SUBBAND IMAGE CODING USING COSINE MODULATED FILTER BANKS WITH PERFECT RECONSTRUCTION AND LINEAR PHASE*

Helmut Bölcskei[a], Thomas Stranz[a], Franz Hlawatsch[a], and Ralph Sucher[b]

[a]INTHFT, Vienna University of Technology, Gusshausstrasse 25/389, A-1040 Vienna, Austria
email: hboelcsk@aurora.nt.tuwien.ac.at

[b]SIEMENS AG, HL WS DSP, Postfach 801760, D-81617 Munich, Germany
email: Ralph.Sucher@hl.siemens.de

Abstract—We propose a lossy image coding scheme based on the recently introduced *even-stacked* cosine modulated filter banks (CMFBs) that allow linear phase filters in all channels. We describe modifications of the JPEG quantization matrix and zig-zag sequence to match the structure of even-stacked CMFBs, and we discuss the efficient design of even-stacked CMFBs. The rate-distortion performance and perceptual performance of the proposed transform coder are demonstrated using simulation results.

1 INTRODUCTION AND OUTLINE

Many existing image coding schemes are based on signal transforms or filter banks (FBs) [1, 2]. *Cosine modulated FBs* (CMFBs) [1]-[5] are particularly attractive since they can be implemented efficiently using the DCT. However, traditional ("odd-stacked") CMFBs [1]-[5] do not allow perfect reconstruction (PR) *and* linear phase filters in all channels [5]. This is important as nonlinear phase filters are known to produce undesirable artifacts in the reconstructed image.

Recently, the new class of *even-stacked* CMFBs has been introduced and shown to allow PR and linear phase filters in all channels [6]-[9]. The analysis FB in a critically sampled, even-stacked CMFB with $2N$ channels and decimation factor $2N$ consists of two partial FBs $\{h_k[n]\}_{k=0,...,N}$ and $\{\bar{h}_k[n]\}_{k=1,...,N-1}$ derived from an analysis prototype $h[n]$ as

$$h_k[n] = \begin{cases} h[n-rN], & k=0 \\ \sqrt{2}\,h[n]\cos\left(\frac{k\pi}{N}n+\phi_k\right), & k=1,..,N-1 \\ h[n-sN]\,(-1)^{n-sN}, & k=N \end{cases}$$

$$\bar{h}_k[n] = \sqrt{2}\,h[n-N]\sin\left(\frac{k\pi}{N}(n-N)+\phi_k\right), \quad k=1,..,N-1$$

where $\phi_k = -\alpha\frac{\pi}{2N}k + r\frac{\pi}{2}$ with $\alpha \in \mathbb{Z}$, $r \in \{0,1\}$, and $s \in \{0,1\}$ with $s = r$ for α even and $s = 1-r$ for α odd. The synthesis FB is defined similarly [7, 8]. Linear phase of all filters $h_k[n]$, $\bar{h}_k[n]$ is guaranteed if the prototype $h[n]$ satisfies $h[\alpha + (2l-1)N - n] = h[n]$ for some $l \in \mathbb{Z}$. Fig. 1(a) shows that, for given $k \in \{1, ..., N-1\}$, the filters $H_k(z)$ and $\bar{H}_k(z)$ are located about the same center frequency. For comparison, the transfer functions in a traditional (odd-stacked) CMFB are depicted in Fig. 1(b).
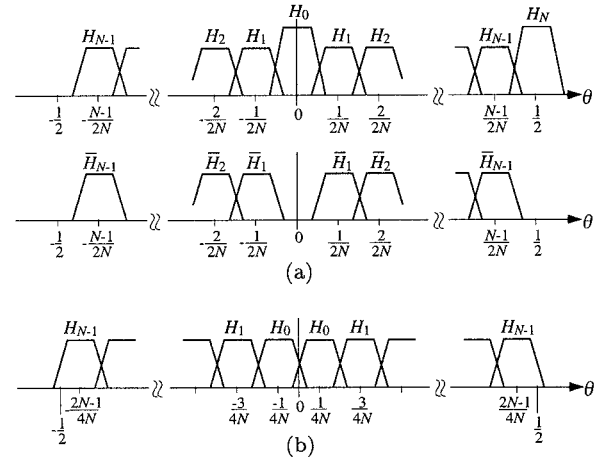
Fig. 1. Transfer functions of the channel filters in
(a) a $2N$-channel even-stacked CMFB and
(b) an N-channel odd-stacked CMFB.

