SEGMENTATION-DRIVEN DIRECTION-ADAPTIVE DISCRETE WAVELET TRANSFORM

Adrian Munteanu, Oana Maria Surdu, Jan Cornelis, Peter Schelkens

Vrije Universiteit Brussel, Department of Electronics and Informatics (ETRO) – Interdisciplinary Institute for Broadband Technology (IBBT), Pleinlaan 2, B-1050 Brussels, Belgium

ABSTRACT

This paper proposes a novel segmentation-driven direction-adaptive discrete wavelet transform (SD DADWT), wherein the adaptation of the directional wavelet bases is performed on the segments describing the natural geometry of the image. First, a multi-resolution segmentation of the image is performed, obtained through an Edgmentation procedure. The optimum lifting directions are then selected for each segment and at each resolution. The proposed SD DADWT retains the inherent advantages offered by a multiresolution representation of the geometric features in the image, and in the same time provides a sparse image representation via DADWT. Preliminary experimental results obtained in a coding application show that the visual quality of the reconstructed image can be further improved by applying a geometrically-oriented transform on segments that approximate the natural borders in the image.

Index Terms — adaptive wavelets, scalable image coding

1. INTRODUCTION

Providing a scalable representation of images and video is of paramount importance in data transmission over heterogeneous networks in order to allow for the adaptation of the source to the inherently variable network conditions and terminal characteristics. In this context, *sparse signal representations* and embedded quantization and coding are the keys in order to enable scalability. As an example, recent years have witnessed the advent of the wavelet-based JPEG 2000 standard for scalable image compression. In this context, the wavelet transform offers a sparse multiresolution image representation, and enables a rich set of functionalities, such as resolution and quality scalability, and region-of-interest coding.

Another challenge in the application of wavelets in both practice and theory is capturing the geometric features in images and accounting for these features. In this context, transforms that perform directional adaptation of the bases include the bandelets of [1] or lifting-based [2] approaches locally-adapting the filtering directions to the geometric flow in the image [3]-[4]. Belonging to this second category, the direction-adaptive discrete wavelet transform (DADWT) of [5] decomposes the image into rectangular regions and adapts the orientation of the wavelet basis in each region. When implemented in a practical coding system, the DADWT significantly improves the compression performance if compared to an equivalent system employing the classical DWT [5].

From a complementary perspective, another important challenge in nowadays applications is capturing and *representing the geometric features* present in images. In this sense, providing a multi-resolution image segmentation and shape extraction are two major pre-processing objectives for any image analysis task or object-based coding approach.

In this paper we propose a multi-resolution image representation that addresses this combined problem. The paper proposes a novel segmentation-driven direction-adaptive discrete wavelet transform (SD DADWT), wherein the adaptation of the directional wavelet bases is performed both in terms of orientation, similar to DADWT, as well as spatially, i.e. on the segments capturing the natural geometry of the image. In this way, the proposed SD DADWT offers a multiresolution representation of the geometric features in the image, via a multiresolution image segmentation, and in the same time provides a sparse image representation via DADWT. The transform is synthesized based on lifting, which enables a fast implementation, as well as a lossless transform. A coding system employing the proposed SD DADWT is designed, and preliminary experimental results are provided.

The remainder of the paper is organized as follows. Section 2 presents the principles of direction-adaptive discrete wavelet transform. Section 3 presents the Edgmentation algorithm employed for image segmentation. In Section 4 we explain the proposed codec architecture. The experimental results are given and discussed in Section 5. Finally, Section 6 concludes our work.

2. DIRECTION-ADAPTIVE DISCRETE WAVELET TRANSFORM (DADWT)

The DADWT proposed in [5] is a critically-sampled discrete wavelet transform implemented with lifting [2], in which the lifting operations are adapted to the local geometry in the image. Basically, in the approach of [5] the image is split into blocks, and within each block a direction-adaptive discrete wavelet transform is performed by employing a "directional" lifting scheme.

