Direction-Adaptive Discrete Wavelet Transform for Image Compression

Chuo-Ling Chang, *Student Member, IEEE*, and Bernd Girod, *Fellow, IEEE*

***Abstract*—We propose a direction-adaptive DWT (DA-DWT) that locally adapts the filtering directions to image content based on directional lifting. With the adaptive transform, energy compaction is improved for sharp image features. A mathematical analysis based on an anisotropic statistical image model is presented to quantify the theoretical gain achieved by adapting the filtering directions. The analysis indicates that the proposed DA-DWT is more effective than other lifting-based approaches. Experimental results report a gain of up to 2.5 dB in PSNR over the conventional DWT for typical test images. Subjectively, the reconstruction from the DA-DWT better represents the structure in the image and is visually more pleasing.**

***Index Terms*—Adaptive filters, anisotropic image model, directional lifting, image coding, wavelet transforms.**

I. Introduction

THE 2-D discrete wavelet transform (DWT) is the most important new image compression technique of the last decade [1]–[3]. Conventionally, the 2-D DWT is carried out as a separable transform by cascading two 1-D transforms in the vertical and horizontal direction. Therefore, vanishing moments of the high-pass wavelet filters exist only in these two directions. The separable transform fails to provide an efficient representation for directional image features, such as edges and lines, not aligned vertically or horizontally since it spreads the energy of these features across subbands.

An early work of an adaptive wavelet transform that adapts the transform directions to image content is presented in [4]. The image is partitioned into blocks. Each image block is then sheared through a reversible resampling filter such that the edges in the sheared block are oriented either vertically or horizontally. The conventional 2-D DWT is applied to the sheared block and, thus, provides vanishing moments along the edges. The more recent work of the *directionlets* achieves this ability by adapting both the wavelet filtering direction and the subsampling grid to the image feature orientation without resampling [5]. The two approaches share the same limitations. First, independent processing of image blocks fails to exploit the correlation across block boundaries and can produce blocking artifacts in the reconstruction. Second, support for spatial scalability is limited since the subsampled low-pass image no longer resides on a regular orthogonal grid.

The *bandelets* approach described in [6], [7] does not have these limitations. The conventional 2-D DWT is first applied to the image, followed by the *bandeletization* procedure that further removes the directional correlation in the high-pass wavelet coefficients. The low-pass image remains the same as in the conventional 2-D DWT, and blocking artifacts are not observed since the block-wise operations are performed in the wavelet domain. However, because the procedure is essentially post-processing of the wavelet coefficients, the inefficiency of the wavelet transform itself remains unresolved.

An important technique that enables a locally adaptive wavelet transform while allowing filtering across block boundaries as well as a regular subsampling grid is *lifting*, a procedure to design DWTs that are ensured to achieve perfect reconstruction [8], [9]. Using lifting, several approaches have been developed to locally adapt the filter coefficients [10]–[12], or the filtering directions [11], [13], [14], such that filtering is not performed across edges in the image. These approaches eliminate the need for signaling the filter selections to the decoder by assuming lossless compression [11] or knowledge of the quantization noise at the encoder [10], [12], [14], or constraining the selection process such that it can be reliably repeated at the decoder [13]. The gain of adaptation is limited due to these assumptions and constraints. No significant improvement on objective quality measurements over the conventional 2-D DWT has been reported, although subjective improvement has been observed.

Other approaches that also adaptively select the filtering directions via lifting choose to explicitly signal the selections to the decoder [15], [16]. In this category, we have independently developed an approach that combines directional lifting with quincunx subsampling [17]. It is then extended to accommodate the case of conventional orthogonal subsampling, and to incorporate directional lifting and the *bandeletization* procedure into a unified framework [18]. Thanks to the efficient representation of the filter selections, these approaches adapt to the image features more effectively and have demonstrated significant objective and subjective quality improvement for texture-rich images.

In this paper, we propose to use a direction-adaptive DWT (DA-DWT) based on directional lifting. The proposed DA-DWT differs from other lifting-based approaches in that the directional filters in the DA-DWT provide a more efficient representation for sharp features in the image. In addition, it requires less computation and the compression performance is less sensitive to image transposition (swapping the x and y direction of the image). Mathematical analysis of the DA-DWT

Manuscript received September 28, 2006; revised September 14, 2006. This work was supported by the Max Planck Center for Visual Computing and Communication. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Giovanni Poggi.

The authors are with the Information Systems Laboratory, Stanford, CA 94305 USA (e-mail: chuoling@stanford.edu; bgirod@stanford.edu).

Digital Object Identifier 10.1109/TIP.2007.894242

is also presented in the paper, together with an anisotropic statistical model of images, to quantify the theoretical gain achieved by adapting the filtering directions. The analysis is also applicable to the lifting-based approaches in [13], [15], and [16]; hence, it provides a direct comparison with our DA-DWT.

In the remainder of this paper, we describe various components of the proposed DA-DWT in Section II. Mathematical analysis of the DA-DWT is presented in Section III. Finally, experimental results are reported in Section IV, demonstrating the objective and subjective improvements over the conventional transform.

II. Direction-Adaptive Discrete Wavelet Transform (DA-DWT)

In this section, we first present a general formulation of the 2-D DWT implemented with lifting. The proposed DA-DWT using directional lifting is then described with this formulation. Finally, coding methods of the overhead for signaling the direction selection are described, followed by discussions about the computational complexity of the DA-DWT.

A. Two-Dimensional Discrete Wavelet Transform With Lifting

The 2-D DWT in general consists of two stages. In the first stage, a 1-D DWT is applied to the image followed by vertical subsampling to obtain the low-pass subband, L, and the high-pass subband, H. In the second stage, another 1-D DWT is applied to L and H, followed by horizontal subsampling to obtain the LL and LH, and the HL and HH subband, respectively.

It is shown in [8] that any two-band biorthogonal DWT can be factored into pairs of lifting steps. In this paper, we only consider the wavelet transforms that can be realized with one pair of the lifting steps, i.e., one prediction step followed by one update step. For instance, the Haar wavelet, the 5/3 wavelet, and the family of interpolating wavelets all belong to this category [9], [19].

Let $\mathbf{s} = \{s[\mathbf{l}] | \mathbf{l} \in \mathbf{\Pi}\}$, where $s[\mathbf{l}] = s[l_x, l_y]$ and $\mathbf{l} = (l_x, l_y)^T$, denote a set of image samples on a 2-D orthogonal sampling grid $\mathbf{\Pi} = \{(l_x, l_y)^T \in \boldsymbol{\mathcal{Z}}^2\}$. The grid $\mathbf{\Pi}$ is composed of four subgrids $\mathbf{\Pi}_{pq} = \{(l_x, l_y)^T \in \mathbf{\Pi} | l_y \bmod 2 = p, l_x \bmod 2 = q\}$.

To apply the 2-D DWT with lifting, we first apply a transform between the even and the odd rows of the image, i.e., between $\mathbf{s}_0 = \{s[\mathbf{l}_0] | \mathbf{l}_0 \in \mathbf{\Pi}_0 = \mathbf{\Pi}_{00} \cup \mathbf{\Pi}_{01}\}$ and $\mathbf{s}_1 = \{s[\mathbf{l}_1] | \mathbf{l}_1 \in \mathbf{\Pi}_1 = \mathbf{\Pi}_{10} \cup \mathbf{\Pi}_{11}\}$. Denote the resulting low-pass subband by $\mathbf{w}_0 = \{w_0[\mathbf{l}_0] | \mathbf{l}_0 \in \mathbf{\Pi}_0\}$ and the high-pass subband by $\mathbf{w}_1 = \{w_1[\mathbf{l}_1] | \mathbf{l}_1 \in \mathbf{\Pi}_1\}$, the transform on $\mathbf{s}$ can generally be expressed as

$$w_1[\mathbf{l}_1] = g_H \cdot \left(s[\mathbf{l}_1] - P_{\mathbf{l}_1}(\mathbf{s}_0)\right), \quad \forall \mathbf{l}_1 \in \mathbf{\Pi}_1 \tag{1a}$$

$$w_0[\mathbf{l}_0] = g_L \cdot \left(s[\mathbf{l}_0] + g_H^{-1} \cdot U_{\mathbf{l}_0}(\mathbf{w}_1)\right), \quad \forall \mathbf{l}_0 \in \mathbf{\Pi}_0 \tag{1b}$$

where the prediction function, $P_{\mathbf{l}_1}(\cdot)$, and the update function, $U_{\mathbf{l}_0}(\cdot)$, are functions of the sample values in the input with a scalar output, and g_L and g_H are scaling factors.

Similarly, the transform between the even and the odd columns of the samples further decomposes $\mathbf{w}_0$ into $\mathbf{w}_{00}$ defined on $\mathbf{\Pi}_{00}$ and $\mathbf{w}_{01}$ defined on $\mathbf{\Pi}_{01}$, $\mathbf{w}_1$ into $\mathbf{w}_{10}$ defined on $\mathbf{\Pi}_{10}$ and $\mathbf{w}_{11}$ defined on $\mathbf{\Pi}_{11}$, where $\mathbf{w}_{00}$, $\mathbf{w}_{01}$, $\mathbf{w}_{10}$, and $\mathbf{w}_{10}$ correspond to the LL, LH, HL, and HH subband of the image, respectively. We refer to this process of decomposing the image samples $\mathbf{s}$ into the four subbands as one level of the 2-D DWT. Multiple levels of the transform may be performed by iteratively applying this process to the resulting LL subband.

B. Directional Lifting

In the conventional 2-D DWT, the prediction and the update function in (1) can be expressed as

$$P_{\mathbf{l}_1}(\mathbf{s}_0) = \sum_{k=-K_P}^{K_P-1} c_{P,k} \cdot s\left[l_{1,x}, l_{1,y} - (2k+1)\right] \tag{2a}$$

$$U_{\mathbf{l}_0}(\mathbf{w}_1) = \sum_{k=-K_U}^{K_U-1} c_{U,k} \cdot w_1\left[l_{0,x}, l_{0,y} - (2k+1)\right] \tag{2b}$$

where K_P, $c_{P,k}$, K_U, and $c_{U,k}$ are determined by the wavelet kernel adopted. Note that only samples in the same column are involved in the transform.

For image compression, for each sample in the high-pass subband, $w_1[\mathbf{l}_1]$, it is generally desirable to select a prediction function $P_{\mathbf{l}_1}(\mathbf{s}_0)$ in (1a) that predicts $s[\mathbf{l}_1]$ from the samples in $\mathbf{s}_0$ such that the magnitude of the residual $w_1[\mathbf{l}_1]$ is minimized. In the proposed DA-DWT, we define directional prediction filters with direction $\mathbf{d} = (d_x, d_y)^T$ from which $P_{\mathbf{l}_1}(\mathbf{s}_0)$ can be adaptively selected by

$$P_{\mathbf{l}_1}^{\mathbf{d}}(\mathbf{s}_0) = \sum_{k=-K_P}^{K_P-1} c_{P,k} \cdot s\left[\mathbf{l}_1 - (2k+1)\mathbf{d}\right] \tag{3}$$

where $\mathbf{d}$ is defined such that

$$\mathbf{l}_1 - (2k+1)\mathbf{d} \in \mathbf{\Pi}_0, \quad \forall \mathbf{l}_1 \in \mathbf{\Pi}_1, k = -K_P, \cdots, K_P - 1. \tag{4}$$

The directional prediction filter corresponds to performing the prediction step along direction $\mathbf{d}$. From (4), $\mathbf{d} \in \boldsymbol{\mathcal{Z}}^2$ and d_y is always odd. We further constrain $\mathbf{d}$ such that the line segment from $(0,0)^T$ to $(d_x, d_y)^T$ does not intersect with any other point in $\mathbf{\Pi}$, i.e., d_x and d_y are coprime integers. Typically, we use nine directions: $\mathbf{d}_{-4} = (-3,1)^T$, $\mathbf{d}_{-3} = (-2,1)^T$, $\mathbf{d}_{-2} = (-1,1)^T$, $\mathbf{d}_{-1} = (-1,3)^T$, $\mathbf{d}_0 = (0,1)^T$, $\mathbf{d}_1 = (1,3)^T$, $\mathbf{d}_2 = (1,1)^T$, $\mathbf{d}_3 = (2,1)^T$, and $\mathbf{d}_4 = (3,1)^T$, as illustrated in Fig. 1.

