LOW BIT-RATE IMAGE CODING BASED ON PYRAMIDAL DIRECTIONAL FILTER BANKS

Yilong Liu, Truong T. Nguyen and Soontorn Oraintara

Dept. of Electrical Engineering, University of Texas at Arlington, 416 Yates St., Arlington, TX, 76019
Email: {yilong, nttruong}@msp.uta.edu, oraintar@uta.edu

ABSTRACT

This paper presents a novel embedded image coding system based on the pyramidal multidirectional image representation. The multiscale filter bank is a combination of the overcomplete pyramidal directional filter bank at higher scales and the traditional maximally-decimated wavelet filter bank at lower scales, to provide a sparse image representation. The coding algorithm then efficiently clusters the significant coefficients using progressive morphological operations. Context models for arithmetic coding are designed to exploit the intra-band dependency and the correlation existing among the neighboring directional subbands. Experimental results show that the proposed coding algorithm outperforms the current state-of-the-art wavelet based coders, such as JPEG2000, for images with directional features.

1. INTRODUCTION

Wavelet and filter bank (FB) have been one of major research topics in signal processing for the last two decades. The discrete wavelet transform (DWT) has been shown to be an optimal representation of one-dimensional (1-D) piece-wise smooth signals [1] and has found widespread uses in many signal and image processing applications. However, the separable wavelets are not effective in capturing line discontinuities since they cannot take advantage of the geometrical regularity of image structures. Image transitions such as edges are expensive to represent by wavelets. Therefore, integrating the geometric regularity in the image representation is a key challenge to improve the performance of current image coders.

Recently, Candès and Donoho constructed the curvelet transform [2], and proved that it is an essentially optimal representation of two-variable functions, which are smooth except at discontinuities along $\boldsymbol{C}^2$ (twice differentiable) curves. The nonlinear approximation of a function f, $f_M^{(c)}$, reconstructed by M curvelet coefficients has an asymptotic decay rate of $\|f - f_M^{(c)}\|^2 \leq CM^{-2}(\log_2 M)^3$. This decay rate of the approximation error is a significant theoretical improvement compared to those by wavelet and Fourier coefficients, which are $O(M^{-1})$ and $O(M^{-1/2})$, respectively [1]. Since the space of smooth functions with singularities along $\boldsymbol{C}^2$ curves is similar to natural images with regions of continuous intensity and discontinuous along edges, there is strong motivation for finding similar transform in the discrete domain.

In [3], a pyramidal directional filter bank (PDFB) is proposed in order to implement the contourlet transform. The proposed structure of the PDFB, which is a combination of the Laplacian pyramid and the conventional directional filter bank (DFB) [4], has the advantages of multiresolution and multidirection. It attempts to separate the low frequency components from the rest directional components, and reiterates with the same DFB in the lowpass band forming a pyramid structure. The transform can achieve the asymptotic optimal approximation power as the curvelet transform. In fact, the PDFB is equivalent to an overcomplete FB with one lowpass and 2^n directional highpass subbands. One criterion for the FB to achieve its theoretical performance is that the directional filters must have very good frequency characteristics.

In this work, we propose a novel embedded image coding scheme with an overcomplete PDFB. The morphological operation is employed progressively to identify clusters of significant coefficients in each bit-plane. Context-based arithmetic coding is used to encode these significant coefficients. We design the context models so that the intra-band and inter-band correlations of the overcomplete PDFB can be well exploited.

The rest of the paper is organized as follows. Section 2 presents the PDFB and the conditions on the lowpass filters needed to reduce aliasing. Section 3 describes the proposed PDFB-based coding algorithm. The experimental results and discussions are presented in section 4. Section 5 summarizes this work.

