Image Coding with Windowed Modulated Filter Banks

Jelena Kovacevic
Department of EE and CTR
Columbia University,
New York, NY 10027, USA

Didier J. Le Gall
Bellcore
445 South street
Morristown, NJ 07960, USA

Martin Vetterli
Department of EE and CTR
Columbia University,
New York, NY 10027, USA

Abstract

Lapped Orthogonal Transforms (LOT) have been introduced to provide a solution to the problem of blocking that occurs with transform coding at low bit-rate. LOT's are a particular case of exact reconstruction filter banks with a paraunitary polyphase representation. For the purpose of Image Coding, a more general class of exact reconstruction filter banks is introduced, that provides an additional degree of freedom in controlling ringing while relaxing the symmetry constraints of LOTs. Windowed modulated filter banks are a very attractive candidate for image coding, since by allowing almost arbitrary selection of the window, they provide a control of the impulse response of the filters in the filter bank. The output of the windowed modulated filter bank are quantized and entropy coded in a way analogous to DCT coefficients. Comparisons are made between the *windowed modulated filter bank* and the Discrete Cosine Transform both from the point of view of mean square error and subjective quality.

I. Introduction

The Discrete Cosine Transform (DCT) is the key element in many commonly used coding algorithms for still images and video [8]. When the transform coefficients are coarsely quantised, two kinds of artifacts occur: a/ noise around sharp discontinuities due to frequency leakage or uncanceled aliasing and b/ blocking effects due to the processing of the image data by block (in general 8x8 or 16x16 picture elements). While the frequency leakage is inherent to a decomposition of the signal in the frequency domain, the blocking effect can be controlled by overlapping data blocs. Lapped Orthogonal Transforms (LOT) have been derived by Cassereau and Malvar [1, 2] and correct the blocking effects by processing blocks of overlapping data. The lapped orthogonal transforms have been recognized [4, 5] as a particular case of perfect reconstruction filter banks with the length of the filter (L) equal to twice the number of channels (N). It turned out that the filters found in [1] and [2] and designed with a square distorsion measure as a goodness criteria had impulse responses with nonzero values at the boundaries and in this case the "blocking effects" while decreased became again visible as ringing.

The theory of perfect recontruction filter banks provides the right framework [5] for the analysis and design of "lapped transforms"; the properties of the polyphase matrix generalise the properties of the transform. A desirable property for the filter bank is that it behaves like an orthogonal transform. When the polyphase matrix $\underline{H}_p(z)$ is paraunitary [6] i.e. satisfies

$$[\underline{H}_p(z^{-1})]^T \cdot \underline{H}_p(z) = \underline{I} \tag{1}$$

then the signal is reconstructed exactly and the reconstruction filter (synthesis) are the same as the analysis filters (with time reversal). Paraunitaricity has been shown in [5] to be equivalent to the "Orthogonality of the Tail" property introduced by Cassereau [1].

The Discrete Cosine Transform is generally preferred to the "optimal" Karhunen-Loeve transform, because of the existence of a fast algorithm; this constraint extends naturally to the lapped transform by demanding that the filters be derived by modulation of a prototype function. Thus, the complexity of implementation of the filter bank will be greatly reduced.

Finally because the mean square error by itself is not a sufficient goodness criteria for the evaluation of image coding techniques, it is desirable that the impulse response of the filters be as smooth as possible. We shall thus look for a modulated filter banks with a paraunitary polyphase matrix, and sufficient freedom in the prototype function (the window). Unfortunately, for all these properties to hold, it has been shown in [5] that the filters cannot be symmetric. With a different approach and independently of the present work, Malvar in [3] has modified the LOT by introducing rotations so as to provide the best possible match to an ideally smooth window-based filter bank, no coding results however have been presented based on this particular LOT. In the next section we shall introduce a class of filter banks that will satisfy all these constraints, the filters will be applied to image coding and the result compared to the DCT on the basis of mean square error and subjective quality.