Denote the direction selected at location $\mathbf{l}_1$ for $P_{\mathbf{l}_1}(\mathbf{s}_0)$ as $\mathbf{d}_{\mathbf{l}_1}^{\star}$. Upon completion of the prediction step of all samples, the corresponding update function in (1b) is defined as

$$U_{\mathbf{l}_0}(\mathbf{w}_1) = \sum_{k=-K_U}^{K_U-1} c_{U,k} \cdot \sum_{\left\{\mathbf{l}_1 | \mathbf{l}_1 - (2k+1)\mathbf{d}_{\mathbf{l}_1}^{\star} = \mathbf{l}_0\right\}} w_1[\mathbf{l}_1]. \tag{5}$$

In other words, wherever an image sample $s[\mathbf{l}_1]$ is predicted by $c_{P,k} \cdot s[\mathbf{l}_0]$, $s[\mathbf{l}_0]$ is updated by $c_{U,k} \cdot w_1[\mathbf{l}_1]$. Note that if $\mathbf{d}_0$ is always selected, i.e., $\mathbf{d}_{\mathbf{l}_1}^{\star} = \mathbf{d}_0, \forall \mathbf{l}_1$, (3) and (5) are equivalent to the functions in (2) and directional lifting is identical to the conventional DWT.

For the second stage of the DWT further applied to $\mathbf{w}_0$ and $\mathbf{w}_1$, the transform can be applied along directions derived from the set of $\mathbf{d}$s, defined by $\breve{\mathbf{d}} = (\breve{d}_x, \breve{d}_y)^T = (d_y, 2d_x)^T$, i.e., $\breve{\mathbf{d}}_{-4} = (1,-6)^T$, $\breve{\mathbf{d}}_{-3} = (1,-4)^T$, $\breve{\mathbf{d}}_{-2} = (1,-2)^T$, $\breve{\mathbf{d}}_{-1} =$

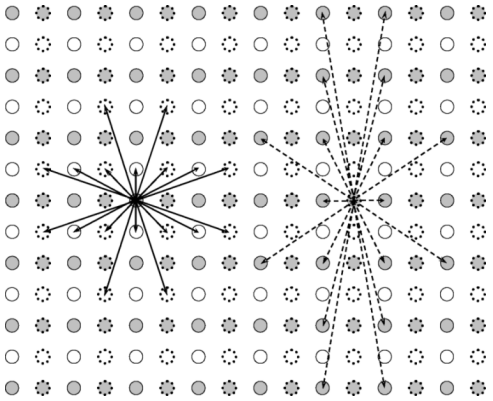


Fig. 1. Proposed set of directions for directional lifting in DA-DWT. The solid arrows denote $\mathbf{d}$ and $-\mathbf{d}$, and the dashed arrows denote $\check{\mathbf{d}}$ and $-\check{\mathbf{d}}$. White-solid dots denote $\mathbf{\Pi}_{00}$, white-dashed dots denote $\mathbf{\Pi}_{01}$, gray-solid dots denote $\mathbf{\Pi}_{10}$, and gray-dashed dots denote $\mathbf{\Pi}_{11}$.

$(3,-2)^T$, $\check{\mathbf{d}}_0 = (1,0)^T$, $\check{\mathbf{d}}_1 = (3,2)^T$, $\check{\mathbf{d}}_2 = (1,2)^T$, $\check{\mathbf{d}}_3 = (1,4)^T$, and $\check{\mathbf{d}}_4 = (1,6)^T$, as illustrated in Fig. 1. Consequently, $\check{d}_y$ is even and $\check{d}_x$ is odd, and

$$\begin{aligned} \mathbf{l}_{01} - (2k+1)\check{\mathbf{d}} \in \overline{\mathbf{\Pi}}_{00}, \quad \mathbf{l}_{11} - (2k+1)\check{\mathbf{d}} \in \overline{\mathbf{\Pi}}_{10} \\ \forall \mathbf{l}_{01} \in \mathbf{\Pi}_{01}, \quad \mathbf{l}_{11} \in \mathbf{\Pi}_{11} \\ k = -K_P, \cdots, K_P - 1. \end{aligned} \tag{6}$$

The direction selected for the transform on $\mathbf{w}_0$ and that for $\mathbf{w}_1$ may be different. The candidates for the prediction function and the corresponding update function are defined similarly as in (3) and (5). Note that since $\check{\mathbf{d}} = (d_y, 2d_x)^T$, directional lifting on $\mathbf{w}_0$ and on $\mathbf{w}_1$ along $\check{\mathbf{d}}_i$ share the same implementation with directional lifting on $\mathbf{s}$ along $\mathbf{d}_i$, if $\mathbf{w}_0$ and $\mathbf{w}_1$ are first vertically subsampled by a factor of 2 and transposed.

We propose $K_P = 3$ with $c_{P,0} = c_{P,-1} = 150/256$, $c_{P,1} = c_{P,-2} = -25/256$ and $c_{P,2} = c_{P,-3} = 3/256$, and $K_U = 3$ with $c_{U,0} = c_{U,-1} = 150/512$, $c_{U,1} = c_{U,-2} = -25/512$ and $c_{U,2} = c_{U,-3} = 3/512$, i.e., the (6,6) interpolating wavelet [19], as they result in the best compression performance in our experiments. Correspondingly, we choose $g_L = \sqrt{1.706}$ and $g_H = \sqrt{0.626}$, computed from the synthesis wavelet filters such that for one level of the transform, if the quantization noise of the wavelet coefficients are uncorrelated, the noise energy is preserved in the reconstructed image.

Several other approaches in the literature that also achieve directional adaptation through lifting essentially adopt a set of sub-pel directions, described by $\mathbf{d}_i = (i/K, 1)^T$, $i = -K, \cdots, K$, where K is typically 2 or 4, and the sub-pel directions are achieved by spatial interpolation [13], [15], [16]. These approaches assume $d_y = 1$ whereas the proposed DA-DWT framework only requires d_y is odd. For instance, for an image feature oriented along 22.5°, our approach would select $(1,3)^T$ (18.4°) whereas the others would select $(1/2,1)^T$ (26.6°). Despite that with $d_y > 1$ in our approach the reference samples are further away from the samples being predicted, experiments indicate that using these samples for prediction is typically more efficient than the interpolated ones in the presence of sharp image features, especially when simple interpolating filters are used. The mathematical analysis that will be presented later in Section III also indicates that the set of directions proposed in this section is preferable to the sub-pel directions, even when an ideal interpolating filter is adopted. Additionally, since our approach does not involve sub-pel interpolation, it also requires less computation.

Furthermore, in the proposed sub-pel approaches only the directions between $\pm 45°$ are considered for $\mathbf{d}$. For horizontal-ish image features, i.e., features oriented beyond 45° from the vertical axis, there is no filtering direction closely aligned with the image feature, and, thus, the energy is spread into the high-pass subband H rather than contained in L. As a result, for image compression they favor vertical-ish image features and, therefore, are sensitive to image transposition [17]. In our previous work [17], quincunx subsampling was proposed to provide directions covering the 2-D space more evenly and has shown reduced sensitivity to image transposition. However, to compress typical images captured by image sensors with a rectangular aperture function, DWT using quincunx subsampling is less efficient [20]. In this work, we adopt the conventional orthogonal subsampling as described in Section II-A but allow directions beyond $\pm 45°$, e.g., $\mathbf{d}_{\pm 3}$ and $\mathbf{d}_{\pm 4}$, to provide a larger span of directions. Similar to the case with quincunx subsampling, reduced sensitivity to transposition is observed. The mathematical analysis in Section III also agrees with this observation.

C. Adaptive Selection of Direction and Block-Partition

For image compression, the filtering directions in the DA-DWT should be selected to minimize the distortion of the reconstructed image for a given rate budget. The rate budget is spent on the overhead required to signal the selection and the rate for coding the wavelet coefficients. Finding the best direction is, therefore, analogous to rate-constrained motion estimation in video coding [21], [22].

To reduce the amount of overhead, we perform the selection in a block-wise fashion. For the transform applied to $\mathbf{s}$, the original grid $\mathbf{\Pi}$ is evenly partitioned into N_b nonoverlapping blocks, each denoted by $\mathbf{B}_b$, $b = 0, \cdots, N_b - 1$. For each block $\mathbf{B}_b$, e.g., a block of 64×64 pixels in our experiments, direction $\mathbf{d}_b^\star$ is selected by minimizing a Lagrangian cost function

$$\mathbf{d}_b^\star = \arg\min_{\mathbf{d}} \left\{ \sum_{\mathbf{l}_1 \in \mathbf{\Pi}_1 \cap \mathbf{B}_b} D\big(g_H \cdot \big(s[\mathbf{l}_1] - P_{\mathbf{l}_1}^{\mathbf{d}}(\mathbf{s}_0)\big)\big) + \lambda R_b^{\mathbf{d}} \right\} \tag{7}$$

where $D(\cdot)$ is a distortion measurement, $\lambda > 0$ is a Lagrangian multiplier, and $R_b^{\mathbf{d}}$ denotes the number of bits spent on signaling the selection $\mathbf{d}_b^\star = \mathbf{d}$. Note that $D(\cdot)$ should reflect the distortion in the reconstructed image. However, to obtain an embedded image representation, perhaps the most important motivation to use wavelet image coding, the decoding rate—hence, the distortion—is not known *a priori* at the encoder [3]. As an approximation, $D(\cdot) = (\cdot)^2$ may be used as the distortion measurement, assuming most of the high-pass coefficients are quantized to zero and the distortion in the wavelet domain is the same as the distortion in the image domain (with proper normalization). Additionally, it has been shown that $D(\cdot) = |\cdot|$ and $D(\cdot) = (\cdot)^2$ result in similar performance for video coding [22]. For the proposed DA-DWT, the former is adopted in the current implementation.

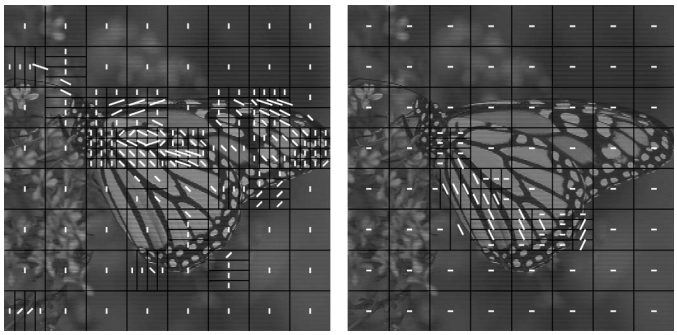


Fig. 2. Selected directions and block-partition for the first level of the transform applied to the 512 × 512 image *Monarch*. The directions shown on the left correspond to $\mathbf{d}$, and those on the right correspond to $\check{\mathbf{d}}$.

Given the $\mathbf{d}_b^\star$ for all blocks, it follows that $\mathbf{d}_{\mathbf{l}_1}^\star = \mathbf{d}_b^\star, \forall \mathbf{l}_1 \in \mathbf{B}_b$, and, thus, the prediction and the update function for every sample can be defined according to (3) and (5). The cost functions for the transform on $\mathbf{w}_0$ and $\mathbf{w}_1$ are similarly defined. Note that although the direction is selected block-wise, filtering in the prediction and update step extends across block boundaries. Therefore, blocking artifacts are not observed in the reconstructed images.

To further increase the efficiency of the prediction step, each block $\mathbf{B}_b$ may be further partitioned into 2 × 1, 1 × 2, 2 × 2, 4 × 1, 1 × 4, 4 × 2, 2 × 4, or 4 × 4 subblocks. The cost function in (7) is modified to incorporate the overhead to signal the block-partition. The best block-partition for each block and the best direction for each subblock are then selected using the modified cost function, similar to the variable block-size motion compensation in video coding [23]. An example of the selected directions and block partition are shown in Fig. 2.

At block (subblock) boundaries, an even sample at $\mathbf{l}_0$ may be used to predict the odd samples in more than one directions. This happens when, for example, $n(\{\mathbf{l}_1 | \mathbf{l}_1 + \mathbf{d}_{\mathbf{l}_1}^\star = \mathbf{l}_0\}) > 1$, where $n(\mathcal{A})$ denotes the cardinality of set $\mathcal{A}$. This is analogous to the situation of multiply-connected pixels in video coding with motion-compensated temporal filtering [24]. Meticulous rules have been developed to cope with this situation in order to make the adaptive transform reversible, both for video coding [24] and image coding [25], [26]. Owing to the lifting structure, the proposed approach does not require these rules and reversibility of the transform is always ensured. The situation of multiply-connected pixels also incurs ambiguity in the direction of the update step. The update functions we propose in (5), analogous to the barbell lifting scheme for video coding [27], has been shown to achieve better compression performance than the update functions in [15] and [17], analogous to using the reversed motion vectors in the update step [27].

D. Coding of Direction and Block-Partition Selection

The direction selection is predicted from the selections of the blocks (subblocks) in the causal neighborhood, and the residual is coded with variable-length coding, again, similar to motion-compensated video coding [28]. The residual is computed with modular N_d arithmetic, where N_d denotes the number of direction candidates, in accordance with the cyclic property of the proposed directions. Each block-partition selection is independently encoded with variable-length coding.

Note that in the scope of this paper, we only consider lossless coding of the direction and block-partition selection. In [29], we have also developed an approach to allow lossy coding of the selections in order to provide a scalable direction representation, analogous to the scalable representation of motion vectors in video coding [30], [31].