Following the notations in [5], let $\mathbf{X}$ denote an image defined on an orthogonal sampling-grid Π composed of 4 different sub-grids $\Pi = \bigcup \Pi_{pq}$, with $\Pi_{pq} = \{(m,n) \mid m \bmod 2 = p,\, n \bmod 2 = q\}$. Similar to the classical DWT, the input signal $\mathbf{X}$ is firstly decomposed into even ($\mathbf{X}_0 = \{l_0 \mid l_0 \in \Pi_0 = \Pi_{00} \cup \Pi_{01}\}$) and odd ($\mathbf{X}_1 = \{l_1 \mid l_1 \in \Pi_1 = \Pi_{10} \cup \Pi_{11}\}$) rows respectively. Then, the lifting scheme predicts the even rows from the odd ones, resulting into the detail signal $\mathbf{H} = \{H[l_1], l_1 \in \Pi_1\}$. This signal is used to update the even rows in order to produce the approximation signal $\mathbf{L} = \{L[l_0], l_0 \in \Pi_0\}$. The 1D vertical lifting steps are thus:

$$\begin{aligned} H[l_1] &= 1/G\left(X[l_1] - P_{X,l_1}(\mathbf{X}_0)\right), \forall l_1 \in \Pi_1 \\ L[l_0] &= G\left(X[l_0] + G \cdot U_{X,l_0}(\mathbf{H})\right), \forall l_0 \in \Pi_0 \end{aligned} \qquad (1)$$

1-4244-1437-7/07/$20.00 ©2007 IEEE

where G is a scaling factor and $P(\cdot), U(\cdot)$ are predict and update functions respectively, taking as input sets of samples in $\mathbf{X}_0$ and $\mathbf{H}$ respectively, and producing a scalar output.

A subsequent lifting step performed on columns decomposes $\mathbf{L}$ into the subbands $\mathbf{LL}$ and $\mathbf{LH}$ defined on the grids Π_{00} and Π_{01} respectively, and $\mathbf{H}$ into the subbands $\mathbf{HL}$ and $\mathbf{HH}$ defined on Π_{10} and Π_{11}, respectively.

The DADWT employs N_c candidate directional predictors, $P^i_{X,l_1}(\cdot)$ for each block. The predict and update functions are defined as [5]:

$$P^i_{X,l_1}(\mathbf{X}_0) = \sum_{k=0}^{K^p-1} c_k^p \cdot \left(X[l_1 - (2k+1)v_i] + X[l_1 + (2k+1)v_i] \right)$$

$$U_{X,l_0}(\mathbf{H}) = \sum_{k=0}^{K^u-1} c_k^u \cdot \Big(\sum_{\{l_1 | l_1 - (2k+1)\overline{v}_{l_1} = l_0\}} H[l_1] + \sum_{\{l_1 | l_1 + (2k+1)\overline{v}_{l_1} = l_0\}} H[l_1] \Big)$$

In the equations above $i = 0 \ldots N_c - 1$ is the direction index, and v_i is the direction candidate; the v_i s are defined such that all $l_1 \pm (2k+1)v_i$ map to an even l_0 index, for any l_1, k (l_1 being odd). In the considered DADWT instantiation, $K^p = K^u = 3$, while c_k^p, c_k^u are the predict- and update-filter coefficients respectively, as specified in [5]. The second equation states that whenever an image sample $X(l_1)$ is predicted by $c_k^p X[l_0]$, the corresponding $L(l_0)$ is updated by $c_k^u H[l_1]$.

For compression purposes, it is desirable to select for each block the predictor for which the magnitude of the residual $H[l_1]$ is minimized. Thus, for every image block, the optimum v_i is chosen as the direction that minimizes the prediction error $H[l_1]$ within the block. Once a prediction direction is chosen, the "flow" vectors v_i are reverted and used in the update step, as given above.

For a multi-level wavelet transform, the vertical and horizontal directional lifting steps are repeated successively for every $\mathbf{LL}$. The inverse transform reconstructs the original signal performing the same steps in reverse order.