While even-stacked CMFBs are advantageous due to their linear phase, they require modifications of some of the elements used in the JPEG standard. Specifically, in Section 2 we describe a modified quantization matrix and zig-zag sequence matched to even-stacked CMFBs. Section 3 introduces an efficient method for designing even-stacked CMFBs and provides a design example. Finally, in Section 4 simulation results are presented that demonstrate the performance of the proposed image transform coder.

2 IMAGE TRANSFORM CODER BASED ON EVEN-STACKED CMFBs

For the sake of simplicity, we consider only *separable*, 2-D, even-stacked CMFBs. Hence, the transform coder applies a 1-D even-stacked CMFB to the rows of the image and subsequently to the columns of the result. The transform coefficients (subband signals) are then quantized, and finally the quantized coefficients are zig-zag scanned and entropy coded using a Huffman coder. The decoder reverses these steps.

The structure of even-stacked CMFBs requires a modification of the quantization matrix and zig-zag sequence used in
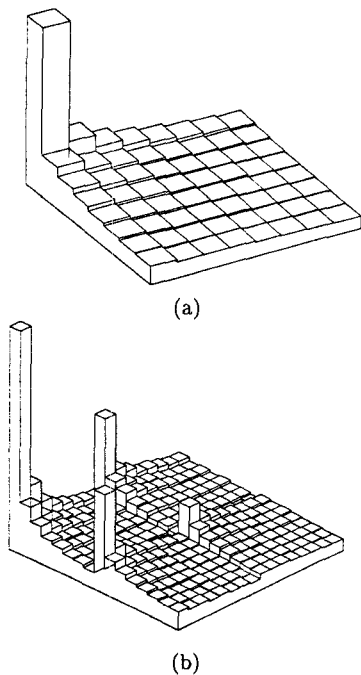
Fig. 2. Typical amplitude (magnitude) distribution in (a) an 8 × 8 DCT-coded block and (b) a 16 × 16 block of subband signals resulting from an even-stacked CMFB.

the JPEG standard. According to Fig. 1(a), an even-stacked CMFB has two filters for each subband, i.e., the filters $H_k(z)$ and $\bar{H}_k(z)$ cover the same frequency band. In order to illustrate this difference from the DCT and the odd-stacked case, typical amplitude distributions in a DCT-coded block and in a block of subband signals resulting from an even-stacked CMFB are shown in Fig. 2(a) and 2(b), respectively. Here, Fig. 2(b) corresponds to an analysis filter vector

$$\mathbf{h}(z) = [H_0(z)\ H_1(z)\ ...\ H_{N-1}(z)\ \bar{H}_1(z)\ \bar{H}_2(z)\ ...$$
$$...\ \bar{H}_{N-1}(z)\ H_N(z)].\quad (1)$$

The reason for this specific ordering (note the position of $H_N(z)$ within $\mathbf{h}(z)$) will be explained further below.

Quantization Matrix. The quantization matrix accounts for the fact that lower frequencies are quantized more accurately than higher frequencies. Based on the amplitude distribution in Fig. 2(b), we now propose a quantization matrix for even-stacked CMFBs. Consider an $(N+1) \times (N+1)$ quantization matrix $\mathbf{Q}_o$ for an $(N+1)$-channel odd-stacked CMFB. ($\mathbf{Q}_o$ can be obtained by suitable linear interpolation or decimation of the entries in the 8 × 8 JPEG quantization matrix.) Then, we define the $2N \times 2N$ quantization matrix for an even-stacked CMFB with $2N$ channels as

$$\mathbf{Q}_e = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix},$$

with the $N \times N$ submatrices $[\mathbf{A}]_{k,l} = [\mathbf{Q}_o]_{k,l}$, $[\mathbf{B}]_{k,l} = [\mathbf{Q}_o]_{k,l+1}$, $[\mathbf{C}]_{k,l} = [\mathbf{Q}_o]_{k+1,l}$, and $[\mathbf{D}]_{k,l} = [\mathbf{Q}_o]_{k+1,l+1}$ where $k,l = 0,1,...,N-1$.