E. Computational Complexity of DA-DWT

From the above discussion, the DA-DWT consists of a 1-D DWT along a certain direction, followed by additional 1-D DWTs on the resulting low-pass and high-pass coefficients, possibly along different directions. The 1-D DWT is realized with directional lifting, which requires the same amount of computation regardless of the selected direction.

At the encoder, to obtain a high-pass coefficient with a particular direction, the prediction step, as expressed in (1a) and (3), requires K_P multiplications and $2K_P$ additions. This is also the computation required in the conventional DWT. Note that scaling by g_H is neglected since it can be equivalently achieved by adjusting the quantization level when encoding the coefficients. For the DA-DWT with N_d direction candidates, the prediction step, together with the computation for the sum of absolute values, needs to be performed N_d times. Therefore, the required computation includes $N_d \cdot K_P$ multiplications and $N_d \cdot (2K_P + 1)$ additions. Note that the Lagrangian cost function is minimized at the block level; hence, the computation required can be neglected. To obtain a low-pass coefficient, the direction in the update step is the same as in the prediction step. As a result, the computation required, for both the conventional DWT and the DA-DWT, includes K_U multiplications and $2K_U$ additions.

Using the (6,6) interpolating wavelet and the set of directions described in Section II-B, $K_P = K_U = 3$ and $N_d = 9$. Hence, at the encoder the conventional DWT requires 6 multiplications and 12 additions for every pair of the lifting steps, whereas the DA-DWT requires 30 multiplications and 69 additions. Therefore, at the encoder the DA-DWT requires about 5 to 6 times more computation than the conventional 2-D DWT. At the decoder, both require six multiplications and 12 additions since the direction is signaled from the encoder.

III. Mathematical Analysis of DA-DWT

In this section, we provide mathematical analysis of the performance gain resulting from adapting the wavelet filtering directions. We first derive the power spectral density (PSD) of the high-pass and the low-pass wavelet coefficients as a function of the PSD of the image samples and the filtering direction. The direction-adaptation gain is then introduced to quantify the performance gain. The analysis is performed for the DWT with vertical subsampling. It is straightforward to extend the analysis to the DWT with horizontal subsampling as well as multiple levels of the DWT. In addition, we propose an anisotropic statistical model of images. Numerical evaluation using the image model is provided to demonstrate the advantages of the direction-adaptive transform over the conventional DWT.

A. Directional Filtering in Lifting Steps

As described in Section II, to apply the 2-D DWT to image samples $s[\mathbf{l}], \mathbf{l} = (l_x, l_y)^T \in \boldsymbol{\mathcal{Z}}^2$, using lifting, $s[\mathbf{l}]$ is first vertically subsampled into the even rows, $s_0[\mathbf{l}] = s[l_x, 2l_y]$, and the odd rows, $s_1[\mathbf{l}] = s[l_x, 2l_y + 1]$. Similar to (1), the prediction and the update step in lifting can be expressed as

$$w_1[\mathbf{l}] = g_H \cdot \left(s_1[\mathbf{l}] - h_P^{\mathbf{d}}[\mathbf{l}] * s_0[\mathbf{l}]\right) \tag{8a}$$

$$w_0[\mathbf{l}] = g_L \cdot \left(s_0[\mathbf{l}] + g_H^{-1} \cdot \left(h_U^{\mathbf{d}}[\mathbf{l}] * w_1[\mathbf{l}]\right)\right) \tag{8b}$$

where $w_0[\mathbf{l}]$ and $w_1[\mathbf{l}]$ denote the high-pass and the low-pass wavelet coefficients, respectively, $h_P^{\mathbf{d}}[\mathbf{l}]$ and $h_U^{\mathbf{d}}[\mathbf{l}]$ denote the 2-D impulse response of the directional filter in the lifting steps along direction $\mathbf{d} = (d_x, d_y)^T$, the symbol $*$ denotes the 2-D convolution operation, and g_L and g_H are scaling factors. In the following discussion, functions denoted by an upper-case letter in the form of $A(e^{j\boldsymbol{\Omega}}) = A(e^{j\Omega_x}, e^{j\Omega_y})$ represent the 2-D discrete-space Fourier transform of the corresponding discrete-space signal $a[\mathbf{l}] = a[l_x, l_y]$.

To facilitate the analysis, we define $\hat{s_0}[\mathbf{l}]$ as the upsampled $s_0[\mathbf{l}]$, i.e.,

$$\hat{s_0}[\mathbf{l}] = \hat{s_0}[l_x, l_y] = \begin{cases} s_0[l_x, l_y/2], & l_y \text{ is even} \\ 0, & l_y \text{ is odd.} \end{cases} \tag{9}$$

Following the discussion in Section II-B, we further define the 2-D directional prediction filters applied to $\hat{s_0}[\mathbf{l}]$ by the impulse response $\hat{h}_P^{\mathbf{d}}[\mathbf{l}]$

$$\hat{h}_P^{\mathbf{d}}[\mathbf{l}] = \begin{cases} c_{P,k}, & \mathbf{l} = (2k+1)\mathbf{d},\ k = -K_P, \cdots, K_P - 1 \\ 0, & \text{otherwise} \end{cases} \tag{10}$$

where K_P and $c_{P,k}$ are determined by the wavelet kernel adopted as discussed in Section II-B. In the conventional DWT, $\hat{h}_P^{\mathbf{d}_0}[\mathbf{l}]$, $\mathbf{d}_0 = (0,1)^T$, is always used, corresponding to vertical filtering with the prototype 1-D prediction filter, $\hat{h}_P[t]$, defined as

$$\hat{h}_P[t] = \hat{h}_P^{\mathbf{d}_0}[0, t], \quad t \in \boldsymbol{\mathcal{Z}}. \tag{11}$$

Denote the 1-D discrete-space Fourier transform of $\hat{h}_P[t]$ by $\hat{H}_P(e^{j\Omega_t})$. We continue the analysis by relating $\hat{H}_P^{\mathbf{d}}(e^{j\boldsymbol{\Omega}})$ with $\hat{H}_P(e^{j\Omega_t})$. For the directions defined in Section II, observe that the line segment from $(0,0)^T$ to $\mathbf{d} = (d_x, d_y)^T$ does not intersect with any other point in the $\boldsymbol{\mathcal{Z}}^2$ grid, d_x and d_y are coprime, i.e., their greatest common divisor is 1. From the Bézout's Identity [32], there is at least one direction, $\bar{\mathbf{d}} = (\bar{d_x}, \bar{d_y})^T \in \boldsymbol{\mathcal{Z}}^2$, such that

$$\bar{d_x} d_y - \bar{d_y} d_x = 1. \tag{12}$$

From basic lattice theory, a generating matrix defined as

$$\mathbf{W} = \begin{pmatrix} \bar{d_x} & d_x \\ \bar{d_y} & d_y \end{pmatrix} \tag{13}$$

is a one-to-one mapping from $\boldsymbol{\mathcal{Z}}^2$ to $\boldsymbol{\mathcal{Z}}^2$, i.e.,

$$\{\mathbf{l}' | \mathbf{l}' = \mathbf{W}\mathbf{l}, \forall \mathbf{l} \in \boldsymbol{\mathcal{Z}}^2\} = \boldsymbol{\mathcal{Z}}^2. \tag{14}$$

With $\mathbf{W}$, $\hat{h}_P^{\mathbf{d}}[\mathbf{l}]$ can be related to $\hat{h}_P^{\mathbf{d}_0}[\mathbf{l}]$ by

$$\hat{h}_P^{\mathbf{d}}[\mathbf{W}\mathbf{l}] = \hat{h}_P^{\mathbf{d}_0}[\mathbf{l}], \quad \hat{h}_P^{\mathbf{d}}[\mathbf{l}] = \hat{h}_P^{\mathbf{d}_0}[\mathbf{W}^{-1}\mathbf{l}]. \tag{15}$$

From (12) and (13), and the fact that d_x and d_y are coprime, one can verify that (15) is consistent with the definition in (10).

From (10) and (11), $\hat{h}_P^{\mathbf{d}_0}[\mathbf{l}] = \delta[l_x] \cdot \hat{h}_P[l_y]$, and it is straightforward that

$$\hat{H}_P^{\mathbf{d}_0}(e^{j\boldsymbol{\Omega}}) = \hat{H}_P(e^{j\Omega_y}). \tag{16}$$

Moreover, from (14)–(16) 16tip05-chang

$$\begin{aligned} \hat{H}_P^{\mathbf{d}}(e^{j\boldsymbol{\Omega}}) &= \sum_{\mathbf{l}' \in \boldsymbol{\mathcal{Z}}^2} \hat{h}_P^{\mathbf{d}}[\mathbf{l}'] e^{-j\boldsymbol{\Omega}^T \mathbf{l}'} = \sum_{\mathbf{l}' \in \boldsymbol{\mathcal{Z}}^2} \hat{h}_P^{\mathbf{d}_0}[\mathbf{W}^{-1}\mathbf{l}'] e^{-j\boldsymbol{\Omega}^T \mathbf{l}'} \\ &= \sum_{\mathbf{l} = \mathbf{W}^{-1}\mathbf{l}' \in \boldsymbol{\mathcal{Z}}^2} \hat{h}_P^{\mathbf{d}_0}[\mathbf{l}] e^{-j\boldsymbol{\Omega}^T \mathbf{W}\mathbf{l}} = \hat{H}_P^{\mathbf{d}_0}\left(e^{j\mathbf{W}^T\boldsymbol{\Omega}}\right) \\ &= \hat{H}_P\left(e^{j\mathbf{d}^T\boldsymbol{\Omega}}\right). \end{aligned} \tag{17}$$

In the Appendix, we relate the transfer function of the filter in (8a), $H_P^{\mathbf{d}}(e^{j\boldsymbol{\Omega}})$, with the equivalent filter in the upsampled domain in (10), $\hat{H}_P^{\mathbf{d}}(e^{j\boldsymbol{\Omega}})$, by

$$H_P^{\mathbf{d}}(e^{j\boldsymbol{\Omega}}) = e^{j\frac{\Omega_y}{2}} \hat{H}_P^{\mathbf{d}}\left(e^{j\Omega_x}, e^{j\frac{\Omega_y}{2}}\right). \tag{18}$$

Finally, combining (17) and (18), we obtain

$$H_P^{\mathbf{d}}(e^{j\boldsymbol{\Omega}}) = e^{j\frac{\Omega_y}{2}} \hat{H}_P\left(e^{j\left(d_x\Omega_x + d_y\frac{\Omega_y}{2}\right)}\right). \tag{19}$$

With a similar derivation, the transfer function of the filter in (8b), $H_U^{\mathbf{d}}(e^{j\boldsymbol{\Omega}})$, can be obtained as

$$H_U^{\mathbf{d}}(e^{j\boldsymbol{\Omega}}) = e^{-j\frac{\Omega_y}{2}} \hat{H}_U\left(e^{j\left(d_x\Omega_x + d_y\frac{\Omega_y}{2}\right)}\right) \tag{20}$$

where $\hat{H}_U(e^{j\Omega_t})$ is the transfer function of the prototype 1-D update filter.

B. Extension to Sub-Pel Directions

The above analysis is applicable to the case that d_x and d_y are coprime integers and d_y is odd, as proposed in Section II-B. In this subsection, we generalize the analysis to the case that Kd_x and d_y are coprime integers, where $K \in \boldsymbol{\mathcal{Z}}^+$, and d_y is odd. The approaches that incorporate sub-pel directions, e.g., $(1/4, 1)^T$, $K = 4$, belongs to this latter case.

To apply filtering along such directions, the image $s[\mathbf{l}]$ is first horizontally upsampled by a factor of K and interpolated horizontally with a 1-D interpolating filter $H_I^K(e^{j\Omega_t})$. Denote the equivalent direction in the upsampled image by $\mathbf{d}^K = (d_x^K, d_y^K)^T = (Kd_x, d_y)^T$. The upsampled and interpolated image is then filtered with $h_P^{\mathbf{d}^K}[\mathbf{l}]$ and finally downsampled horizontally by K. Since Kd_x and d_y are coprime integers,

(19) can be applied to derive the overall transfer function

$$\begin{aligned} H_P^{K,\mathbf{d}}(e^{j\mathbf{\Omega}}) &= \frac{1}{K}\sum_{k=0}^{K-1} H_P^{\mathbf{d}^K}\left(e^{j(\Omega_x+2k\pi)/K}, e^{j\Omega_y}\right) \\ &\quad \times H_I^K\left(e^{j(\Omega_x+2k\pi)/K}\right) \\ &= \frac{1}{K}\sum_{k=0}^{K-1} H_P^{\mathbf{d}}\left(e^{j(\Omega_x+2k\pi)}, e^{j\Omega_y}\right) \\ &\quad \times H_I^K\left(e^{j(\Omega_x+2k\pi)/K}\right) \end{aligned} \tag{21}$$

where $H_P^{\mathbf{d}}(e^{j\mathbf{\Omega}})$ is expressed as in (19) with d_x being a rational number rather than limited to integers.

