OVER-SAMPLED AND UNDER-SAMPLED PRE-/POST-FILTERS FOR BLOCK DCT CODERS

Chengjie Tu[1], Trac D. Tran[2], and Jie Liang[3]

[1]Windows Digital Media Division, Microsoft Corporation
Redmond, WA 98052. Email: chentu@microsoft.com
[2]ECE Department, The Johns Hopkins University
Baltimore, MD 21218. Email: trac@jhu.edu
[3]School of Engineering Science, Simon Fraser University
Burnaby, BC, V5A 1S6, Canada. Email: jiel@sfu.ca

ABSTRACT

Pre-/post-filtering operators have been shown to improve the coding efficiency as well as to mitigate blocking artifacts in traditional DCT-based block coders. Pre-/post-filters which preserve the system sampling rate have been extensively investigated. This paper explores the two non-critically-sampled signal decomposition cases – under-sampling and over-sampling – via the pre-/post-processing perspective. We discuss various design issues, efficient structures, and present two application examples: under-sampled pre-/post-filters for very low bit-rate image coding and over-sampled pre-/post-filters for error resilient image transmission. Preliminary experimental results illustrate that non-critically-sampled pre-/post-filtering does provide much improved coding performances than its critically-sampled counterpart for the aforementioned specific applications.

1. INTRODUCTION

Block coding based on the Discrete Cosine Transform (DCT) is very popular in image and video compression. Its success stems from the DCT's excellent energy compaction capability, low computational complexity, and a high degree of flexibility. The main disadvantage of block coders is that correlation between neighboring blocks has not been taken into account, leading to sub-optimal coding efficiency and the manifestation of blocking artifacts at low bit rates.

The key to improving the coding efficiency of block-DCT coders is to employ transforms with overlapped basis functions. The wavelet transform (WT) and the lapped transform (LT) are two signal decomposition schemes with such property [1, 2]. Recently, we have shown that various lapped transforms can be constructed by adding pre- and post-processing operators along the block boundaries of the traditional DCT coding framework [3]. This time-domain processing approach operates outside the existing

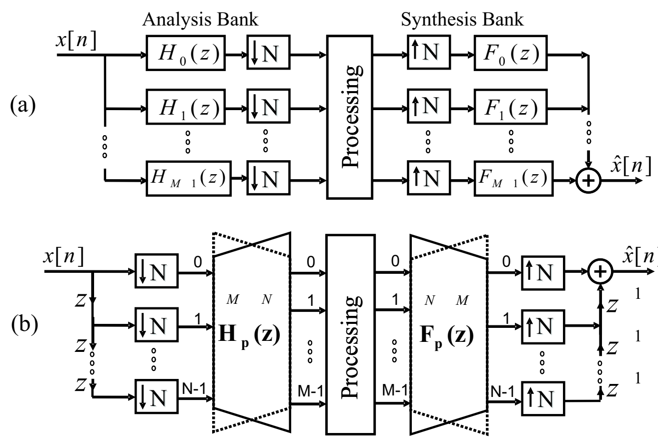

Fig. 1. General signal decomposition via an M-channel filter bank. (a) Traditional filtering representation. (b) Equivalent polyphase representation.

DCT-based block coding infrastructure. Hence, compliance to standards can be easily achieved. Furthermore, intuitive time-domain interpretation facilitates the transform design as well as implementation significantly. It has also been demonstrated that by turning on pre-/post-filtering and optimize adaptive context-based entropy coding appropriately, a DCT-based block coder can achieve competitive coding performance as the wavelet-based JPEG2000 coder [4].

In this paper, we investigate non-critically-sampled pre-/post-filters. Section 2 develops efficient methods to design pre-/post-filters with several desirable properties such as high coding gain, perfect reconstruction, or minimal reconstruction error under the linear phase constraint. Section 2 provides design examples and coding results to demonstrate the power of under- and over-sampled pre-/post-filters in two applications: low bit-rate image coding and error resilient image communication.