II. Windowed Modulated Filter Bank

The Windowed Modulated Filter Banks that we propose to use in this paper are a particular case of Pseudo Quadrature Mirror Filter (PQMF) when the filter length is restricted to twice the number of channel $L = 2N$. Pseudo-QMF filters have been proposed as an extension to N channels of the classical two-band QMF filters [9, 10, 11]. Pseudo-QMF analysis/synthesis systems achieve in general only the cancellation of the main aliasing term. However when the filter length is restricted to $L = 2N$, they can achieve perfect reconstruction under certain conditions. The main advantages of pseudo-QMF filters are their low computational complexity as well as the fact that the window function can be tuned to satisfy additional design constraints while maintaining the exact reconstruction property.

A family of pseudo-QMF filter banks that achieves main aliasing cancellation has been designed in [11] and is of the form:

$$h_k(n) = h_{pr}(n) \cdot cos(\frac{2\pi(2k+1)}{4N}\left(n - \left(\frac{L-1}{2}\right)\right) + \phi_k) \tag{2}$$

CH2673-2/89/0000-1949 $1.00 © 1989 IEEE

for the analysis filters ($h_{pr}(n)$ is the impulse response of the window). In the general case, the main aliasing term is canceled for the value of the phase:

$$\phi_k = \frac{\pi}{4} + k\frac{\pi}{2} \quad (3)$$

In the special case $L = 2N$ the filter bank provides exact reconstruction for this value of the phase. For an arbitrary symmetric window $h_{pr}(n)$ exact reconstruction is preserved if $h^2_{pr}(i) + h^2_{pr}(N-1-i)$ is non zero for $i = 0, 1, \cdots, N-1$ furthermore paraunitaricity of the polyphase matrix holds if and only if

$$h^2_{pr}(i) + h^2_{pr}(N-1-i) = 1 \quad (4)$$

This property is important because if it fails to hold, an inverse window will appear at the synthesis that might reduce the benefits of windowing. The fact that a filter bank derived by modulation of a window function can lead to exact reconstruction in the case $L = 2N$ and the importance of eq.(4) had also been recognized by Princen et al. [7] in a slightly different context. Independently of this work, Schiller has recently applied Prinzen's Filter Bank to image coding [12].

A. Selection of a window

Three windows are designed to evaluate the effect of windowing, all windows are symetric, and satisfy eq. (4) thus maintaining the paraunitaricity of the filter bank. The first window W1 is a flat window (in fact no window) in order to verify the ringing effect in absence of adequate windowing; poor results are to be expected with that window. The frequency response of the unwindowed filter bank is represented in figure 1. and clearly exhibits significant crosstalk between the channels. This crosstalk is due to very important sidelobes extending over the whole frequency range. The second window W2 is designed by sampling a sine function in a way as to span the whole length of the filter and ensure maximum smoothness; the window

$$h_{pr}(i) = sin(\frac{\pi(i)}{2(N-1)}) \quad i = 0, \cdots, N-1 \quad (5)$$

satisfies naturally equation (4). The frequency response of the modulated filter bank with window W2 is illustrated in figure 2., the window has considerably reduced the side lobes and the crosstalk between channels. A third window (W3) is designed also by sampling a sine but by limiting the span of the window to the $2N-4$ central values, thus constraining the four central values to 1.

III. Experiments with windowed Modulated filter bank

In order to have an element of comparison a simple quantization and entropy coding technique will be used and applied to both the outputs of the DCT and the Windowed Modulated Filter Bank. In order to be as simple and general as possible, the same uniform quantizer will be applied to all channels (except DC) and the quantized coeffients will be entropy coded according to the technique proposed in the paper of Chen and Pratt [8]; the two dimensional coefficients are zig-zag scanned, then, run-length and amplitude coded. Rather than designing Huffman codes based on a limited number of test pictures, and solely for the purpose of comparison, the entropy of run-length and amplitude symbols will be computed. With this simple technique, quality and bit rate are controlled by a single parameter: the quantization step. The quantizer will be adjusted to yield a given bit-rate and the results compared both in term of MSE and of subjective quality. We shall use separable 2D-filter banks with 64 channels generated from a one-dimensional modulated filter bank with N=8, L=16; these filter banks will be compared to 8x8 two dimensional DCT.

A. Influence of the window

The full span sine window (W2) of eq. 5 provides the best mean square error as well as subjective quality, the limited span sine window (W3) yields a quite similar result but with a somewhat inferior mean square error. The ringing with the flat window is very annoying.