Zig-Zag Scanning. In the DCT-based JPEG standard, the quantized transform coefficients are zig-zag scanned before entropy encoding. This produces long runs of zeros which can be encoded very efficiently. For even-stacked
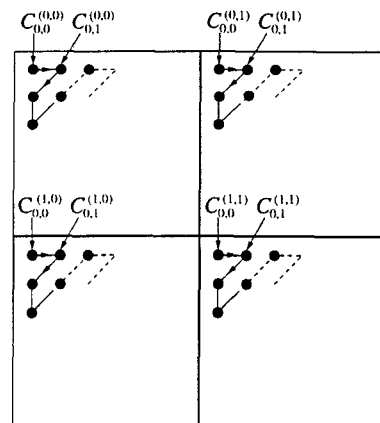

Fig. 3. Subblock zig-zag sequences to be interleaved in the case of even-stacked CMFBs.

CMFBs, the fact that the filters $H_k(z)$ and $\bar{H}_k(z)$ for given $k \in \{1,...,N-1\}$ cover the same frequency band requires a modified zig-zag sequence since standard zig-zag scanning would result in significantly shorter runs of zeros.

The specific filter ordering in (1) yields 4 subblocks of equal size $N \times N$, whose amplitude distributions resemble that in one DCT-coded block (cf. Fig. 2). This suggests to define the modified zig-zag sequence by interleaving the zig-zag scanned values of the four $N \times N$ subblocks (see Fig. 3). That is, the modified zig-zag sequence for even-stacked CMFBs is

$$\mathbf{s} = \left(C_{0,0}^{(0,0)}\ C_{0,0}^{(0,1)}\ C_{0,0}^{(1,0)}\ C_{0,0}^{(1,1)}\ C_{0,1}^{(0,0)}\ C_{0,1}^{(0,1)}\ C_{0,1}^{(1,0)}\ C_{0,1}^{(1,1)}\ ...\right),$$

where the superscript $^{(i,j)}$ $(i,j = 0,1)$ and subscript $_{k,l}$ $(k,l = 0,1,...,N-1)$ in $C_{k,l}^{(i,j)}$ identify the subblock and the frequency index, respectively.

An alternative ordering of the analysis filters is

$$\mathbf{h}(z) = [H_0(z)\ H_1(z)\ \bar{H}_1(z)\ ...\ H_{N-1}(z)\ \bar{H}_{N-1}(z)\ H_N(z)].$$

This ordering, like the one in (1), requires a suitable modification of the quantization matrix. However, there is no need for a modified zig-zag sequence, i.e., a standard JPEG-type zig-zag scanning can be performed on the resulting $2N \times 2N$ block of subband signals. We observed that both orderings yield similar performance.

3 CMFB DESIGN

In this section, we present an efficient method for the design of even-stacked CMFBs. This method extends a method for the design of odd-stacked CMFBs with near-PR [10] to the design of paraunitary, even-stacked CMFBs. We assume an FIR prototype $h[n]$ of length $L = \alpha + (2m + 1)N + 1$ with some $m \in \mathbb{Z}$ and α according to the definition of ϕ_k (cf. Section 1). Our design minimizes a quadratic cost function $C_h = \mathbf{h}^T \mathbf{P} \mathbf{h}$, where $\mathbf{h}$ is the vector of filter coefficients $h[n]$ and $\mathbf{P}$ is a matrix depending on the choice of C_h. The (necessary and sufficient) condition for paraunitarity [7, 8],

$$N \sum_{i=-\infty}^{\infty} h[n - iN]\ h[n - (i + 2l)N] = \delta[l],\quad (2)$$