Note that if the interpolating filter is an ideal rectangular filter, i.e., $H_I^K(e^{j\Omega_t}) = \sqcap_K(e^{j\Omega_t})$ where

$$\sqcap_K(e^{j\Omega_t}) = \begin{cases} K, & \exists k \in \boldsymbol{\mathcal{Z}} : -\frac{\pi}{K} \le \Omega_t + 2k\pi < \frac{\pi}{K} \\ 0, & \text{otherwise} \end{cases} \tag{22}$$

the following equality holds:

$$H_P^{K,\mathbf{d}}(e^{j\mathbf{\Omega}}) = H_P^{\mathbf{d}^K}\left(e^{j\frac{\Omega_x}{K}}, e^{j\Omega_y}\right) = H_P^{\mathbf{d}}(e^{j\mathbf{\Omega}}). \tag{23}$$

C. Power Spectral Density of Wavelet Coefficients

Assume $s[\mathbf{l}]$ is a stationary 2-D discrete-space random field. Denote the 2-D discrete-space autocorrelation function of $s[\mathbf{l}]$ by $R_{ss}[\Delta\mathbf{l}] = E\{s[\mathbf{l}+\Delta\mathbf{l}]s[\mathbf{l}]\}$. The 2-D discrete-space PSD of $s[\mathbf{l}]$, i.e., the 2-D discrete-space Fourier transform of $R_{ss}[\Delta\mathbf{l}]$, is denoted by $\Phi_{ss}(e^{j\mathbf{\Omega}})$ (with the shorthand notation Φ_{ss}). Since $s_0[\mathbf{l}]$ and $s_1[\mathbf{l}]$ are vertical subsamples of $s[\mathbf{l}]$ with different phases

$$R_{s_0s_0}[\Delta\mathbf{l}] = R_{s_1s_1}[\Delta\mathbf{l}] = R_{ss}[\Delta l_x, 2\Delta l_y] \tag{24}$$

$$R_{s_0s_1}[\Delta\mathbf{l}] = R_{ss}[\Delta l_x, 2\Delta l_y - 1]. \tag{25}$$

Therefore

$$\begin{aligned} \Phi_{s_0s_0}(e^{j\mathbf{\Omega}}) &= \Phi_{s_1s_1}(e^{j\mathbf{\Omega}}) \\ &= \frac{1}{2}\sum_{k=0}^{1}\Phi_{ss}\left(e^{j\Omega_x}, e^{j\left(\frac{\Omega_y}{2}+k\pi\right)}\right) \end{aligned} \tag{26}$$

$$\begin{aligned} \Phi_{s_0s_1}(e^{j\mathbf{\Omega}}) &= \frac{1}{2}\sum_{k=0}^{1}(-1)^k e^{-j\frac{\Omega_y}{2}}\Phi_{ss} \\ &\quad \times \left(e^{j\Omega_x}, e^{j\left(\frac{\Omega_y}{2}+k\pi\right)}\right). \end{aligned} \tag{27}$$

From (8), (19), and (20), and the property that $\Phi_{ss}(e^{j\mathbf{\Omega}})$ is real, the PSD of the high-pass and the low-pass wavelet coefficients when wavelet filtering is applied along direction $\mathbf{d}$, denoted by $\Phi_{w_1w_1}^{\mathbf{d}}$ and $\Phi_{w_0w_0}^{\mathbf{d}}$, respectively, can be expressed as

$$\Phi_{w_1w_1}^{\mathbf{d}} = g_H^2 \cdot \left(\left(1+\left|H_P^{\mathbf{d}}\right|^2\right)\Phi_{s_0s_0} - 2H_P^{\mathbf{d}}\Phi_{s_0s_1}\right) \tag{28}$$

$$\begin{aligned} \Phi_{w_0w_0}^{\mathbf{d}} &= g_L^2 \cdot \left(\left(\left|1-H_U^{\mathbf{d}}H_P^{\mathbf{d}}\right|^2 + \left|H_U^{\mathbf{d}}\right|^2\right)\Phi_{s_0s_0}\right. \\ &\quad \left. +2\left(1-H_U^{\mathbf{d}}H_P^{\mathbf{d}}\right)H_U^{\mathbf{d}*}\Phi_{s_0s_1}\right) \end{aligned} \tag{29}$$

where $H_U^{\mathbf{d}*}$ denotes the complex conjugate of $H_U^{\mathbf{d}}$.

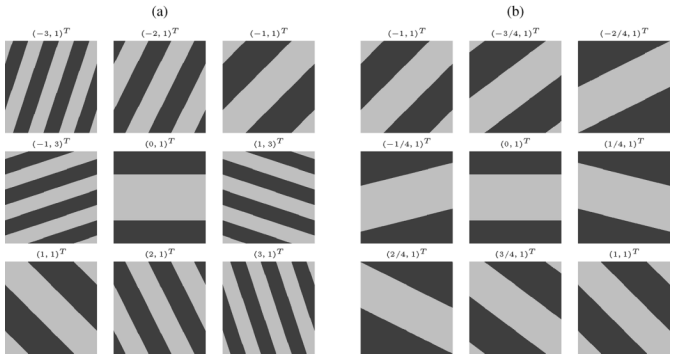


Fig. 3. $\hat{H}_P^{\mathbf{d}}(e^{j\mathbf{\Omega}})$ in (31) for: (a) the proposed directions; (b) the sub-pel directions. Each square corresponds to $-\pi \le \Omega_x \le \pi$ from left to right, and $-\pi \le \Omega_y \le \pi$ from bottom to top. Light gray and dark gray represent magnitude 1 and -1, respectively.

As an example, assume that the prototype 1-D prediction filter has impulse response

$$\hat{h}_P[t] = \begin{cases} 0, & t \text{ is even} \\ sinc\left(\frac{t}{2}\right), & t \text{ is odd} \end{cases} \tag{30}$$

and, hence, $\hat{H}_P(e^{j\Omega_t}) = \sqcap_2(e^{j\Omega_t}) - 1$. In addition, assume that the prototype 1-D update filter has impulse response $\hat{h}_U[t] = (1/2)\hat{h}_P[t]$. From (17)

$$\hat{H}_P^{\mathbf{d}}(e^{j\mathbf{\Omega}}) = \begin{cases} 1, & q^{\mathbf{d}}(\Omega_x,\Omega_y) \text{ is even} \\ -1, & q^{\mathbf{d}}(\Omega_x,\Omega_y) \text{ is odd} \end{cases} \tag{31}$$

$$\hat{H}_U^{\mathbf{d}}(e^{j\mathbf{\Omega}}) = \frac{1}{2}\hat{H}_P^{\mathbf{d}}(e^{j\mathbf{\Omega}}) \tag{32}$$

where

$$q^{\mathbf{d}}(\Omega_x,\Omega_y) = \left\lfloor \left(d_x\Omega_x + d_y\Omega_y + \frac{\pi}{2}\right)/\pi \right\rfloor \tag{33}$$

and $\lfloor\ \rfloor$ denotes the floor function. Plots of $\hat{H}_P^{\mathbf{d}}(e^{j\mathbf{\Omega}})$ for the directions described in Section II-B, as well as the sub-pel directions,are shown in Fig. 3.

Finally, combining (19), (20), (26)–(29), (31), and (32), the PSD of the wavelet coefficients in this example is

$$\begin{aligned} \Phi_{w_1w_1}^{\mathbf{d}}(e^{j\mathbf{\Omega}}) &= 2g_H^2 \\ &\quad \cdot \Phi_{ss}\left(e^{j\Omega_x}, e^{j\left(\frac{\Omega_y}{2}+\left(q^{\mathbf{d}}\left(\Omega_x,\frac{\Omega_y}{2}\right)+1\right)\pi\right)}\right) \end{aligned} \tag{34}$$

$$\Phi_{w_0w_0}^{\mathbf{d}}(e^{j\mathbf{\Omega}}) = \frac{1}{2}g_L^2 \cdot \Phi_{ss}\left(e^{j\Omega_x}, e^{j\left(\frac{\Omega_y}{2}+q^{\mathbf{d}}\left(\Omega_x,\frac{\Omega_y}{2}\right)\pi\right)}\right). \tag{35}$$

Note that, for this example, $g_L = g_H^{-1} = \sqrt{2}$ to achieve an orthonormal transform.

D. Direction-Adaptation Gain

To reduce the energy in the high-pass coefficients, the prototype prediction filter is typically designed such that $\hat{H}_P(e^{j\mathbf{0}}) = 1$. In consequence, from (8) and assuming $s[\mathbf{l}]$ is zero-mean, $E\{w_0[\mathbf{l}]\} = E\{w_1[\mathbf{l}]\} = 0$. Given $\Phi_{w_1w_1}^{\mathbf{d}}(e^{j\mathbf{\Omega}})$ and $\Phi_{w_0w_0}^{\mathbf{d}}(e^{j\mathbf{\Omega}})$, the variance of the high-pass and the low-pass wavelet coefficients, denoted by $\sigma_{w_1}^{2,\mathbf{d}}$ and $\sigma_{w_0}^{2,\mathbf{d}}$ respectively, can be calculated simply by averaging the corresponding PSD over $-\pi \le \Omega_x < \pi$ and $-\pi \le \Omega_y < \pi$.

Following the discussion in Section II-C, for the direction-adaptive transform the best filtering direction, denoted by $\mathbf{d}^\star$, is selected by minimizing the Lagrangian cost function in (7). For high-rate applications, the corresponding λ is small and (7) can be approximated by selecting the direction that results in the minimum $\sigma^{2,\mathbf{d}}_{w_1}$.

Assuming high-rate scalar quantization, memoryless coding of the wavelet coefficients and orthonormality of the transform, the MSE distortion is proportional to the geometric mean of the subband variances, i.e., $(\sigma^{2,\mathbf{d}}_{w_0} \cdot \sigma^{2,\mathbf{d}}_{w_1})^{1/2}$ [3]. To quantify the performance gain of adapting the wavelet filtering directions, we define the direction-adaptation gain, denoted by G_D, as the ratio between the distortion obtained with the nonadaptive transform and the adaptive transform at the same rate, i.e.,

$$G_D = \left(\frac{\sigma^{2,\mathbf{d}_0}_{w_0} \cdot \sigma^{2,\mathbf{d}_0}_{w_1}}{\sigma^{2,\mathbf{d}^\star}_{w_0} \cdot \sigma^{2,\mathbf{d}^\star}_{w_1}} \right)^{\frac{1}{2}}. \tag{36}$$

Although (36) is expressed for two-band orthonormal transforms, similar to the transform coding gain expression, it is straightforward to extend the direction-adaptation gain to multiband transforms as well as biorthogonal transforms in general [3].

E. Statistical Model of Images

Assume $s[\mathbf{l}]$ are 2-D samples of an image source. The image source may be modeled as a stationary, zero-mean 2-D continuous-space random field $\tilde{s}(\mathbf{p}) = \tilde{s}(p_x, p_y)$, $\mathbf{p} = (p_x, p_y)^T \in \boldsymbol{\mathcal{R}}^2$, with autocorrelation function $R_{\tilde{s}\tilde{s}}(\boldsymbol{\tau})$, where $\boldsymbol{\tau} \in \boldsymbol{\mathcal{R}}^2$ and $R_{\tilde{s}\tilde{s}}(\boldsymbol{\tau}) = E\{\tilde{s}(\mathbf{p}+\boldsymbol{\tau})\tilde{s}(\mathbf{p})\}$. Typically, autocorrelation functions, averaged over many images, resemble an isotropic function

$$R^I_{\tilde{s}\tilde{s}}(\boldsymbol{\tau}) = \sigma^2 \exp(-\omega_0 \sqrt{\boldsymbol{\tau}^T \boldsymbol{\tau}}) \tag{37}$$

where variance σ^2 and parameter $\omega_0 \geq 0$ can be measured for a set of images [33].

Locally, individual image blocks may be far from isotropic. For instance, if an image block contains edges with a certain orientation, the autocorrelation orthogonal to the edge orientation typically decays faster while, along the edge orientation, it decays slower. To accommodate these cases, we extend (37) and propose an anisotropic model

$$R^A_{\tilde{s}\tilde{s}}(\boldsymbol{\tau}) = \sigma^2 \exp(-\sqrt{\boldsymbol{\tau}^T \boldsymbol{\Gamma} \boldsymbol{\tau}}) \tag{38}$$

where

$$\boldsymbol{\Gamma} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \omega_a^2 & 0 \\ 0 & \omega_b^2 \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}^T \tag{39}$$

and $\omega_a \geq \omega_b \geq 0$. In (38), contours of equal autocorrelation form ellipses where the angle θ represents the orientation of the minor axis. The length of the minor axis and that of the major axis are proportional to ω_a^{-1} and ω_b^{-1}, respectively, and the autocorrelation with a one-unit shift along the minor axis is $\sigma^2 \exp(-\omega_a)$ whereas that along the major axis is $\sigma^2 \exp(-\omega_b)$. Using this model to represent the source of an image block, the minor axis is aligned to the direction with the dominant intensity transition. For instance, if edges appear in an image block, the minor axis is aligned orthogonal to the edge orientation, and ω_a increases with the sharpness of the edges as well as the density of edges in the block.