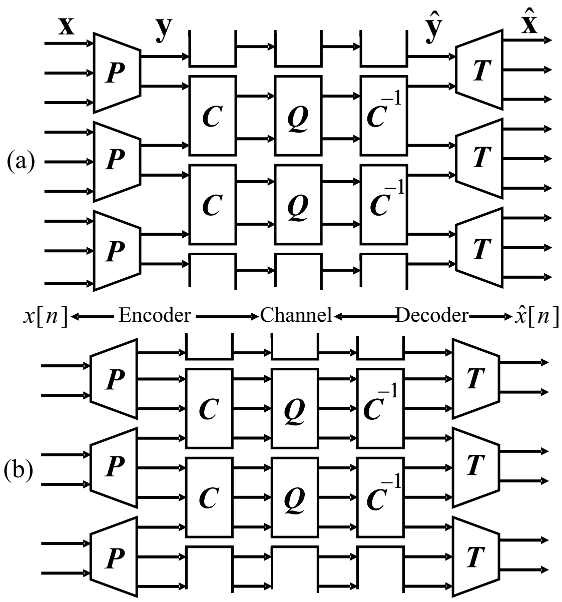

Fig. 2. Under-sampled (a) and over-sampled (b) pre-/post-filtering framework.

Notation-wise, let $\mathbf{I}_M$, $\mathbf{J}_M$, $\mathbf{0}_{M\times N}$ denote the $M \times M$ identity, the $M \times M$ reversal, and the $M \times N$ null matrix, respectively. Let $\mathbf{C}_M$ and $\mathbf{C}_M^{-1}$ be the M-point type-II DCT and IDCT matrices. Another common matrix is the $2m \times 2m$ butterfly matrix, $\mathbf{W}_{2m} \triangleq \frac{\sqrt{2}}{2}\begin{bmatrix} \mathbf{I}_m & \mathbf{J}_m \\ \mathbf{J}_m & -\mathbf{I}_m \end{bmatrix}$. $\mathbf{A}^+$ denotes the pseudo inverse of matrix $\mathbf{A}$.

2. DESIGN ISSUES

A general signal decomposition via an M-channel filter bank is depicted in Fig. 1. The system is critically-sampled when $M = N$, under-sampled if $M < N$, and over-sampled if $M > N$. A large subset of first-order polyphase matrices $\mathbf{H}_p(z)$ and $\mathbf{F}_p(z)$ can be realized from the framework of pre- and post-filtering as illustrated Fig. 2. We intentionally let the pre-/post-filters handle the rate-change responsibility. Assuming that M and N are even, the $M \times N$ pre-filter $\mathbf{P}$ and the $N \times M$ post-filter $\mathbf{T}$ can be partitioned into four quadrants as $\mathbf{P} = \begin{bmatrix} \mathbf{P}_{00} & \mathbf{P}_{01} \\ \mathbf{P}_{10} & \mathbf{P}_{11} \end{bmatrix}$, $\mathbf{T} = \begin{bmatrix} \mathbf{T}_{00} & \mathbf{T}_{01} \\ \mathbf{T}_{10} & \mathbf{T}_{11} \end{bmatrix}$. Then they generates an LT with the following forward and inverse transform matrices

$$\begin{aligned} \mathbf{H} &= \mathbf{C}_M \begin{bmatrix} \mathbf{P}_{10} & \mathbf{P}_{11} & \mathbf{0}_{\frac{M}{2}\times\frac{N}{2}} & \mathbf{0}_{\frac{M}{2}\times\frac{N}{2}} \\ \mathbf{0}_{\frac{M}{2}\times\frac{N}{2}} & \mathbf{0}_{\frac{M}{2}\times\frac{N}{2}} & \mathbf{P}_{00} & \mathbf{P}_{01} \end{bmatrix}, \\ \mathbf{F} &= \begin{bmatrix} \mathbf{T}_{01} & \mathbf{0}_{\frac{N}{2}\times\frac{M}{2}} \\ \mathbf{T}_{11} & \mathbf{0}_{\frac{N}{2}\times\frac{M}{2}} \\ \mathbf{0}_{\frac{N}{2}\times\frac{M}{2}} & \mathbf{T}_{00} \\ \mathbf{0}_{\frac{N}{2}\times\frac{M}{2}} & \mathbf{T}_{10} \end{bmatrix} \mathbf{C}_M^{-1}. \end{aligned} \tag{1}$$