2. THE REDUCED-ALIASING PDFB

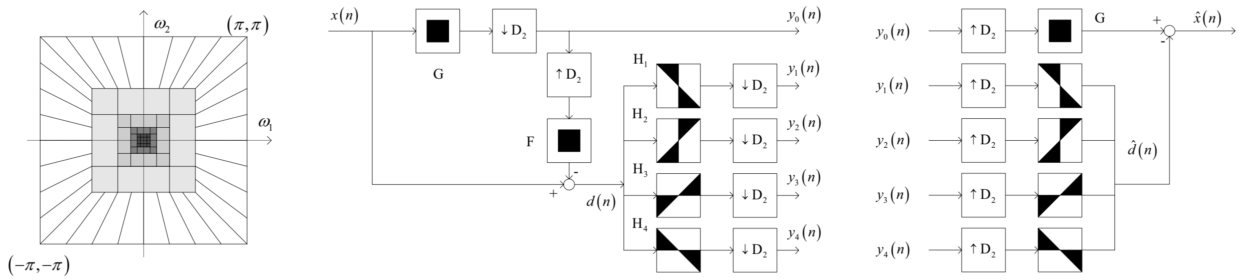
A PDFB consisting of a Laplacian pyramid and a four-channel DFB (Fig. 1(b)) is equivalent to an overcomplete five-channel FB with a decimation matrix $\boldsymbol{D}_2 = 2\boldsymbol{I}$, ($\boldsymbol{I}$ is the identity matrix). The overall directional filters can be expressed in a closed-form formula from the Laplacian and the DFB filters [5]. It is shown that the aliasing occurs in these filters is from iterations in the Laplacian pyramid. It can be reduced or removed by requiring the two filters of the pyramid to satisfy the Nyquist property. This condition means that the frequency responses of the two lowpass filters in the pyramid ($F(\boldsymbol{\omega})$ and $G(\boldsymbol{\omega})$ in Figs. 1(b) and (c)) should have the passband regions (including the transition bands) limited in $[-\pi/2, \pi/2]^2$. $F(\boldsymbol{\omega})$ and $G(\boldsymbol{\omega})$ in this paper are designed to have transition bands in the region $0.3\pi < |\omega_i| < 0.6\pi$.

Fig. 1(a) depicts the frequency supports of the image decomposition, in which the PDFB is used at the highest resolution, and the DWT is used in the next four resolutions. To illustrate the construction, the analysis and synthesis sides of a PDFB with four directional filters are described in Fig. 1(b) and (c), respectively. Note that the conventional DFB is implemented by a binary tree of two-channel FBs [4], and the four filters $H_i(\boldsymbol{z})$ in Fig. 1(b) represent the equivalent directional filters of the DFB tree.

3. IMAGE CODING USING HYBRID OVERCOMPLETE PDFB AND WAVELET FB

Since significant coefficients of the PDFB are sparser than those of wavelet, utilizing the statistical properties of the transform coefficients is a crucial task in the design of high-performance PDFB-based coders. An explicit way to solve this problem is to classify

This work was supported in part by NSF Grant ECS-0528964.



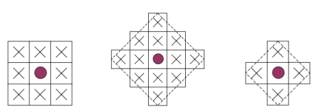
(a) (b) (c)

Fig. 1. The PDFB employed in image coding. (a) The partitioning of the frequency plane by a hybrid PDFB and wavelet decomposition, and (b) The analysis side of a PDFB; the DFB has four directional subbands and (c) the synthesis PDFB corresponds to analysis PDFB in (a).

the coefficients of each subband into two subsets that separate insignificant and significant coefficients. Thus the quantization and estimated models can be adapted to each subset independently.

A potential approach to exploit this behavior has been introduced with the morphological operation for wavelet coding [6]. Based on the observation that clusters tend to grow both in spatial and in frequency domain, the previously detected significant coefficients are used as seeds for the search of new significant ones. The clustering trend of significant coefficients also exists in the PDFB bands. This suggests using a morphological dilation to identify the significant coefficients in the PDFB subbands before the coding step. Thus different probability models can be estimated for significant and insignificant coefficients separately. On the other hand, although the overcomplete transform introduces more coefficients to be coded, strong correlations exist between neighboring directional subbands that can be adopted to improve probability modeling.