enters into the minimization of C_h as a (quadratic) side constraint. Adopting an iterative approach to the constrained
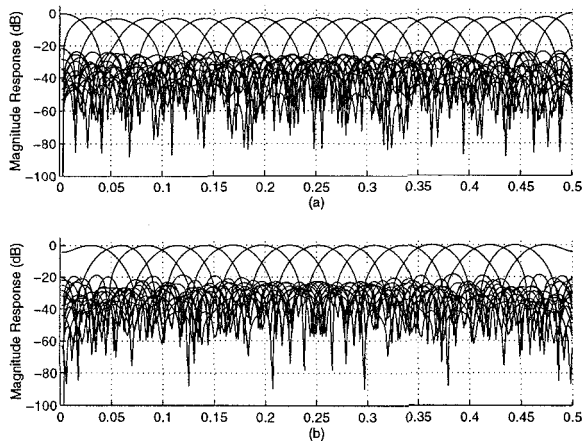
Fig. 4. Design of paraunitary even-stacked CMFB (number of channels $2N = 36$, filter length $L = 67$) via linearized method: Frequency responses of (a) filters $H_k(z)$ $(k = 0, ..., N)$, (b) filters $\bar{H}_k(z)$ $(k = 1, ..., N-1)$.

minimization of C_h, the quadratic side constraint (2) can be linearized by approximating (2) as

$$c_{h_k,h_{k-1}}[l] \stackrel{\triangle}{=} N \sum_{i=-\infty}^{\infty} h_{k-1}[n - iN] \, h_k[n - (i + 2l)N] = \delta[l],$$

$$(3)$$

where k is the iteration index. This is indeed linear in the current impulse response $h_k[n]$. The kth iteration of the constrained optimization now goes as follows:

1. Minimize C_h subject to the linear side constraint (3),

$$\mathbf{h}_k = \arg \min_{\mathbf{h}} \mathbf{h}^T \mathbf{P} \mathbf{h} \quad \text{subject to } c_{h_k,h_{k-1}}[l] = \delta[l].$$

Note that $\mathbf{h}_k$ depends on $\mathbf{h}_{k-1}$ via the side constraint.

2. If $\|\mathbf{h}_k - \mathbf{h}_{k-1}\| < \epsilon$, where ϵ is a prescribed tolerance level, stop the iteration. The result is $\mathbf{h}_{k-1}$. Otherwise, set $\mathbf{h}_k \to \tau \mathbf{h}_k + (1 - \tau)\mathbf{h}_{k-1}$ (with some fixed τ between 0 and 1) and go back to Step 1 with $k \to k + 1$.

This iteration is initialized using a linear-phase lowpass prototype h_0 that is designed by means of a standard method such as the Remez exchange algorithm. It should be noted that the iteration stops only when $\mathbf{h}_k \approx \mathbf{h}_{k-1}$, in which case (3) is a good approximation to the true paraunitarity condition (2) and thus paraunitarity is approximately satisfied. By choosing the tolerance level ϵ sufficiently small, paraunitarity can be achieved with arbitrary accuracy.

This algorithm can be expected to converge once that the prototype changes little in each iteration. Otherwise, the approximation of the paraunitarity side constraint will be poor and the algorithm may not converge. In our design experiments, we observed that $\tau = 0.5$ (as recommended in [10]) yields good convergence properties in most cases; however, in some cases the algorithm did not converge at all.

Fig. 4 shows an even-stacked CMFB that has been obtained by our iterative design method. The cost function was defined as $C_h = \int_{\theta_s}^{1/2} \theta^2 \, |H(e^{j2\pi\theta})|^2 \, d\theta$ with $\theta_s = 1/18$ (this can easily be expressed as a quadratic form $\mathbf{h}^T \mathbf{P} \mathbf{h}$). The tolerance level was chosen as $\epsilon = 10^{-4}$, which resulted in an excellent approximation to a truly paraunitary CMFB.
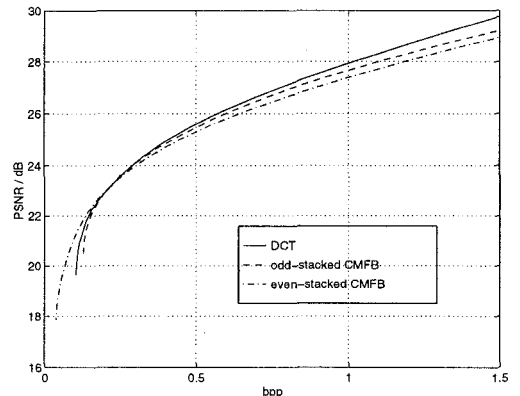

Fig. 5. Rate-distortion characteristics of the DCT and odd- and even-stacked CMFBs (for image "Lena").

