can be found at http://saigon.ece.wisc.edu/~waveweb/Coder/index.html along with more details on the coding scheme.

4. CONCLUSIONS

We have presented in this paper a general, albeit minimal, lattice structure for \( M \)-channel \( KM \)-length LPPFEB. The novel lattice based on the SVD provides a robust implementation and a friendly design procedure for all lapped transforms with arbitrary integer overlapping factor \( K \).

REFERENCES

Table 1. Objective coding result comparison

<table>
<thead>
<tr>
<th>Comp. Ratio</th>
<th>SPIHT</th>
<th>8 x 40 GenLOT</th>
<th>8 x 16 GLBT</th>
<th>16 x 32 GLBT</th>
<th>SPIHT</th>
<th>8 x 40 GenLOT</th>
<th>8 x 16 GLBT</th>
<th>16 x 32 GLBT</th>
<th>SPIHT</th>
<th>8 x 40 GenLOT</th>
<th>8 x 16 GLBT</th>
<th>16 x 32 GLBT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:8</td>
<td>40.41</td>
<td>40.43</td>
<td>40.35</td>
<td>40.43</td>
<td>36.55</td>
<td>36.80</td>
<td>36.69</td>
<td>36.78</td>
<td>36.41</td>
<td>38.08</td>
<td>37.84</td>
<td>38.43</td>
</tr>
<tr>
<td>1:16</td>
<td>37.21</td>
<td>37.32</td>
<td>37.28</td>
<td>37.33</td>
<td>33.13</td>
<td>33.36</td>
<td>33.31</td>
<td>33.42</td>
<td>31.40</td>
<td>33.47</td>
<td>33.02</td>
<td>33.94</td>
</tr>
<tr>
<td>1:32</td>
<td>34.11</td>
<td>34.23</td>
<td>34.14</td>
<td>34.27</td>
<td>30.56</td>
<td>30.79</td>
<td>30.70</td>
<td>30.84</td>
<td>27.58</td>
<td>29.53</td>
<td>29.04</td>
<td>30.18</td>
</tr>
<tr>
<td>1:100</td>
<td>29.35</td>
<td>29.31</td>
<td>29.14</td>
<td>29.38</td>
<td>27.38</td>
<td>27.40</td>
<td>27.33</td>
<td>27.62</td>
<td>23.76</td>
<td>24.95</td>
<td>24.55</td>
<td>25.39</td>
</tr>
</tbody>
</table>

Figure 4. Embedded coding results of Barbara at 1:32 compression ratio (a) original image (b) SPIHT, 27.58 dB (c) embedded 8 x 16 GLBT, 29.04 dB (d) embedded 16 x 32 GLBT, 30.18 dB.