3.1. Progressive morphological dilation



(a) SE_1 (b) SE_2 (c) SE_3

Fig. 2. Structuring elements used in the progressive morphological dilation.

Erosion and dilation are two morphological operators frequently adopted [7]. Let I be a binary-valued image where dilation will be applied. Dilation of a given set $A \subseteq I$ with set $B \subseteq I$ is defined by

$$A \oplus B = \cup_{b \in B} A_b, \tag{1}$$

where B is a binary-valued array called a *structuring element* (SE), A_b denotes the translation of A to a point b [8]. The dilation operation produces an enlarged set, $A \oplus B$, which can also be written as $A \cup (A \oplus B \setminus A)$, where $(A \oplus B \setminus A)$ represents the set of new points obtained by dilation. If A is the set of previously detected significant points, the points in set $(A \oplus B \setminus A)$ have a much higher probability to be recognized as significant.

Organizing and representing each subband of PDFB as irregular shaped clusters of significant coefficients provide an efficient way for encoding. Exploiting cross-scale dependency and correlation between neighboring DFB bands of the overcomplete pyramid can further improve the detection accuracy of significant coefficients. Moreover, shaping the clustering boundaries with less cost should be also considered with adaptive SE's [9]. A progressive embedded coding algorithm is proposed to code the significant coefficients based on morphological dilation. Six passes are performed for coding the coefficients at each bit-plane. Three different SE's, as shown in Fig. 2, are adopted in these coding passes for variant steps of dilation.

1. *Significance detection* (SD) pass: The intra-band dependency is exploited in this pass. The square SE_1 in Fig. 2(a) is adopted to detect new significant positions based on significant neighbors in the previous bit-planes.

2. *Cross-band prediction* (CBP) pass: Based on the fact that the inter-band correlation exists between the successive decomposition levels, a significant cluster in a children subband can be predicted by those in the parent subband. The diamond SE_2 in Fig. 2(b) is employed to dilate around the associated children positions corresponding to each significant position in the parent subband.

3. *Neighboring correlation prediction* (NCP) pass: This pass is designed to capture the redundancy within the overcomplete directional subbands. If the two spatial filters have small angle difference in their principal directions, their corresponding decimated subbands exhibit significant dependency between those coefficients at the same positions relative to their upper-left corners. Hence, the distribution of significant coefficients of each directional subband is highly correlated to their neighboring subbands (*cousin subbands*). For the current subband, those coefficients associated with the identified significant coefficients in the cousin subbands are coded. SE_1 is employed to implement the dilation around the resulting new significant coefficients.

4. *Boundary shaping* (BS) pass: Typically, on the boundaries of a large cluster, there are a few scattered significant coefficients located in small isolated clusters. It is difficult to forecast the dimension of these isolated clusters. Hence adaptive dilation is expected to search around a cluster. The smaller rood SE_3 in Fig. 2(c) is adopted for dilation on the previously formed cluster boundaries which are identified by the insignificant positions. The boundary extension is adaptively controlled based on the occurrence of new significant positions detected. It stops when the recursive dilation results in no more new significant positions.

5. *Sparse significant coefficients detection* (SSD) pass: Although most of the significant coefficients have been recognized by

the previous four passes, there are still few sparse significant coefficients remained undetected. Those coefficients that have not been processed are scanned and coded in a raster order. The dilation using SE_1 is implemented if a new significant coefficient is found.

6. *Magnitude refinement* (MR) pass: Refine those significant coefficients that have been recognized in previous bit-planes.

3.2. Context modeling

According to the conditional probability theory, context modeling can exploit the inter-symbol redundancy by switching between different probability models [10]. In embedded coding systems, the coding is conducted on a series of significance maps that correspond to a set of decreasing thresholds. After each pass of coding, all coefficients are quantized to specific values and can be used as context information. Thus, a significant context template, as described in Fig. 3, is defined to exploit the coefficients that have been coded by the previous passes or bit-planes.