One notes that in order to produce an integer version of (1), for the particular DADWT instantiation used in this paper, we can multiply by G both the low- and band-pass terms in (1) and employ the integer-part operation $\lfloor \cdot \rfloor$ as follows:

$$\begin{aligned} H'[l_1] &= X[l_1] - \left\lfloor P_{X,l_1}(\mathbf{X_0}) \right\rfloor, \forall l_1 \in \Pi_1 \\ L'[l_0] &= G^2 \cdot \left(X[l_0] + \left\lfloor U_{X,l_0}(\mathbf{H}') \right\rfloor \right), \forall l_0 \in \Pi_0 \end{aligned} \quad (2)$$

For the employed interpolating transforms [5] one has $G = \sqrt{2}$, implying that the coefficients produced by (2) are indeed integers.

3. EDGMENTATION

In the proposed SD DADWT any type of multiresolution segmentation algorithm can be practically employed. Basically, the segmentation tool only drives the directional lifting process at every resolution level. Hence, the choice of the employed segmentation tool and/or segmentation criteria can be made according to the scope of the targeted image analysis application.

In our SD DADWT implementation, our previously developed Edgmentation algorithm [6]-[8] is employed. Edgmentation can be considered a modified split-and-merge algorithm, where the splitting step is performed given the edge information in the image and the merging step occurs according to the grey-value characteristics of the segment.

3.1. Splitting procedure

The splitting step of the Edgmentation algorithm involves a "segment and refine" method. The image partitioning is coordinated by the edge information, as described below.

First, the local maxima in the reverse of the edge image are considered as roots. Every root is labeled with the grey value of the corresponding pixel from the original image. Afterwards, pixels are linked to one another along the steepest path in the reverse edge image, until every pixel is linked to a root [6].

In the refinement procedure, the segments are checked for their grey value homogeneity. The employed measure of homogeneity is the absolute difference between the minimum and the maximum of the grey values existing in a region [7], [8]. The segments that do not satisfy the homogeneity criteria are further segmented.

After the splitting step is performed, a large number of segments are produced, as shown in the example of Fig.2 (a). Many of these segments are meaningless and some of them correspond to parts of the same object. Further on, a merging step is necessary in order to solve this drawback.

3.2. Merging procedure

Two neighboring regions are merged if the absolute value of a cost calculated between these regions is less than an imposed threshold T_m [8]:

$$\left| C_{ij} \right| \le T_m ,$$

where C_{ij} is a cost function expressing the differences between the pixels inside adjacent regions i and j:

$$C_{ij} = GS_{ij} \times GR_{ij} \times SL_{ij} \times SM_{ij} ,$$

where:

$$GS_{ij} = \left| \mu_i - \mu_j \right|$$

is the grey-level similarity measure between regions i and j, and μ_i is the mean grey value in region i;

$$GR_{ij} = gradient_{ij} \big/ shared\_contour\_length_{ij}$$

is a measure of the strength of the edge separating the regions i and j;

$$SL_{ij} = \min(n_i, n_j) \big/ shared\_contour\_length_{ij}$$

controls the size of the produced regions. n_i, n_j are the sizes of regions i and j respectively;

$$SM_{ij} = shared\_contour\_length_{ij} \big/ no\_of\_vertices$$

represents the contour-coding "easiness" factor [8], and $no\_of\_vertices$ indicates the number of vertices needed to approximate the contour using a polygon approximation.

An example of the result of the merging step is given in Fig.2 (b).

4. SEGMENTATION-DRIVEN DADWT

The proposed segmentation-driven DADWT and the design of a scalable image coding system that makes use of this representation are given in the following.

The block diagram of our system is depicted in Figure 3. First, a label image is produced by segmenting the input image using the previously described Edgmentation algorithm.

(a) Over-segmented image (b) After merging (c) Approximated contours

Figure 1. Edgmentation results for “Barbara”, (a) after the splitting step, and (b) after merging. (c) Linear approximation of the contours.