B. Comparison with the DCT

Overall, even with the best window, the performance of the windowed modulated filter bank is somewhat poorer than the performance of the DCT from the point of view of mean square error and even subjective quality (blocking has been replaced by periodic spots on the face of "Lenna"). Carefull examination of the filter bank showed that the baseband filter gave a less than appropriate DC term and that some of the DC energy was "leaking" to the higher coefficients. To a large extend this phenomenon was not as strong with the filters centered at higher frequencies because images always contain DC energy (with the exception of frame differences) but contain less often narrow peaks of higher frequency. This suggests a modification to the Windowed Modulated Filter Bank by removing the DC energy Prior to applying the filter bank.

C. Modified Scheme: DC substraction

In this modified scheme, the average value of the 8*8 block is substracted from the input image prior to applying the filter bank (see figure 3.); the average value is coded separately. Strictly speaking the Modified Windowed Modulated filter bank is no longer critically decimated since it has N^2+1 channels. For large number of channel (transform coefficients) (64 or 256) the coding loss caused by an additional channel is almost negligible. In fact the results after DC substraction are very good, while the mean square error is slightly above the mean square error of DCT coding (see table 1.), the blocking effects have disappeared even at very low bit rate and the technique is clearly superior on smooth areas (lady's face figure 4. and 5.) and adequate on textures and edges. Figure 6. and 7. show enlargement of the face of the image *Lenna*: the windowed filter bank provides a much smoother face completely free of blocking, while the DCT suffers from heavy blocking particularly annoying in the near uniform regions (cheeks); in both cases the quantization error (very coarse quantizers are used in order to achieve 0.5 b/p) results in uncancelled aliasing taking the form of strange periodic patterns.

IV. Conclusions

The proposed coding scheme with a windowed modulated filter bank has shown a performance in many ways comparable to the discrete Cosine transform. The subjective quality is enhanced by the quasi-disappearance of the so-called blocking effect, while the mean square error is slighthly higher. For low bit rate applications were the blocking effects of DCT greatly affect the subjective quality, Windowed Modulated Filter Banks provide an interesting alternative that should be studied in more detail.

Acknowledgements: The first author would like to thank Pacific Bell who supported this research in part. Support from the American National Science Foundation under grants CDR-84-21402 and MIP-88-08277 is also acknowledged.

References

[1] P. Cassereau, *"A new class of optimal unitary transforms for image processing,"* S.M. Thesis, Dept. EE and CS, Massachusetts Institute of Technology, May 1985.

[2] H.S. Malvar, *"Optimal pre- and post-filtering in noisy sampled data systems,"* Ph.D. Thesis, Dept. EE and CS, Massachusetts Institute of Technology, Aug. 1986.

[3] H.S. Malvar *"The LOT: a Link between Block Transform Coding and Multirate Filter Bank"* Proc. IEEE Int. Symposium on Circuit and Systems, Helsinki, June 1988.

[4] M. Vetterli and D. Le Gall, *"Perfect reconstruction FIR filter banks: lapped transforms, pseudo-QMF's and paraunitary matrices,"* Proc. IEEE Int. Symposium on Circuit and Systems, Helsinki, June 1988.

[5] M. Vetterli and D. Le Gall, *"Perfect Reconstruction FIR Filter Banks: Some Properties and Factorizations"*, IEEE transaction on ASSP, July 1989.

[6] P.P. Vaidyanathan, *"Quadrature Mirror Filter Banks, M-band Extensions and Perfect-Reconstruction Technique,"* IEEE ASSP Magazine, Vol. 4, No. 3, pp.4-20, July 1987.

[7] J. Princen, A. Johnson and A. Bradley, *"Sub-band/ transform coding using filter bank designs based on time domain aliasing cancellation,"* Proc. ICASSP-87, pp.2161-2164, Dallas, April 1987.

[8] W.H. Chen and W. Pratt, *"Scene adaptive Coder"* IEEE Transactions on Communications, COM-32, 1984.

[9] H.J. Nussbaumer, *"Pseudo QMF filter Bank,"* IBM Technical Disclosure Bulletin, 1981.

[10] H.J. Nussbaumer and M. Vetterli, *"Computationally Efficient QMF Filter Banks,"* Proc. 1984 Int. IEEE Conf. on ASSP, San Diego, March 1984.