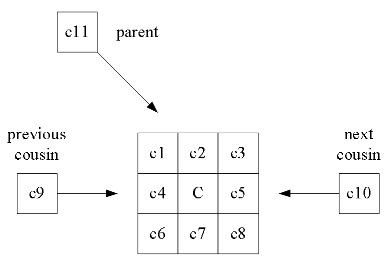


Fig. 3. Context modeling template.

In Fig. 3, C is the current symbol that needs to be coded. $c_1, ..., c_{11}$ are reconstructed values up to the current coding pass of the corresponding coefficients. We define a series of reconstruction matrices $\mathbf{Y}_s^n$, in which each element is the reconstructed value of the corresponding coefficient, n indicates the order of coding passes, s denotes the subband index. Assume C is the (i, j)th element of the binary significance map in pass n of subband s. Thus the intra-band correlation is derived from the 8 neighboring coefficients,

$$c_m = y_s^{n'}(i + a, j + b),\ a, b = \text{-1, 0, or 1},\ m = 1, ..., 8. \tag{2}$$

where $n' = n - 1$ or n since the morphological dilation does not guarantee a raster scan order for the coefficients that have been coded. The coefficients of parent and cousin subbands are also used to exploit the inter-band dependency

$$\begin{aligned} c_9 &= y_{s-1}^{n}(N_c(i), N_c(j)), \\ c_{10} &= y_{s+1}^{n-1}(N_c(i), N_c(j)), \\ c_{11} &= y_{P(s)}^{n}(P_c(i), P_c(j)), \end{aligned} \tag{3}$$

where $P(\cdot)$ specifies the parent of band s, $N_c(\cdot)$ and $P_c(\cdot)$ are functions for the cousin and parent subbands, respectively, that specify the coordinates corresponding to the positions in the current subband. Considering the weak correlation among the cousin subbands for wavelet, c_9 and c_{10} are not used for coding wavelet coefficients in the simulation. Thus the NCP pass is executed only for those directional subbands.

3.3. Algorithm summary

In this section, we summarize the coding algorithm. Six passes are designed in this approach. The produced progressive bitstream can be truncated at any pass. Let $w_{ij}(k)$ be the coefficient with the coordinate (i, j) relative to the upper-left corner of subband k. $LSC(k)$ and $LIC(k)$ represent the lists of significant and insignificant coefficients, respectively. $D(SE_m)$, $m = 1, 2, 3$, denotes the operation of morphological dilation with the structuring element specified by SE_m. $V[w_{ij}(k)|SE_m] = w_{ij}(k) \oplus SE_m \setminus w_{ij}(k)$, defines the vicinity of $w_{ij}(k)$ generated by dilation with SE_m. Let $P(k)$ and $N(k)$ be the parent and cousin bands of subband k, respectively. $C[w_{ij}(k)]$ and $N[w_{ij}(k)]$ denote the corresponding coefficients in $P(k)$ and $N(k)$ associated with $w_{ij}(k)$.

The PDFB-based coding algorithm is summarized as follows.