4 CODER PERFORMANCE

We finally present coding experiments that demonstrate the performance of even-stacked CMFBs.

Rate-Distortion Performance. Fig. 5 demonstrates that the rate-distortion performance achieved with even-stacked CMFBs is comparable to that of the DCT-based JPEG standard and of odd-stacked CMFBs if (i) the modified quantization matrix and zig-zag sequence as described in Section 2 are used, and (ii) the number of channels in the even-stacked CMFB is twice that of the odd-stacked CMFB.[1]

Perceptual Performance. Fig. 6(a) shows the image "Bridge." The result of coding this image using the DCT-based JPEG standard at 0.127 bpp (PSNR = 21.37 dB) is depicted in Fig. 6(b) and seen to exhibit heavy blocking artifacts due to the very low bitrate. Fig. 6(c) shows "Bridge" coded at 0.137 bpp (as before, PSNR = 21.37 dB) using an odd-stacked CMFB with $N = 8$ channels and filter length $L = 32$. Here, blocking artifacts are not visible but the image appears blurred; this is essentially due to the nonlinear phase of the filters. Finally, Fig. 6(d) shows "Bridge" coded at 0.144 bpp (again, PSNR = 21.37 dB) using an even-stacked CMFB with $2N = 16$ channels and the same prototype as in the odd-stacked CMFB. Here, the details appear much sharper than in Fig. 6(c). Thus, for equal PSNR, the even-stacked CMFB outperforms the odd-stacked CMFB and also the DCT from a perceptual point of view.

5 CONCLUSION

For subband image coding, the recently introduced even-stacked CMFBs are advantageous over conventional (odd-stacked) CMFBs due to their linear phase. This paper proposed a lossy subband image coding scheme based on even-stacked CMFBs. It was demonstated that, relative to JPEG and odd-stacked CMFBs, the proposed coder has similar rate-distortion performance and superior perceptual performance. Furthermore, an efficient algorithm for the design of even-stacked, paraunitary CMFBs was presented.

---

[1] An even-stacked CMFB with $2N$ channels has effectively N subbands (i.e., center frequencies) and the filter bandwidth equals that in an N-channel odd-stacked CMFB. Since the rate-distortion performance of subband coders depends critically on the filter bandwidth (the channels are coded mutually independently), it is not surprising that the performance of an N-channel odd-stacked CMFB is similar to that of a $2N$-channel even-stacked CMFB.
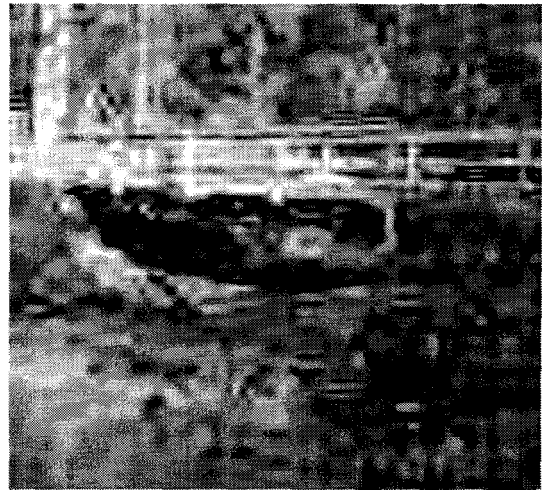
(a)
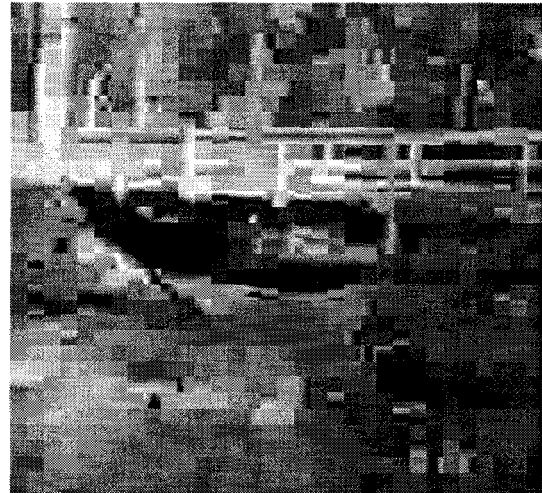
(b)
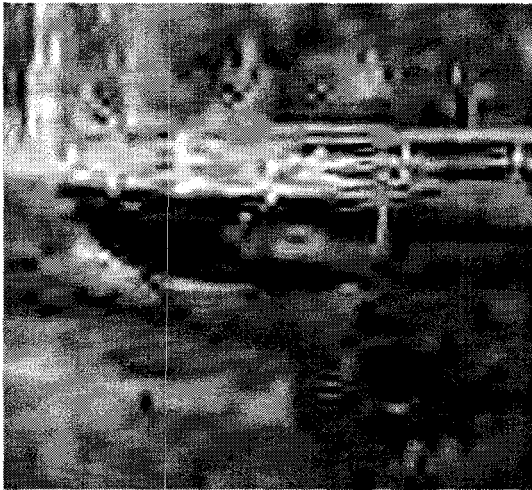
(c)