The PSD of the proposed anisotropic model, i.e., the 2-D continuous-space Fourier transform of $R^A_{\tilde{s}\tilde{s}}(\boldsymbol{\tau})$, is

$$\Phi^A_{\tilde{s}\tilde{s}}(\boldsymbol{\omega}) = \frac{2\pi\sigma^2}{\omega_a \omega_b} (1 + \boldsymbol{\omega}^T \overline{\boldsymbol{\Gamma}} \boldsymbol{\omega})^{-\frac{3}{2}} \tag{40}$$

where $\boldsymbol{\omega} = (\omega_x, \omega_y)^T \in \boldsymbol{\mathcal{R}}^2$ and

$$\overline{\boldsymbol{\Gamma}} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \omega_a^{-2} & 0 \\ 0 & \omega_b^{-2} \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}^T. \tag{41}$$

The contours of equal PSD also form ellipses with the major axis aligned to the direction with the dominant intensity transition. The length of the major axis and that of the minor axis are proportional to ω_a and ω_b, respectively. Note that, as a special case of (40), the PSD of the isotropic model is

$$\Phi^I_{\tilde{s}\tilde{s}}(\boldsymbol{\omega}) = \frac{2\pi\sigma^2}{\omega_0^2} \left(1 + \frac{\boldsymbol{\omega}^T \boldsymbol{\omega}}{\omega_0^2} \right)^{-\frac{3}{2}}. \tag{42}$$

To derive the 2-D discrete-space anisotropic PSD model, let f_{sx} and f_{sy} denote the sampling frequency along the x axis and the y axis, respectively; thus, $s[\mathbf{l}] = s[l_x, l_y] = \tilde{s}(l_x/f_{sx}, l_y/f_{sy})$. We further assume that $\tilde{s}(\mathbf{p})$ is band limited with $\Phi^A_{\tilde{s}\tilde{s}}(\boldsymbol{\omega})$ modified to

$$\Phi^A_{\tilde{s}\tilde{s}}(\boldsymbol{\omega}) = \begin{cases} \frac{2\pi\sigma^2}{\omega_a \omega_b} (1 + \boldsymbol{\omega}^T \overline{\boldsymbol{\Gamma}} \boldsymbol{\omega})^{-\frac{3}{2}}, & |\omega_x| \leq (2B+1)\pi f_{sx} \\ & |\omega_y| \leq (2B+1)\pi f_{sy} \\ 0, & \text{otherwise.} \end{cases} \tag{43}$$

As a result, the discrete-space anisotropic PSD model is

$$\Phi^A_{ss}(e^{j\boldsymbol{\Omega}}) = \sum_{r_x=-B}^{B} \sum_{r_y=-B}^{B} \Phi^A_{\tilde{s}\tilde{s}} \times ((\Omega_x - 2r_x\pi) f_{sx}, (\Omega_y - 2r_y\pi) f_{sy}). \tag{44}$$

Examples of $\Phi^A_{ss}(e^{j\boldsymbol{\Omega}})$ with $B = 3$ are shown in Fig. 5 (a_0), (b_0), and (c_0). In Fig. 5, (b_0) and (c_0), for instance, spatial aliasing due to undersampling of sharp image features can be clearly observed.

We have used the anisotropic model to analyze images and assess the potential coding gain by exploiting directionality on a block-by-block basis. Towards this end, we first divide the images into 32×32 blocks and then remove the mean in each block. Assuming that the source is stationary and $f_{sx} = f_{sy} = f_s$, we measure the autocorrelation function with discrete shifts up to two sampling periods both horizontally and vertically, i.e., $R_{\tilde{s}\tilde{s}}(\boldsymbol{\tau})$, $\boldsymbol{\tau} \in \{-2/f_s, \cdots, 2/f_s\}^2$. Additionally, we normalize ω_a and ω_b by a reference $\omega_r = -\ln(0.85) f_s$, where the number 0.85 is obtained by averaging the measured autocorrelation coefficients (autocorrelation divided by variance) with a horizontal or vertical shift of one sampling period over all mean-removed blocks in the test set. Variance σ^2 in the model is obtained directly from $R_{\tilde{s}\tilde{s}}((0,0)^T)$. The three parameters, ω_a/ω_r, ω_b/ω_r, and θ are fit to the other 12 measured unique autocorrelation

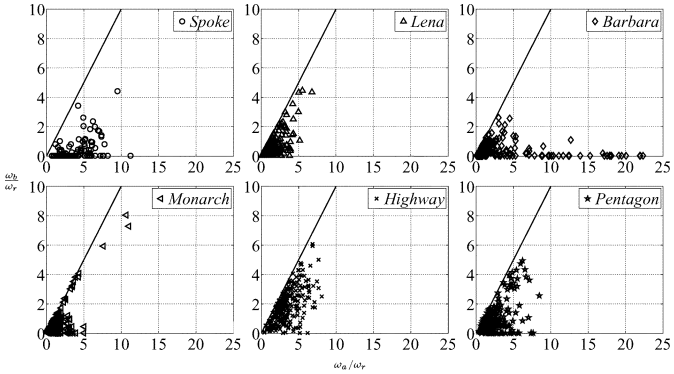


Fig. 4. Parameters of the anisotropic statistical model fit to image blocks in the test set. The solid line corresponds to $\omega_a = \omega_b$.

TABLE I
MODELING IMAGES WITH THE ANISOTROPIC STATISTICAL MODEL

Image	$\frac{\overline{\omega}_a}{\omega_r}$	$\frac{\overline{\omega}_b}{\omega_r}$	$\exp(-\frac{\overline{\omega}_a}{f_s})$	$\exp(-\frac{\overline{\omega}_b}{f_s})$	$\overline{G}_{D_1}$ (dB)	$\overline{G}_{D_2}$ (dB)
Spoke	2.24	0.06	0.70	0.99	3.23	1.91
Lena	0.86	0.20	0.87	0.97	1.00	0.88
Barbara	1.92	0.34	0.73	0.95	1.59	1.30
Monarch	1.25	0.44	0.82	0.93	0.63	0.55
Highway	3.13	1.70	0.60	0.76	0.24	0.22
Pentagon	1.97	0.69	0.73	0.89	0.68	0.47
Overall	2.01	0.41	0.72	0.93	1.48	1.03

values by solving nonlinear least squares. Fig. 4 plots the distribution of ω_a/ω_r and ω_b/ω_r for the test images shown in Fig. 9. Parameter θ in the model for the test set is approximately uniformly distributed.

To further characterize the entire image, in addition to modeling individual image blocks, we average the model autocorrelation with a shift of one sampling period along θ, i.e., $\sigma^2 \exp(-(\omega_a/f_s))$, over all blocks in the image, and denote the average by $\overline{R}_a$. Similarly, denote such an average orthogonal to θ by $\overline{R}_b$, and the average variance by $\overline{\sigma^2}$. Let $\overline{\omega}_a = -\ln(\overline{R}_a/\overline{\sigma^2})f_s$ and $\overline{\omega}_b = -\ln(\overline{R}_b/\overline{\sigma^2})f_s$, the image-wise parameters, expressed in the form of $\overline{\omega}_a/\omega_r$ and $\overline{\omega}_b/\omega_r$, and the corresponding autocorrelation coefficients $\exp(-(\overline{\omega}_a/f_s))$ and $\exp(-(\overline{\omega}_a/f_s))$, for the test set are listed in Table I. The local directionality of the images is well captured by $\overline{\omega}_a$ and $\overline{\omega}_b$. For instance, in *Spoke* the intensity changes rapidly across the lines in each 32×32 block but remains almost constant along the lines and in the flat area; therefore, a large $\overline{\omega}_a$ and a small $\overline{\omega}_b$ are observed. The image *Highway* represents a complex scene and each block typically contains sharp transitions in various directions; therefore, both $\overline{\omega}_a$ and $\overline{\omega}_b$ are large.

F. Numerical Evaluation

We consider the case that one level of the direction-adaptive transform is applied to the image samples $\mathbf{s}$. $\mathbf{s}$ is first filtered along direction $\mathbf{d}^\star$ and subsampled vertically to obtain $\mathbf{w}_0$ and $\mathbf{w}_1$. The direction $\mathbf{d}^\star$ is selected from $\mathbf{d}_i = (d_{i,x}, d_{i,y})^T$, such that the variance of $\mathbf{w}_1$ is minimized, as discussed in Section III-D. Furthermore, $\mathbf{w}_0$ and $\mathbf{w}_1$ are filtered along $\mathbf{d}'^\star$ and $\mathbf{d}''^\star$, respectively. Similarly, $\mathbf{d}'^\star$ and $\mathbf{d}''^\star$ are both selected from $\mathbf{d}'_i = (d'_{i,x}, d'_{i,y})^T = (d_{i,y}, d_{i,x})^T$. As discussed in Section II-B, the equivalent direction of $\mathbf{d}'_i$ in the original grid that $s[\mathbf{1}]$ resides in is denoted by $\breve{\mathbf{d}}_i = (d_{i,y}, 2d_{i,x})^T$.

Two sets of directions are considered in the numerical evaluation. The first set contains the nine directions described in Section II-B, denoted by $\mathbf{D}_1$ in this subsection. The other contains the sub-pel directions, denoted by $\mathbf{D}_2 = \{\mathbf{d}_i | \mathbf{d}_i = (i/4, 1)^T, i = -4, \cdots, 4\}$, assuming the required sub-pel interpolating filter is the ideal rectangular filter with transfer function $\sqcap_4(e^{j\Omega_t})$. We further assume that the ideal rectangular filter is adopted for the 1-D prototype prediction and update filter, as described in (30)–(33). Therefore, the four subbands partition the 2-D discrete-space frequency $-\pi \le \Omega_x \le \pi$ and $-\pi \le \Omega_y \le \pi$.

The direction-adaptation gain using $\mathbf{D}_1$ for image sources with $\omega_b/\omega_r = 0.4$ and different ω_a/ω_r and θ is plotted in Fig. 6(a). $\omega_b/\omega_r = 0.4$ is chosen as a representative value of the model parameter since it approximately corresponds to the overall autocorrelation with a shift of one sampling period orthogonal to θ, averaged over all blocks in the test set, as shown in Table I. Each curve, from light to dark, corresponds to $\omega_a/\omega_r =$ 1, 2, 4, 8, and 16. For θ, only the values from 0 to 90° are shown since the direction-adaptation gain is symmetric among the four quadrants. Additionally, the five dashed vertical lines (from left to right) depict $\theta = \tan^{-1}(d_{i,x}/d_{i,y})$, $i = 0, \cdots, 4$, corresponding to the directions orthogonal to $\mathbf{d}_0 \sim \mathbf{d}_4$. The five dashed-dotted vertical lines (from right to left) depict $\theta = \tan^{-1}(d_{i,y}/(2d_{i,x}))$, $i = 0, \cdots, 4$, corresponding to the directions orthogonal to $\breve{\mathbf{d}}_0 \sim \breve{\mathbf{d}}_4$. It can be observed that the direction-adaptation gain reaches local maxima at these directions since filtering is applied along the orientation of the image sources. For comparison, the direction-adaptation gain using $\mathbf{D}_2$ is plotted in Fig. 6(b). Note that the five dashed lines range from 0° to around 71.6° in Fig. 6(a), whereas those in Fig. 6(b) range from 0° to around 45°, indicating a more evenly-distributed coverage of the directions in $\mathbf{D}_1$. Similarly, the five dashed-dotted lines in Fig. 6(a) range from around 9.5° to 90°, whereas those in Fig. 6(b) range from around 26.6° to 90°. Moreover, the direction-adaptation gain in Fig. 6(a) is more symmetric about $\theta = 45°$ than that in Fig. 6(b) in general, indicating that the coding performance of the proposed DA-DWT using $\mathbf{D}_1$ is less sensitive to image transposition (swapping the x and y direction) than using $\mathbf{D}_2$.

From Fig. 6, the direction-adaptation gain provided by $\mathbf{D}_1$ is in general larger than that by $\mathbf{D}_2$, except for image sources with a small ω_a/ω_r and $0 < \theta < 45°$. To further interpret the difference between Fig. 6(a) and (b), we shall examine the subband-partitioning resulting from the direction-adaptive transform. Fig. 5 (a_1) and Fig. 5 (a_2) illustrate the subband-partitioning for the image model shown in Fig. 5 (a_0) using $\mathbf{D}_1$ and $\mathbf{D}_2$, respectively. For $\omega_a/\omega_r = 1$ and $\theta = 30°$, $\mathbf{d}^\star = \mathbf{d}_2 = (1/2, 1)^T$ (26.6°) with $\mathbf{D}_2$, whereas $\mathbf{d}^\star = \mathbf{d}_0 = (0, 1)^T$ with $\mathbf{D}_1$ since there is no candidate for $\mathbf{d}^\star$ close to 30° in $\mathbf{D}_1$. As a result, the direction-adaptation gain using $\mathbf{D}_2$ is 0.43 dB, higher than using $\mathbf{D}_1$ where there is nearly no gain.