1. (*Initialization*): Decompose the image with PDFB. Find the maximum number of bit-planes M and set $n \leftarrow M$.
2. (*SD pass*): If $n = M$, go to step 6).
   Apply $D(SE_1)$ to each entry of $LSP(k)$. Encode all $w \in V[w_{ij}(k)|SE_1]$ that have not been scanned with arithmetic coding, if w is insignificant, attach it to the end of $LIC(k)$; else encode the sign of w, and add it to the end of $LSC(k)$.
3. (*CBP pass*): If $k = 0$, i.e. the LL band, go to step 5).
   For the significant coefficients of $P(k)$, apply $D(SE_2)$ to each entry of $C[w_{ij}(P(k))]$. Similar to *SD pass*, encode each newly scanned coefficient and add these coefficients to the end of either $LSC(k)$ or $LIC(k)$ depending on whether they are significant or not.
4. (*NCP pass*): If subband k is a wavelet band, go to step 5).
   For the significant coefficients of $N(k)$, apply $D(SE_1)$ to each entry of $N[w_{ij}(N(k))]$. Encode each scanned coefficient. Add the coefficient to the end of $LSC(k)$ if it is found to be a new significant coefficient; otherwise append it to $LIC(k)$.
5. (*BS pass*): Apply $D(SE_3)$ to each entry of $LIC(k)$. If a new significant coefficient is found, the recursive dilation is implemented around it until no more new significant coefficient is detected. Encode the scanned coefficients and update $LSC(k)$ and $LIC(k)$ correspondingly.
6. (*SSD pass*): The remaining coefficients are scanned and coded in raster order. However, if one significant coefficient is found, apply $D(SE_3)$ at once, then continue to scan the following coefficients that have not been coded. Update $LSC(k)$ correspondingly.
7. (*MR pass*): If $n = M$, go to step 8);
   Encode the nth bit of those significant coefficients recognized by previous bit-planes.
8. (*New bit-plane*): Empty all LIC's, set $n \leftarrow n - 1$, and go back to step 2).

4. SIMULATION RESULTS

For the sake of comparison, the proposed PDFB-based algorithm is compared with three state-of-art wavelet based coding schemes: SPIHT [11] with arithmetic coding, JPEG2000 [12] and the embedded coder of morphological representation of wavelet data (MRWD) [6]. Two 512×512 grayscale images, Barbara and Lena, are tested. A 5-level decomposition is used for SPIHT, MRWD and JPEG2000 with 9/7 Daubechies wavelet filters. On the other hand, a 5-level

PDFB decomposition, as described in Fig. 1(a), is applied in the proposed scheme, where the four coarser levels are decomposed using 9/7 filters and the level with the finest resolution is decomposed into 16 directional subbands using PDFB.

Table 1. Performance comparison (PSNR [dB]) for Barbara.

Rate(bpp)	SPIHT	MRWD	JPEG2000	PDFB
0.10	24.24	24.15	24.66	25.24
0.15	25.63	25.32	25.93	26.74
0.20	26.63	26.86	27.31	27.84
0.25	27.56	27.51	28.36	28.94
0.30	28.54	28.16	29.24	29.86
0.40	30.09	30.18	30.83	30.97
0.50	31.38	31.31	32.26	32.42

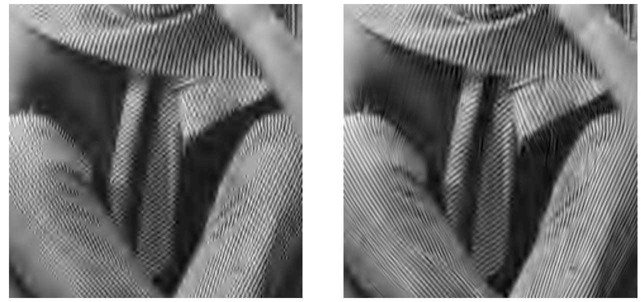
Table 2. Performance comparison (PSNR [dB]) for Lena.

Rate(bpp)	SPIHT	MRWD	JPEG2000	PDFB
0.10	30.17	30.18	29.87	29.86
0.15	31.87	31.55	31.69	31.87
0.20	33.12	33.14	32.97	33.07
0.25	34.09	33.90	34.13	33.98
0.30	34.93	34.57	34.75	34.91
0.40	36.22	36.17	36.09	35.96
0.50	37.20	37.01	37.22	36.78