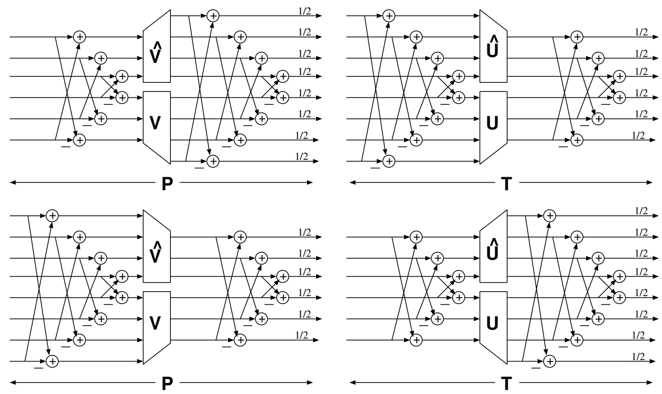

Fig. 3. Structure of pre- and post-filters: under-sampled case (top) and over-sampled case (bottom).

2.1. Coding Gain

The coding performance of a pre-/post-filter pair is measured by the coding gain of its corresponding LT:

$$G_{TC} \triangleq 2^{-R\left(1-\frac{N}{M}\right)} \times 10 \times log_{10} \frac{\sum_{i=0}^{M-1} \sigma_i^2}{\prod_{i=0}^{M-1} \sigma_i^2 ||\mathbf{F}_{.i}||_2^2}. \tag{2}$$

Here, R is the bit-rate; σ_i^2 is the variance of the i^{th} subband, $\mathcal{R}_{xx}$ is the autocorrelation matrix of the input signal; and $\mathbf{F}_{.i}$ is the i^{th} column of $\mathbf{F}$, or the i^{th} synthesis filter.

2.2. Linear Phase Property

Follow similar derivations as in [3], it can be shown that the linear phase property implies

$$\begin{aligned} \mathbf{P} &= \mathbf{W}_M \begin{bmatrix} \hat{\mathbf{V}} & \mathbf{0}_{\frac{M}{2}\times\frac{N}{2}} \\ \mathbf{0}_{\frac{M}{2}\times\frac{N}{2}} & \mathbf{V} \end{bmatrix} \mathbf{W}_N, \\ \mathbf{T} &= \mathbf{W}_N \begin{bmatrix} \hat{\mathbf{U}} & \mathbf{0}_{\frac{N}{2}\times\frac{M}{2}} \\ \mathbf{0}_{\frac{N}{2}\times\frac{M}{2}} & \mathbf{U} \end{bmatrix} \mathbf{W}_M, \end{aligned} \tag{3}$$

which leads to the structure of the pre-/post-filter as in Fig. 3. All design parameters appear in the $\frac{M}{2} \times \frac{N}{2}$ matrices $\mathbf{U}$ and $\mathbf{V}$ of $\mathbf{P}$ and the $\frac{N}{2} \times \frac{M}{2}$ matrices $\hat{\mathbf{U}}$ and $\hat{\mathbf{V}}$ of $\mathbf{T}$.

2.3. Under-Sampled Pre-/Post-Filter Design

In the under-sampled case ($M < N$), the pre-filter $\mathbf{P}$ is a "fat" matrix, thus it has no left inverse, i.e., perfect reconstruction is not achievable. In this case, pre- and post-filters can be designed to minimize the reconstruction error. Since $\mathbf{y} = \mathbf{P}\mathbf{x}$, if $\hat{\mathbf{y}} = \mathbf{y}$ then $\hat{\mathbf{x}} = \mathbf{T}\hat{\mathbf{y}} = \mathbf{T}\mathbf{P}\mathbf{x}$, and the reconstruction error is

$$\mathbf{e} = \hat{\mathbf{x}} - \mathbf{x} = (\mathbf{I} - \mathbf{T}\mathbf{P})\mathbf{x} = \mathbf{E}\mathbf{x}, \tag{4}$$

where $\mathbf{E} \stackrel{\triangle}{=} \mathbf{I} - \mathbf{TP}$. For each pair of $\mathbf{P}$ and $\mathbf{T}$, it is clear that to minimize the mean-square error