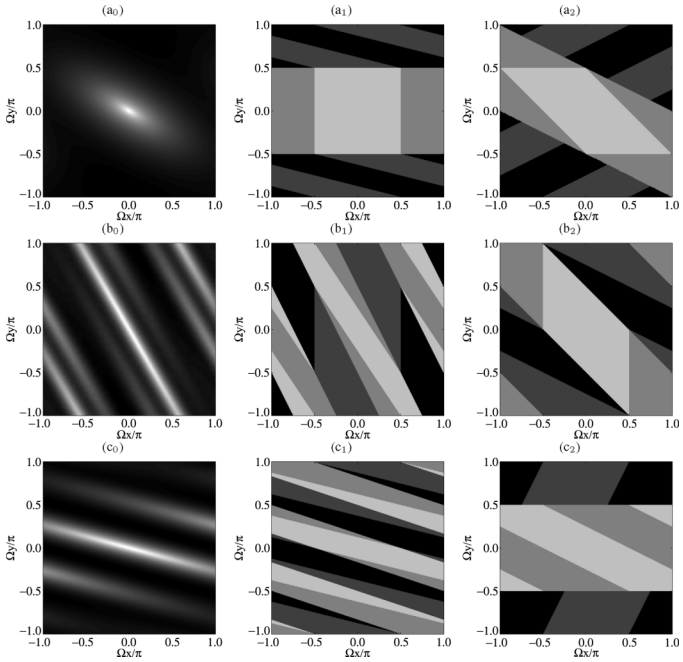


Fig. 5. (a_0), (b_0), and (c_0) show $\log_{10}\Phi_{ss}^{A}(e^{j\mathbf{\Omega}})$ from the anisotropic image model in (44) with $B = 3$, $\omega_b/\omega_r = 0.4$, and $\omega_a/\omega_r = 1$ and $\theta = 30°$ in (a_0); $\omega_a/\omega_r = 16$ and $\theta = 60°$ in (b_0); and $\omega_a/\omega_r = 8$ and $\theta = 15°$ in (c_0). In each plot, $\log_{10}\Phi_{ss}^{A}(e^{j\mathbf{\Omega}})$ is shifted and scaled to fit within the gray-scale range from 0 (black) to 1 (white). (a_1) and (a_2) illustrate the subband-partitioning for the image model shown in (a_0) using directions in $\mathbf{D}_1$ and $\mathbf{D}_2$, respectively. The colors from light to dark represent the subbands in the order of LL, LH, HL, and HH. Similarly, (b_1) and (b_2), and (c_1) and (c_2) illustrate the subband-partitioning for the image model shown in (b_0) and (c_0), respectively.

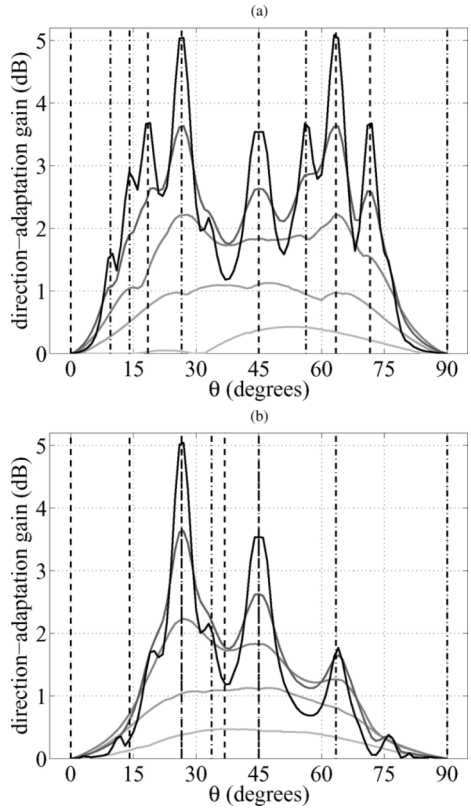


Fig. 6. Direction-adaptation gain of the anisotropic image model with the set of directions: (a) proposed $\mathbf{D}_1$, (b) sub-pel $\mathbf{D}_2$. $\omega_b/\omega_r = 0.4$ and a darker line corresponds to a larger $\omega_a/\omega_r = 1, 2, 4, 8, 16$. The five dashed vertical lines from left to right correspond to the directions orthogonal to $\mathbf{d}_0 \sim \mathbf{d}_4$. The five dashed-dotted vertical lines from right to left correspond to the directions orthogonal to $\check{\mathbf{d}}_0 \sim \check{\mathbf{d}}_4$.

For the image source in Fig. 5 (b_0), $\omega_a/\omega_r = 16$ and $\theta = 60°$. With $\mathbf{D}_1$, $\mathbf{d}^\star = \mathbf{d}_3 = (2,1)^T$ since the passband of the resulting prediction filter $\hat{H}_P^{(2,1)^T}(e^{j\mathbf{\Omega}})$, shown in Fig. 3(a), is closely aligned with the image PSD (63.4°). In addition, as illustrated in Fig. 5 (b_1), even the energy from the aliasing components is contained in LL and LH. With $\mathbf{D}_2$, $\mathbf{d}^\star = \mathbf{d}_4 = (1,1)^T$ (45°) since there is no direction closer to 60°. In Fig. 3(b) and Fig. 5 (b_2), the passband cannot account for the aliasing components and the aliasing energy is spread into HL and HH. As expected, the direction-adaptation gain with $\mathbf{D}_1$ is 2.84 dB, higher than the 0.83 dB with $\mathbf{D}_2$.

In Fig. 5, (c_0), $\omega_a/\omega_r = 8$ and $\theta = 15°$. Although the direction $\mathbf{d}_1 = (1/4,1)^T$ in $\mathbf{D}_2$ is closely aligned with θ (14°), direction $\mathbf{d}_0 = (0,1)^T$ is selected as $\mathbf{d}^\star$ instead. This is because the prediction filter $\hat{H}_P^{(1/4,1)^T}(e^{j\mathbf{\Omega}})$, shown in Fig. 3(b), although aligned with the image feature, fails to include in the passband the strong aliasing energy, for instance, near $(\Omega_x, \Omega_y) = (\pi, \pi/4)$. On the contrary, the direction $\mathbf{d}_1 = (1,3)^T$ in $\mathbf{D}_1$, although less aligned with the image feature (18.4°), results in a passband that contains major parts of the aliasing energy as shown in Fig. 3(a) and Fig. 5 (c_1). Furthermore, $\mathbf{d}'^\star = \check{\mathbf{d}}_3 = (d_{3,y}, 2d_{3,x})^T = (1,4)^T$ is selected from $\mathbf{D}_1$ ($\sim 14°$), whereas the closest direction in $\mathbf{D}_2$ is $\check{\mathbf{d}}_4 = (1,2)^T$ (26.6°). Due to better energy concentration, the direction-adaptation gain with $\mathbf{D}_1$ is 1.94 dB, higher than the 0.75 dB with $\mathbf{D}_2$.

The direction-adaptation gain from one level of the adaptive transform for each of the test images is also computed, using the model parameters fit to the image blocks. Since the distortion of each reconstructed image block is proportional to the geometric mean of the subband variances in the block (Section III-D), the distortion of the entire image is proportional to the sum of the block-wise geometric mean. Therefore, we define the model direction-adaptation gain of the entire image as the ratio (in decibels) between the sum of geometric mean from the conventional 2-D DWT and that from the direction-adaptive transform. The image-wise direction-adaptation gain using $\mathbf{D}_1$ and $\mathbf{D}_2$, denoted by $\overline{G}_{D_1}$ and $\overline{G}_{D_2}$, respectively, for each test image are listed in Table I.

To summarize, from the mathematical analysis and the numerical evaluation presented, we make the following conclusions. For anisotropic image sources, the DA-DWT realizes the direction-adaptation gain over the conventional 2-D DWT. The set of directions described in Section II-B is preferable to the sub-pel directions in the DA-DWT. For sharp image features, where the DA-DWT is most advantageous, the former achieves a higher direction-adaptation gain, while requiring less computational complexity, because it accounts for the aliasing energy due to undersampling. In addition, compression performance of the former is less sensitive to image transposition. Note that the analysis in this subsection assumes only one level of the transform with the ideal rectangular prediction and update filters, among other simplifications. The analysis does not aim to provide a quantitative prediction of the performance in practical systems, but rather a qualitative study of the improvement in

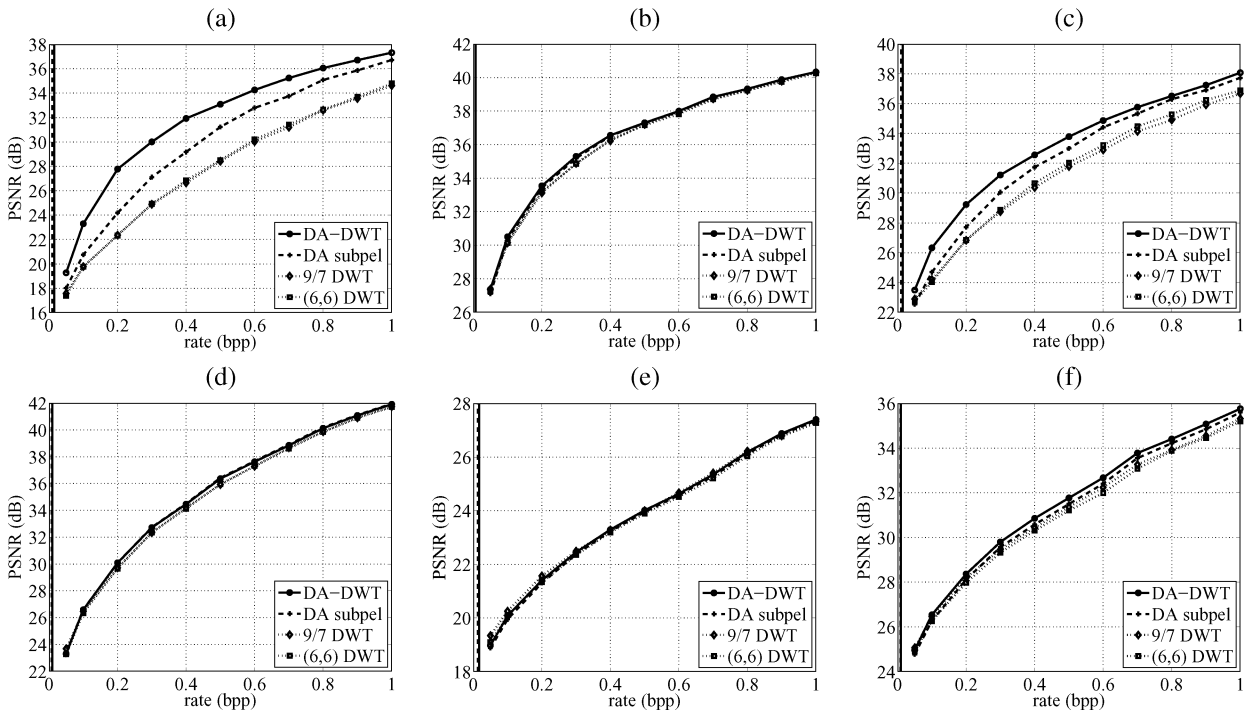


Fig. 7. Compression performance of the proposed DA-DWT, the sub-pel approach, and the conventional DWTs. The solid and the dashed vertical lines on the left of each plot represent the overhead spent for the DA-DWT and for the sub-pel approach, respectively. (a) *Spoke*; (b) *Lena*; (c) *Barbara*; (d) *Monarch*; (e) *Highway*; (f) *Pentagon*.

compression efficiency through exploiting the directional correlation in images. In Section IV, we shall present results from the practical implementation of DA-DWT to further justify the advantages of the proposed adaptive transform.

IV. Experimental Results

In the experiments, four levels of the DA-DWT, with the directional lifting filters and block-partitioning described in Section II, are applied to the 512×512 test images listed in Table I. The wavelet coefficients are encoded by the TCE embedded bitplane coder, which has been reported to achieve performance comparable to JPEG2000 [34]. To decompose the H subband further into HL and HH, we observe that adaptive selection of the filtering directions can be neglected without much loss in the compression performance. Therefore, $\breve{\mathbf{d}}_0 = (1,0)^T$ is always selected. Similarly, in the last level of the transform, $\mathbf{d}_0 = (0,1)^T$ and $\breve{\mathbf{d}}_0 = (1,0)^T$ are always selected to reduce the overhead.