Table 1 shows the performance of Barbara with different systems in terms of PSNR. The performances are calculated by truncating the embedded bitstreams at different rates during decoding. In the comparison, the proposed algorithm consistently outperforms all the other three schemes. Although the overcomplete transform is employed in the proposed scheme, the improvements are still remarkable for rich-edgy images. This is because of that the PDFB can represent the geometrical regularity of image structures with fewer coefficients and smaller magnitudes. Hence, the proposed algorithm is more efficient at very low bit rates when a few of significant coefficients are actually used for reconstruction. For instance, the PDFB-based coder gains 0.81 dB in PSNR over JPEG2000 at 0.15 bpp for Barbara image. Fig. 4 shows the reconstructed Barbara images at 0.15 bpp for JPEG2000 and the proposed coder. On the other hand, the PDFB-based codec exhibits competitive performance for smooth images, such as Lena. As shown in Table 2, the performance of our codec is comparable to that of the other three coders, and the performance difference is getting smaller as the bit-rate decreases.

5. CONCLUSION

In this paper, we propose a novel embedded image coding scheme using hybrid overcomplete PDFB. The improved design of the PDFB helps to reconstruct images with directional features by fewer coefficients than other image decompositions. Based on the progressive morphological dilation of significant coefficients, the highly concentrated PDFB significant coefficients are identified by exploiting the intra-band statistical dependencies. The correlation exists among the neighboring directional subbands is also adopted to reduce the high redundancy of the overcomplete transform. Experimental results justify that the proposed algorithm is superior to the state-of-art wavelet-based coders for rich-edgy images.



(a) (b)

Fig. 4. Reconstructed Barbara images at 0.15 bpp using (a) JPEG2000, PSNR = 25.93 dB and (b) PDFB, PSNR = 26.74 dB.

6. REFERENCES

[1] S. Mallat, *A Wavelet tour of signal processing*, 2nd ed. Academic Press, 1999.

[2] E. J. Candès and D. L. Donoho, "Curvelets - a surprisingly effective nonadaptive representation for objects with edges," in *Curves and Surfaces*, L. L. S. et al., Ed. Nashville, TN: Vanderbilt University Press, 1999.

[3] M. N. Do and M. Vetterli, "The contourlet transform: An efficient directional multiresolution image representation," *IEEE Transactions on Image Processing*, p. in press, 2005.

[4] R. H. Bamberger and M. J. T. Smith, "A filter bank for the directional decomposition of images: theory and design," *IEEE Transactions on Signal Processing*, vol. 40, no. 7, pp. 882 – 893, Apr. 1992.

[5] T. T. Nguyen and S. Oraintara, "On the aliasing effect of the contourlet filter banks," in *Submitted to the International Symposium on Circuits and Systems (ISCAS'06)*, May 2006.

[6] S. D. Servetto, K. Ramchandran, and M. T. Orchard, "Image coding based on a morphological representation of wavelet data," *IEEE Transaction on Image Processing*, vol. 8, no. 9, pp. 1161–1174, Sept. 1999.

[7] L. Vincent, "Morphological grayscale reconstruction in image analysis: Applications and effective algorithms," *IEEE Transaction on Image Processing*, vol. 2, pp. 176–201, Apr. 1993.

[8] R. M. Haralick and L. G. Shapiro, *Computer and Robot Vision*. Reading, MA: Addison-Wesley, 1992.

[9] J. M. Zhong, C. H. Leung, and Y. Y. Tang, "Wavelet image coding based on significance extraction using morphological operations," in *Proc. Inst. Elect. Eng.– Vis. Image Signal Processing*, vol. 146, no. 4, 1999, pp. 206–210.

[10] X. Wu, "High-order context modeling and embedded conditional entropy coding of wavelet coefficients for image compression," in *Proc. 31st Asilomar Conf. Signals, Systems, Computers*, Nov. 1997, pp. 1378–1382.

[11] A. Said and W. A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, pp. 23–50, June 1996.

[12] T. Acharya and P. Tsai, *JPEG2000 Standard for Image Compression*. John Wiley and Sons, Inc., 2005.