(d)

Fig. 6. Subband image coding using the DCT and odd- and even-stacked CMFBs: (a) original image "Bridge," (b) "Bridge" coded at 0.127 bpp (PSNR = 21.37 dB) using a DCT of block size 8, (c) "Bridge" coded at 0.137 bpp (PSNR = 21.37 dB) using an odd-stacked CMFB with N = 8 channels, (d) "Bridge" coded at 0.144 bpp (PSNR = 21.37 dB) using an even-stacked CMFB with 2N = 16 channels. Note that in (b)-(d) the PSNR is identical.

References

[1] M. Vetterli and J. Kovačević, *Wavelets and Subband Coding.* Englewood Cliffs (NJ): Prentice Hall, 1995.

[2] P. P. Vaidyanathan, *Multirate Systems and Filter Banks.* Englewood Cliffs (NJ): Prentice Hall, 1993.

[3] H. S. Malvar, *Signal Processing with Lapped Transforms.* Artech House, 1992.

[4] R. D. Koilpillai and P. P. Vaidyanathan, "Cosine-modulated FIR filter banks safisfying perfect reconstruction," *IEEE Trans. Signal Processing,* vol. 40, pp. 770-783, April 1992.

[5] R. A. Gopinath and C. S. Burrus, "Some results in the theory modulated filter banks and modulated wavelet tight frames," *Applied and Computational Harmonic Analysis,* vol. 2, pp. 303-326, 1995.

[6] R. A. Gopinath, "Modulated filter banks and wavelets — A unified theory," in *Proc. IEEE ICASSP-96,* (Atlanta, GA), pp. 1585-1588, May 1996.

[7] H. Bölcskei and F. Hlawatsch, "Oversampled cosine-modulated filter banks with linear phase," in *Proc. IEEE ISCAS-97,* (Hong Kong), pp. 357-360, June 1997.

[8] H. Bölcskei and F. Hlawatsch, "Oversampled cosine modulated filter banks with perfect reconstruction: Analysis and design," *IEEE Trans. Circuits and Systems II,* submitted Dec. 1996.

[9] Y.-P. Lin and P. P. Vaidyanathan, "Linear phase cosine modulated maximally decimated filter banks with perfect reconstruction," *IEEE Trans. Signal Processing,* vol. 42, pp. 2525-2539, Nov. 1995.

[10] H. Xu, W. S. Lu, and A. Antoniou, "Efficient iterative design method for cosine-modulated QMF banks," *IEEE Trans. Signal Processing,* vol. 44, pp. 1657-1668, July 1996.