The wavelet transforms, together with embedded bitplane coding, provide a scalable representation of images such that different reconstruction qualities are obtained by truncating a single bitstream at different rates. The compression performance of the DA-DWT for the test images is shown in Fig. 7. In addition, results from the conventional 9/7 wavelet adopted in JPEG2000 [35], and the (6,6) wavelet [19], equivalent to always selecting $\mathbf{d}_0$ and $\breve{\mathbf{d}}_0$ for the DA-DWT, are also shown. We also include the sub-pel approach proposed by others for comparison [15], [16]. The sub-pel approach we implemented differs from the DA-DWT only in the candidates for filtering directions, as described in Section II-B and Section III-F. They share the same prototype (6,6) wavelet, block-partition candidates, as well as the same formulation of the Lagrangian cost function. For spatial interpolation required in the sub-pel approach, the six-tap upsampling filter defined in the H.264 video coding standard is adopted [28].

In the experiments reported in this section, we do not optimize the choice of the Lagrangian multiplier for a specific reconstruction quality or for individual images. The same Lagrangian multiplier, 40, chosen empirically, is used for all test images. In Fig. 7, the overhead for the side information indicating the selected block-partition and filtering directions is included in the rate calculation, and represented by the solid and the dashed vertical lines, on the left of each plot, for the DA-DWT and the sub-pel approach respectively, ranges from 0.007 to 0.017 bpp. Note that in the experiments, the entirety of the side information is transmitted to the decoder at all decoding rates. At lower rates, the overhead takes a significant portion of the total rate available, leaving limited rate for the wavelet coefficients. In [29], we have developed a scalable direction representation such that a coarser direction representation, obtained with a larger Lagrangian multiplier, may be transmitted at lower rates to improve the performance.

In Fig. 7, the DA-DWT outperforms the conventional 9/7 DWT by up to 5.4 dB in PSNR for *Spoke*, 0.4 dB for *Lena*, 2.5 dB for *Barbara*, 0.4 dB for *Monarch*, and 0.5 dB for *Pentagon*. However, for *Highway*, the DA-DWT is inferior to the 9/7 DWT at lower rates mainly due to the cost of signaling the overhead. In addition to the natural images in the test set, significant gain of more than 1 dB is also observed for fingerprint images, as well as medical images where sharp edges occur at the boundary between the object of interest and the background. In general, the gain is larger for images rich of features with large intensity transition in one direction and small variation in

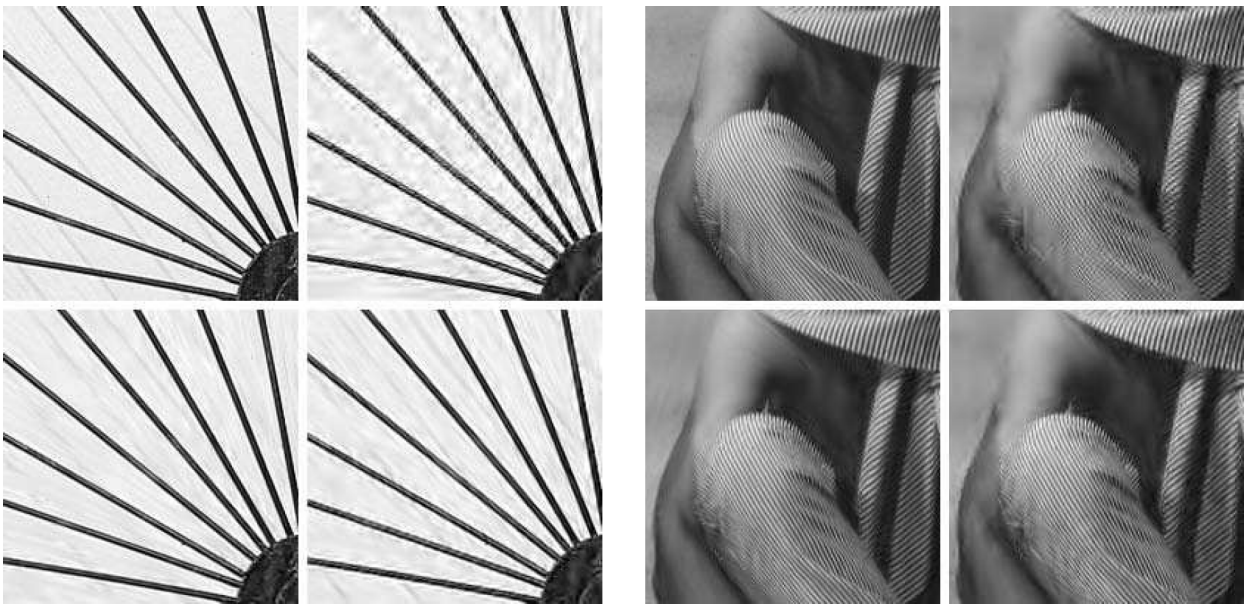


Fig. 8. Original and the reconstructions at 0.3 bpp of part of the test image *Spoke* and *Barbara*. For each test image, the four sub-images correspond to (top-left) the original; (top-right) 9/7 DWT; (bottom-left) DA-DWT; (bottom-right) sub-pel approach.

the orthogonal direction, qualitatively consistent with the direction-adaptation gain analysis using the anisotropic model presented in Section III (G_{D_1} in Table I).

The DA-DWT also outperforms the sub-pel approach by up to 3.6 dB for *Spoke*, 1.6 dB for *Barbara*, and 0.3 dB for *Pentagon*, and the gain is usually less than 0.1 dB for the other test images. These results are, again, consistent with the analysis in Section III, where $G_{D_1} - G_{D_2}$ in Table I also follows this order.

Part of the *Spoke* image is shown on the left of Fig. 8, along with the reconstructions at 0.3 bpp. The subjective quality from the DA-DWT is significantly better than the 9/7 DWT. Furthermore, in the reconstruction from the sub-pel approach, ringing artifacts appear around the lines in the spoke, especially at the horizontal-ish lines, whereas for the DA-DWT the artifacts are mostly absent. These results show that, compared to the sub-pel approach, the DA-DWT better handles sharp image features and is less sensitive to the feature orientation, as also suggested by the mathematical analysis. Similarly, on the right of Fig. 8, DA-DWT is clearly superior to both the 9/7 DWT and the sub-pel approach, around the arm and along the stripes on the fabrics in *Barbara*.

The original and the reconstructions of the six test images, using the 9/7 DWT and the DA-DWT, are shown in Fig. 9. For each image, we enlarge two 125 $\times$ 125 regions to show the difference in subjective quality between the 9/7 DWT and the DA-DWT, reconstructed at 0.05, 0.1, 0.2, and 0.4 bpp. At lower rates, the reconstruction error from the DA-DWT exhibits brushstroke-like artifacts along image features. Different from the ringing artifacts in the conventional transform, the brushstroke artifacts follow the geometric flow in the image and, therefore, better preserve the geometric structure, as shown in Fig. 9, for example, around the hat in *Lena* and the books in *Barbara*. As another example, for *Highway* reconstructed at 0.05, 0.1, and 0.2 bpp, in terms of the PSNR measurement the DA-DWT performs worse than the 9/7 DWT as plotted in Fig. 7. However, we argue that the reconstruction from the DA-DWT better represents the geometric content in the image and is visually more pleasing as shown in Fig. 9. This phenomenon is observed in all the images in the test set as well as other images not reported in this paper.

V. Conclusion

The DA-DWT provides an efficient representation for directional image features, such as edges and lines in the image, by locally adapting the filtering directions to the image content. In our experiments, the DA-DWT outperforms the conventional 2-D DWT by up to 2.5 dB in PSNR for typical test images, and by up to 5.4 dB for special classes of images. In general, the gain from exploiting directional correlation in the image is larger for image features with large intensity transition in one direction and small variation in the orthogonal direction, qualitatively consistent with the mathematical analysis using the proposed anisotropic image model.

In addition, compared to other lifting-based approaches with sub-pel directions, the DA-DWT accounts for the aliasing energy due to undersampling of sharp image features, thus provides higher compression efficiency while requiring less computation. Compression performance of the DA-DWT is also less sensitive to the orientation of image features.

Significant subjective improvement over the conventional 2-D DWT is observed. At low rates, the reconstruction error in the DA-DWT exhibits brushstroke-like artifacts. Different from the ringing artifacts typically observed in the conventional transform, the brushstroke artifacts better preserve the geometric structure in the image. The ability to better depict the image content at low rates makes the DA-DWT especially suitable for progressive transmission of images, where a low-quality preview is first reconstructed and, as more data are received, refined to achieve higher qualities.

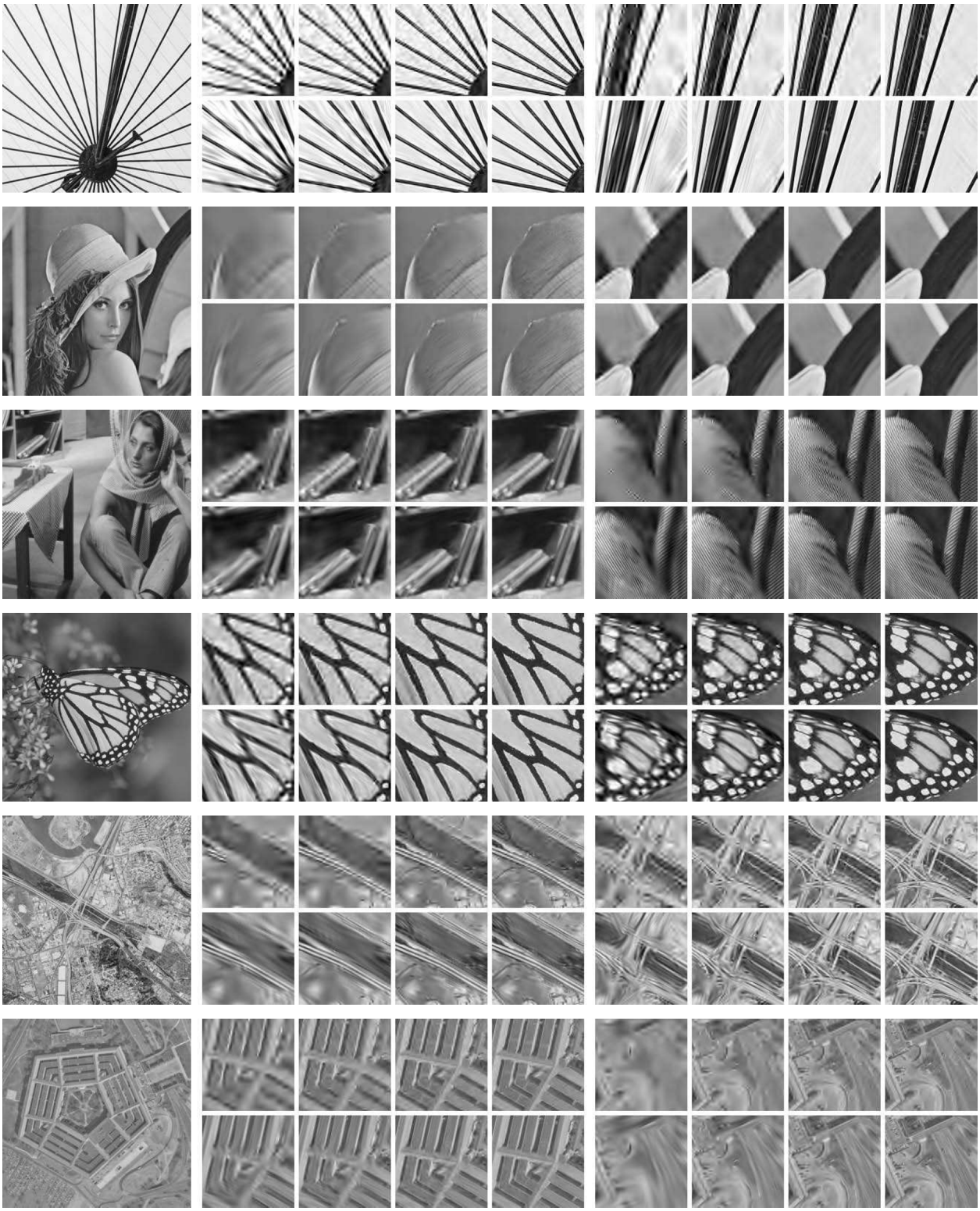


Fig. 9. Images on the left are the originals of the 512 × 512 test images, from top to bottom, *Spoke*, *Lena*, *Barbara*, *Monarch*, *Highway*, and *Pentagon*. For each test image, two 125 × 125 regions of the reconstruction from the conventional 9/7 DWT and the proposed DA-DWT are shown at the top and the bottom row, respectively, reconstructed at 0.05, 0.1, 0.2, and 0.4 bpp from left to right for each region.