$$||\mathbf{e}||_2^2 = \mathbf{x}^T \mathbf{E}^T \mathbf{E} \mathbf{x}, \tag{5}$$

the following relation should hold

$$\mathbf{P} = \mathbf{T}^+. \tag{6}$$

Subsitituting (3) into (6) and noticing the orthogonality of $\mathbf{W}_M$ and $\mathbf{W}_N$, we have

$$\hat{\mathbf{V}} = \hat{\mathbf{U}}^+, \quad \mathbf{V} = \mathbf{U}^+. \tag{7}$$

Obviously, the post filter $\mathbf{T}$ has maximal number of parameters when rank($\mathbf{T}$) = M, i.e., when rank($\hat{\mathbf{U}}$) = rank($\mathbf{U}$) = $\frac{M}{2}$. Hence, in what follows, we will only consider this case. Suppose the Smith forms of $\hat{\mathbf{U}}$ and $\mathbf{U}$ are

$$\hat{\mathbf{U}} = \hat{\mathbf{A}} \begin{bmatrix} \hat{\mathbf{\Lambda}} \\ \mathbf{0}_{\frac{N-M}{2} \times \frac{M}{2}} \end{bmatrix} \hat{\mathbf{B}}, \mathbf{U} = \mathbf{A} \begin{bmatrix} \mathbf{\Lambda} \\ \mathbf{0}_{\frac{N-M}{2} \times \frac{M}{2}} \end{bmatrix} \mathbf{B}, \tag{8}$$

where $\{\hat{\mathbf{\Lambda}}, \mathbf{\Lambda}\}$ are $\frac{M}{2} \times \frac{M}{2}$ diagonal matrices with nonzero diagonal entries; $\{\hat{\mathbf{A}}, \mathbf{A}\}$ are $\frac{N}{2} \times \frac{N}{2}$ orthonormal matrices; and $\{\hat{\mathbf{B}}, \mathbf{B}\}$ are $\frac{M}{2} \times \frac{M}{2}$ orthonormal matrices. It is easy to show that

$$\mathbf{E} = \mathbf{E}^T \mathbf{E} = \mathbf{W}_N \begin{bmatrix} \hat{\mathbf{A}}_l \hat{\mathbf{A}}_l^T & \mathbf{0}_{\frac{N}{2} \times \frac{N}{2}} \\ \mathbf{0}_{\frac{N}{2} \times \frac{N}{2}} & \mathbf{A}_l \mathbf{A}_l^T \end{bmatrix} \mathbf{W}_N, \tag{9}$$

where $\hat{\mathbf{A}}_l$ and $\mathbf{A}_l$ contain the first $\frac{M}{2}$ columns of $\hat{\mathbf{A}}$ and $\mathbf{A}$, respectively. Therefore, the minimal reconstruction error $||\mathbf{e}||_2^2$ is only related to $\mathbf{A}_l \mathbf{A}_l^T$ and $\hat{\mathbf{A}}_l \hat{\mathbf{A}}_l^T$. The remaining parameters in $\hat{\mathbf{A}}$ and $\mathbf{A}$ as well as in $\hat{\mathbf{B}}$, $\mathbf{B}$, $\hat{\mathbf{\Lambda}}$, and $\mathbf{\Lambda}$ can be used to obtain other desired properties such as coding gain.

In practice, instead of minimizing $||\mathbf{e}||_2^2$ directly, we minimize $E\{||\mathbf{e}||_2^2\}$ according to a signal model. Note that

$$E\{||\mathbf{e}||_2^2\} = \sum_{i=0}^{N-1} \sigma_i^2, \tag{10}$$

where σ_i^2 is the i^{th} diagonal entry of the auto correlation matrix of the error signal, $R_{\mathbf{ee}}$, i.e.,

$$R_{\mathbf{ee}} = E\{\mathbf{e}\mathbf{e}^T\} = \mathbf{E} E\{\mathbf{x}\mathbf{x}^T\} \mathbf{E}^T = \mathbf{E} R_{\mathbf{xx}} \mathbf{E}^T. \tag{11}$$

So, to design an under-sampled pre-/post-filter pair, we first search for $\mathbf{A}_l \mathbf{A}_l^T$ and $\hat{\mathbf{A}}_l \hat{\mathbf{A}}_l^T$, which minimize $E\{||\mathbf{e}||_2^2\}$. Then the remaining free parameters of $\hat{\mathbf{U}}$ and $\hat{\mathbf{V}}$ are optimized to achieve other desired properties.