[11] J. Masson, and Z. Picel, *"Flexible Design of Computationaly Efficient Nearly Perfect QMF Filter Banks,"* Proc. 1985 IEEE Conf. on ASSP, Tampa, March 1985.

[12] H. Schiller *"Overlapping Block transform for image coding preserving equal number of sample,"* Proc. SPIE 1001, Visual Communications and Image Processing '88, Cambridge, MA.

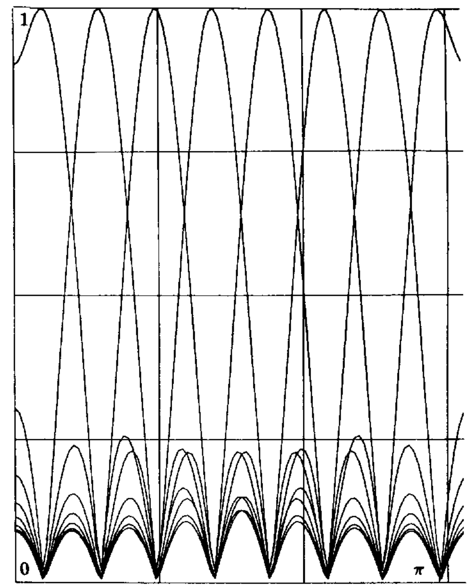

Figure 1: Amplitude frequency response of modulated filter bank with constant window (W1) ($N = 8$, $L = 16$)

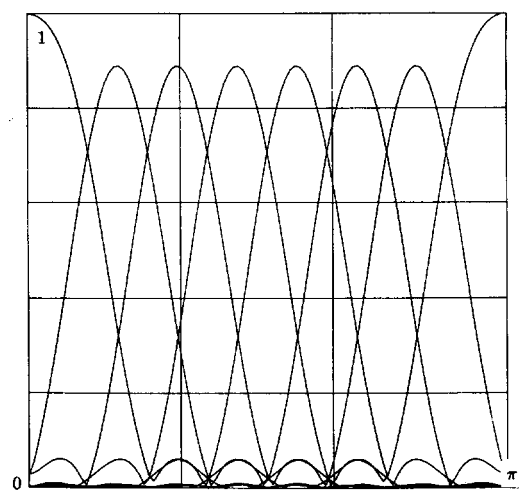

Figure 2: Amplitude frequency response of modulated filter bank with sine window (W2) ($N = 8$, $L = 16$)

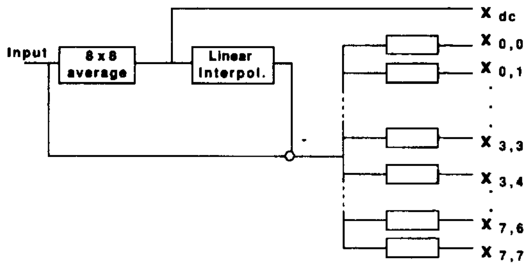

Figure 3: Modified Windowed Modulated Filter Bank with DC Substraction (8x8 Transform)

Figure 4: "lenna" coded at 0.5 bit/p with filter bank (L=16) and window (W2)

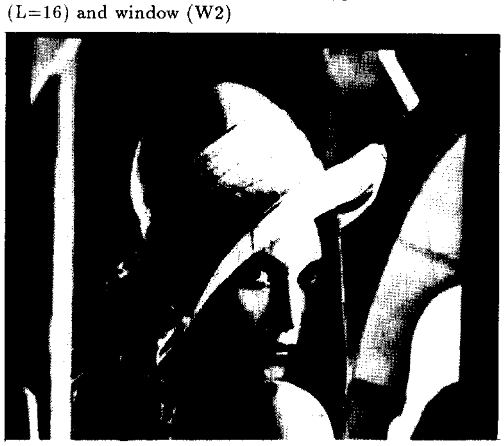

Figure 5: "lenna" coded at 0.5 bit/p with DCT

bits/p	0.7	0.6	0.5
DCT	0.000247	0.000260	0.000285
Window W2	0.000252	0.000269	0.000314
Window W3	0.000258	0.000279	0.000325

Table 1: Mean Square Error as a function of the bit rate for three filter banks

Figure 6: Detail of "lenna" coded at 0.5 bit/p with DCT

Figure 7: Detail of "lenna" coded at 0.5 bit/p with filter bank and window W2