APPENDIX

From (8) and (10), convolving $h_P^d[l]$ with $s_0[l]$ is equivalent to convolving $\hat{h}_P^d[l]$ with $\hat{s_0}[l]$ followed by a one-pixel advance and downsampling vertically, i.e.,

$$\left(h_P^d * s_0\right)[l] = \left(\hat{h}_P^d * \hat{s_0}\right)[l_x, 2l_y + 1]. \tag{45}$$

Hence

$$H_P^d(e^{j\Omega}) \cdot S_0(e^{j\Omega}) = \frac{1}{2}\sum_{k=0}^{1} \hat{H}_P^d\left(e^{j\Omega_x}, e^{j\left(\frac{\Omega_y}{2}+k\pi\right)}\right) \cdot e^{j\left(\frac{\Omega_y}{2}+k\pi\right)} \cdot \hat{S}_0\left(e^{j\Omega_x}, e^{j\left(\frac{\Omega_y}{2}+k\pi\right)}\right). \quad (46)$$

From (9)

$$\hat{S}_0(e^{j\Omega}) = S_0(e^{j\Omega_x}, e^{j2\Omega_y}). \quad (47)$$

In addition, from the fact that d_y is always an odd integer and (10), $\hat{h}_P^d[l] = 0$ if l_y is even. Therefore

$$\hat{H}_P^d(e^{j\Omega}) = -\hat{H}_P^d\left(e^{j\Omega_x}, e^{j(\Omega_y+k\pi)}\right), \quad \forall k : k \text{ is odd}. \quad (48)$$

Combining (46)–(48), we obtain

$$H_P^d(e^{j\Omega}) = \frac{1}{2}\sum_{k=0}^{1}(-1)^k e^{j\frac{\Omega_y}{2}} \hat{H}_P^d\left(e^{j\Omega_x}, e^{j\left(\frac{\Omega_y}{2}+k\pi\right)}\right) = e^{j\frac{\Omega_y}{2}} \hat{H}_P^d\left(e^{j\Omega_x}, e^{j\frac{\Omega_y}{2}}\right). \quad (49)$$

REFERENCES

[1] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform," *IEEE Trans. Image Process.*, vol. 1, no. 2, pp. 205–220, Apr. 1992.

[2] 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, no. 3, pp. 243–250, Jun. 1996.

[3] D. S. Taubman and M. W. Marcellin, *JPEG2000: Image Compression Fundamentals, Standards and Practice*. Norwell, MA: Kluwer, 2002.

[4] D. Taubman and A. Zakhor, "Orientation adaptive subband coding of images," *IEEE Trans. Image Process.*, vol. 3, no. 4, pp. 421–437, Jul. 1994.

[5] V. Velisavljevic, B. Beferull-Lozano, M. Vetterli, and P. L. Dragotti, "Directionlets: Anisotropic multi-directional representation with separable filtering," *IEEE Trans. Image Process.*, vol. 15, no. 7, pp. 1916–1933, Jul. 2006.

[6] G. Peyre and S. Mallat, "Surface compression with geometric bandelets," *ACM Trans. Graph.*, vol. 24, Aug. 2005.

[7] ——, "Discrete bandelets with geometric orthogonal filters," in *Proc. IEEE Int. Conf. Image Processing*, Genova, Italy, 2005, pp. I-65–I-68.

[8] W. Sweldens, "The lifting scheme: A construction of second generation wavelets," *SIAM J. Math. Anal.*, vol. 29, no. 2, pp. 511–546, 1998.

[9] I. Daubechies and W. Sweldens, "Factoring wavelet transforms into lifting steps," *J. Fourier Anal. Appl.*, vol. 4, no. 3, pp. 245–267, 1998.

[10] O. N. Gerek and A. E. Cetin, "Adaptive polyphase subband decomposition structures for image compression," *IEEE Trans. Image Process.*, vol. 9, no. 10, pp. 1649–1660, Oct. 2000.

[11] N. V. Boulgouris, D. Tzovaras, and M. G. Strintzis, "Lossless image compression based on optimal prediction, adaptive lifting and conditional arithmetic coding," *IEEE Trans. Image Process.*, vol. 10, no. 1, pp. 1–14, Jan. 2001.

[12] R. L. Claypoole, G. M. Davis, W. Sweldens, and R. G. Baraniuk, "Nonlinear wavelet transforms for image coding via lifting," *IEEE Trans. Image Process.*, vol. 12, no. 12, pp. 1449–1459, Dec. 2003.

[13] D. Taubman, "Adaptive, non-separable lifting transforms for image compression," in *Proc. IEEE Int. Conf. Image Processing*, Kobe, Japan, Oct. 1999, vol. 3, pp. 772–776.

[14] O. N. Gerek and A. E. Cetin, "A 2-D orientation-adaptive prediction filter in lifting structures for image coding," *IEEE Trans. Image Process.*, vol. 15, no. 1, pp. 106–111, Jan. 2006.

[15] W. Ding, F. Wu, and S. Li, "Lifting-based wavelet transform with directionally spatial prediction," presented at the Picture Coding Symp., San Francisco, CA, Dec. 2004.

[16] D. Wang, L. Zhang, A. Vincent, and F. Speranza, "Curved wavelet transform for image coding," *IEEE Trans. Image Process.*, vol. 15, no. 8, pp. 2413–2421, Aug. 2006.

[17] C.-L. Chang, A. Maleki, and B. Girod, "Adaptive wavelet transform for image compression via directional quincunx lifting," in *Proc. IEEE Workshop on Multimedia Signal Processing*, Shanghai, China, Oct. 2005.

[18] C.-L. Chang and B. Girod, "Direction-adaptive discrete wavelet transform via directional lifting and bandeletization," presented at the IEEE Int. Conf. Image Processing, Atlanta, GA, Oct. 2006.

[19] R. Calderbank, I. Daubechies, W. Sweldens, and B.-L. Yeo, "Wavelet transforms that map integers to integers," *Appl. Comput. Harmon. Anal.*, vol. 5, no. 3, pp. 332–369, 1998.

[20] A. Gouze, M. Antonini, M. Barlaud, and B. Macq, "Design of signal-adapted multidimensional lifting scheme for lossy coding," *IEEE Trans. Image Process.*, vol. 13, no. 12, pp. 1589–1603, Dec. 2004.

[21] B. Girod, "Rate-constrained motion estimation," in *Proc. SPIE Visual Communications and Image Processing*, Chicago, IL, 1994, pp. 1026–1034.

[22] T. Wiegand and B. Girod, "Lagrange multiplier selection in hybrid video coder control," in *Proc. IEEE Int. Conf. Image Process.*, Thessaloniki, Greece, Oct. 2001, vol. 3, pp. 542–545.

[23] G. J. Sullivan and T. Wiegand, "Rate-distortion optimization for video compression," *IEEE Signal Process. Mag.*, vol. 15, pp. 74–90, Nov. 1998.

[24] J.-R. Ohm, "Three-dimensional subband coding with motion compensation," *IEEE Trans. Image Process.*, vol. 3, no. 5, pp. 559–571, Sep. 1994.

[25] D. Wang, L. Zhang, and A. Vincent, "Improvement of JPEG2000 using curved wavelet transform," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing*, Philadelphia, PA, Mar. 2005, vol. 2, pp. 365–368.

[26] E. L. Pennec and S. Mallat, "Sparse geometric image representations with bandelets," *IEEE Trans. Image Process.*, vol. 14, no. 4, pp. 423–438, Apr. 2005.

[27] R. Xiong, F. Wu, J. Xu, S. Li, and Y.-Q. Zhang, "Barbell lifting wavelet transform for highly scalable video coding," presented at the Picture Coding Symp., San Francisco, CA, Dec. 2004.

[28] T. Wiegand, G. J. Sullivan, G. Bjntegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 560–576, Jul. 2003.

[29] T. Xu, C.-L. Chang, and B. Girod, "Scalable direction representation for image compression with direction-adaptive discrete wavelet transform," presented at the SPIE Visual Communications and Image Processing, San Jose, CA, Jan. 2007.

[30] A. Secker and D. Taubman, "Highly scalable video compression with scalable motion coding," *IEEE Trans. Image Process.*, vol. 13, no. 8, pp. 1029–1041, Aug. 2004.

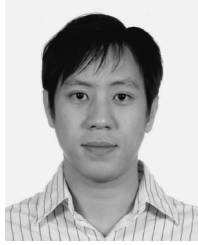
[31] R. Xiong, J. Xu, F. Wu, S. Li, and Y.-Q. Zhang, "Layered motion estimation and coding for fully scalable 3D wavelet video coding," in *Proc. IEEE Int. Conf. Image Processing*, Singapore, Oct. 2004, pp. 2271–2274.

[32] G. A. Jones and J. M. Jones, *Elementary Number Theory*. New York: Springer-Verlag, 1998.

[33] J. B. O'Neal, Jr. and T. R. Natarajan, "Coding isotropic images," *IEEE Trans. Inf. Theory*, vol. 23, no. 6, pp. 697–707, Nov. 1977.

[34] C. Tian and S. S. Hemami, "An embedded image coding system based on tarp filter with classification," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Montréal, QC, Canada, May 2004, vol. 3, pp. 49–52.

[35] A. Cohen, I. Daubechies, and J.-C. Feauveau, "Biorthogonal bases of compactly supported wavelets," *Commun. Pure Appl. Math.*, vol. 45, pp. 485–560, 1992.



Chuo-Ling Chang (S'03) received the B.S. degree in electrical engineering from the National Taiwan University, Taipei, Taiwan, R.O.C., in 1998, and the M.S. degree in electrical engineering from Stanford University, Stanford, CA, in 2002, where he is currently pursuing the Ph.D. degree in electrical engineering at the Information Systems Laboratory.

His research interests include scalable coding and streaming of multimedia data.

Bernd Girod (S'80–M'80–M'97–F'98) received the M.S. degree ("with highest honors") from the Georgia Institute of Technology (Georgia Tech), Atlanta, in 1980, and the Engineering Doctorate from the University of Hannover, Germany, in 1987.

He is a Professor of electrical engineering in the Information Systems Laboratory of Stanford University, Stanford, CA, where he also holds a courtesy appointment with the Department of Computer Science. He serves as Director of both the Stanford Center for Image Systems Engineering (SCIEN) and the Max Planck Center for Visual Computing and Communication. Since 2004, he has served as the Chairman of the Steering Committee of the new Deutsche Telekom Laboratories, Technical University of Berlin, Berlin, Germany. Until 1987, he was a Member of the Research Staff at the Institut für Theoretische Nachrichtentechnik und Informationsverarbeitung, University of Hannover. In 1988, he joined Massachusetts Institute of Technology, Cambridge, first as a Visiting Scientist with the Research Laboratory of Electronics, then as an Assistant Professor of media technology at the Media Laboratory. From 1990 to 1993, he was a Professor of Computer Graphics and the Technical Director of the Academy of Media Arts, Cologne, Germany, jointly appointed with the Computer Science Section. He was a Visiting Adjunct Professor with the Digital Signal Processing Group, Georgia Tech, in 1993. From 1993 to 1999, he was Chaired Professor of electrical engineering/telecommunications at the University of Erlangen-Nuremberg, Germany, and the Head of the Telecommunications Institute I, co-directing the Telecommunications Laboratory. He served as the Chairman of the Electrical Engineering Department from 1995 to 1997, and as Director of the Center of Excellence "3-D Image Analysis and Synthesis" from 1995 to 1999. He was a Visiting Professor with the Information Systems Laboratory, Stanford University, during the 1997–1998 academic year. As an entrepreneur, he has worked successfully with several start-up ventures as Founder, investor, Director, or Advisor. Most notably, he has been a Co-Founder and Chief Scientist of Vivo Software, Inc., Waltham, MA (1993-1998); after Vivo's aquisition (1998–2002), he was Chief Scientist of RealNetworks, Inc. (Nasdaq: RNWK). He served on the Board of Directors for 8x8, Inc., Santa Clara, CA, (Nasdaq: EGHT) from 1996 to 2004, and for GeoVantage, Inc., Swampscott, MA, from 2000 to 2005. He is currently an Advisor to the start-up companies Mobilygen, Santa Clara, and NetEnrich, Inc., Santa Clara. He has authored or coauthored one major textbook (printed in three languages), three monographs, and over 350 book chapters, journal articles, and conference papers, and he holds over 20 U.S. patents. His research interests include video coding and networked media systems.

Prof. Girod has served as on the editorial boards for several journals in his field, among them as Area Editor for Speech, Image, Video, and Signal Processing of the IEEE TRANSACTIONS ON COMMUNICATIONS. He has served on numerous conference committees, e.g., as Tutorial Chair for ICASSP 1997 in Munich, Germany; for ICIP 2000 in Vancouver, BC, Canada; as General Chair for the 1998 IEEE Image and Multidimensional Signal Processing Workshop in Alpbach, Austria; as General Chair for the Visual Communication and Image Processing Conference (VCIP) in San Jose, CA, in 2001; and as General Chair of Vision, Modeling, and Visualization (VMV) in Stanford in 2004. He was a member of the IEEE Image and Multidimensional Signal Processing Technical Committee from 1989 to 1997. He was elected Fellow of the IEEE in 1998 "for his contributions to the theory and practice of video communications." He was named Distinguished Lecturer for the year 2002 by the IEEE Signal Processing Society. He received the 2002 EURASIP Best Paper Award (with J. Eggers) and the 2004 EURASIP Techical Achievement Award.