2.4. Over-Sampled Pre-/Post-Filter Design

In the over-sampled case ($M > N$), since $\mathbf{P}$ is now a "tall" matrix, its left inverse is not unique. So, many post-filters can be coupled with a given pre-filter as far as perfect reconstruction is concerned. This extra degree of freedom complicates the over-sampled design. A typical design approach is optimizing some design criteria (e.g. coding gain) using the matrices $\hat{\mathbf{V}}$, $\mathbf{V}$, $\hat{\mathbf{U}}$ and $\hat{\mathbf{U}}$ as free parameters under the perfect reconstruction constraint

$$\hat{\mathbf{U}} \hat{\mathbf{V}} = \mathbf{U} \mathbf{V} = \mathbf{I}. \tag{12}$$

For each pair of $\hat{\mathbf{U}}$ and $\hat{\mathbf{V}}$ (or $\mathbf{U}$ and $\mathbf{V}$), this involves $\frac{MN}{2}$ free parameters and $\frac{N^2}{4}$ constraints.

To make the design problem more tractable, suppose that the Smith forms of $\hat{\mathbf{V}}$ and $\mathbf{V}$ are

$$\hat{\mathbf{V}} = \hat{\mathbf{A}} \begin{bmatrix} \hat{\mathbf{\Lambda}} \\ \mathbf{0}_{\frac{M-N}{2} \times \frac{N}{2}} \end{bmatrix} \hat{\mathbf{B}}, \mathbf{V} = \mathbf{A} \begin{bmatrix} \mathbf{\Lambda} \\ \mathbf{0}_{\frac{M-N}{2} \times \frac{N}{2}} \end{bmatrix} \mathbf{B}, \tag{13}$$

where $\{\hat{\mathbf{\Lambda}}, \mathbf{\Lambda}\}$ are $\frac{N}{2} \times \frac{N}{2}$ diagonal matrices with nonzero diagonal entries, while $\{\hat{\mathbf{A}}, \mathbf{A}\}$ are $\frac{M}{2} \times \frac{M}{2}$ orthonormal matrices, and $\{\hat{\mathbf{B}}, \mathbf{B}\}$ are $\frac{N}{2} \times \frac{N}{2}$ orthonormal matrices. The perfect reconstruction condition implies

$$\begin{aligned} \hat{\mathbf{U}} &= \hat{\mathbf{B}}^{-1} \begin{bmatrix} \hat{\mathbf{\Lambda}}^{-1} & \hat{\mathbf{M}} \end{bmatrix} \hat{\mathbf{A}}^{-1}, \\ \mathbf{U} &= \mathbf{B}^{-1} \begin{bmatrix} \mathbf{\Lambda}^{-1} & \mathbf{M} \end{bmatrix} \mathbf{A}^{-1}. \end{aligned} \tag{14}$$

Here $\hat{\mathbf{M}}$ and $\mathbf{M}$ are arbitrary $\frac{N}{2} \times \frac{M-N}{2}$ matrices. So we can use $\hat{\mathbf{M}}$ and $\mathbf{M}$ instead of $\hat{\mathbf{U}}$ and $\mathbf{U}$ as design parameters. For each pair of $\hat{\mathbf{U}}$ and $\hat{\mathbf{V}}$ (or $\mathbf{U}$ and $\mathbf{V}$), this parameterization successfully reduces the number of free design parameters from $\frac{MN}{2}$ to $\frac{MN}{4} + \frac{N(M-N)}{4} = \frac{2MN-N^2}{4}$ (from 40 to 24 for the case $M = 10$ and $N = 8$).

3. APPLICATIONS

Potential applications of under- and over-sampled pre-/post-filters include low bit-rate image/video coding, transcoding, multi-rate signal conversion, denoising, and error concealment. Two specific application examples in image coding are considered in this paper.

3.1. Very Low Bit-Rate Image Coding

A low bit-rate image coding example is depicted in Fig. 4: the original *Lena* image is compressed to 2048 bytes. The pre-/post-filter pair employed in the left reconstructed portion is an 8-point critically-sampled one with the known highest coding gain [3, 4]. On the other hand, the pre-/post-filter pair employed in the right reconstructed portion is an under-sampled one ($N = 10$, $M = 8$, i.e., the pre-filter $\mathbf{P}$

Fig. 4. Reconstructed 0.125 bpp Lena portions coded by: critically-sampled pre-/post-filter (left, 25.9 dB) and under-sampled pre-/post-filter (right, 27.1 dB).

maps every 10-sample block into an 8-sample block). The pair of operators $\{\mathbf{P}, \mathbf{T}\}$ is designed to minimize the reconstruction error and to maximize the resulting coding gain as discussed earlier. The coding algorithm is from the L-CEB coder presented in [4]. The under-sampled pre-/post-filter pair clearly outperforms the critically-sampled one objectively (a 1.2 dB improvement) as well as subjectively.

3.2. Error Resilient Image Transmission

Fig. 5. Portions of reconstructed Lena image (coded at 0.25 bpp) suffering 25% random block loss: critically- (left, 27.13 dB) and over-sampled case (right, 28.35 dB)

Over-sampled pre-/post-filters add redundancy which has an adverse effect on compression performance. So we rarely use over-sampled pre-/post-filters for image/video compression in the case of perfect communication channels. However, with unreliable communication channels, the added redundancy provides more error resilience. Well-designed coding algorithm should fully exploit the added redundancy and the coding performance degradation should be small. The gain from the increased error resilience can outweigh the cost of decreased coding performance.

The problem of designing error-resilient critically-sampled pre-/post-filters for image concealment has been exploited in [5]. Except for the fact that $\mathbf{P}$ and $\mathbf{T}$ have different structure and sizes of some matrices need to be modified, the design procedure for the over-sampled case remains the same. Since a pre-filter can be coupled with many post-filters to achieve perfect reconstruction, we can dynamically switch post-filters at the decoder without sacrificing perfect reconstruction. This is not possible for critically-sampled pre-/post-filters.

We use an 8×6 pre-filter and a 6×8 post-filter design, which maximizes the coding gain, as an illustrative example. The post-filter is applied if lost coefficients are not involved. However, when coefficient loss occurs, another post-filter, which is designed to minimize the reconstruction error incurred by the lost coefficients, is applied instead.

Fig. 5 shows portions of reconstructed *Lena* image coded at 0.25 bpp using the best critically-sampled error resilient pre-/post-filter P3 in the literature [5] (left) and our new over-sampled pre-/post-filter (right) under the same coding setup in [5]. Clearly, the over-sampled pre-/post-filter outperform its critically-sampled cousin visually as well as objectively (a 1.1 dB improvement).

4. CONCLUSION

This paper generalizes the design of pre-/post-filters for the block DCT coding framework to the over-sampled and under-sampled cases. We demonstrate with practical coding experiments that over- and under-sampled pre-/post-filters can offer much better performances than critically-sampled ones for some applications.

5. REFERENCES

[1] M. Vetterli and J. Kovacevic, *Wavelets and Subband Coding*, Prentice Hall, 1995.

[2] H. S. Malvar, *Signal processing with lapped transforms*, Artech House, Norwood, MA, 1992.

[3] T. D. Tran, J. Liang, and C. Tu, "Lapped transform via time-domain pre- and post-processing," *IEEE Trans. on Signal Processing*, vol. 51, pp. 1557–1571, Jun. 2003.

[4] C. Tu and T. D. Tran, "Context based entropy coding of block transform coefficients for image compression," *IEEE Trans. on Image Processing*, vol. 11, pp. 1271–1283, Nov. 2002.

[5] C. Tu, T. D. Tran, and J. Liang, "Error resilient pre-/post-filtering for DCT based block coding systems," *Proc. IEEE ICASSP 2003*, Hong Kong, Apr. 2003.