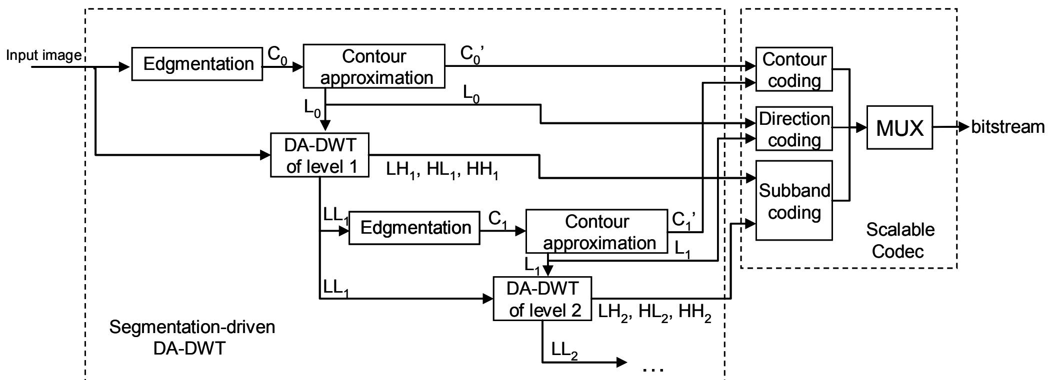

Figure 3. Block diagram of the proposed SD DADWT and scalable codec architecture making use of this representation.

The label image is the result of a splitting and merging step controlled by two parameters, one specifying the final number of produced regions, and a second one specifying the smallest accepted size for any segment. The label image is further transformed into a contour map C_0, representing the contours of the regions in the segmentation produced at the original resolution level (level 0).

Further on, the contours in C_0 are simplified by keeping only 8-connected pixel contours, and then approximated by polygons [8], in order to reduce the rate needed to represent C_0. This means that the contours are approximated by sets of points (vertices) that are connected by straight lines [8]. The approximation is produced by ensuring that the distance between the original contours and the approximated ones does not exceed a pre-defined threshold. An example of such an approximation is given in Fig. 2(c). We note that in our approach we use linear approximations, but higher order approximations (e.g. splines) are also feasible.

The result of the approximation technique is a contour image C_0'. Based on C_0' one generates a label image L_0 indicating the corresponding predictors (and updaters) to be used for the DADWT of level 1. In our implementation we consider $N_c = 11$ direction candidates [5] for the vertical lifting step, as well as for the horizontal one. The predictors are chosen segment-wise. That is, for each segment in L_0, we perform an exhaustive search among all possible directions and choose the direction that minimizes the mean square error (MSE) between the original segment and the predicted one. Later developments will include the minimization of a functional accounting both for the distortion and rate needed to encode the predictors, similar to [5].

As it concerns the proposed coding system, the predictors are encoded using a first-order entropy coder. The contour image C_0' is encoded using a high-order arithmetic coder. The information sent to the contour coder is the starting point’s coordinates for each contour, and the coordinate differences between every two consecutive vertices. Triple points are used [8] in order to signal the fact that a point could be part of more than one contour.

The subbands obtained by transforming the initial image, are encoded using the QT-L coder of [9], except the LL subband, which is recursively decomposed and coded, until the required number of decomposition levels is attained.

We observe that, in order to reduce the rate needed to encode the contour maps at all the resolution levels, it is possible to apply Edgmentation only at the highest resolution level (level 0), and reuse the segmentation results at the lower resolution levels, by an appropriate downsampling of the label image L_0.

Additionally, instead of approximating the contours in C_0 by using polygons, an alternative solution is to employ a scalable contour coding technique. In this way, we can losslessly decode C_0, if needed, but also extract C_0' as a lower-rate version of C_0. Designing such a scalable contour codec, and studying the optimum rate allocation between the contour and subband codecs are left as topics of future investigation.

5. EXPERIMENTAL RESULTS

This section reports preliminary experimental results obtained with a scalable coding system employing the proposed SD DADWT and compares them against those obtained using (*i*) the classical lifting-based implementation of the DWT [2], and (*ii*) the conventional block-based DADWT of [5]. All systems employ the same scalable subband codec, which is the QT-L codec of [9].

Coding rate	Conventional DADWT	Proposed codec, without contour coding	Proposed codec, with contour coding	Classical lifting
0.125 bpp	26,27 dB	26,39 dB	25,19 dB	23,57 dB
0.25 bpp	29,32 dB	29,67 dB	28,88 dB	28,08 dB
0.3 bpp	30,40 dB	30,75 dB	29,79 dB	28,96 dB
1 bpp	38,17 dB	38,47 dB	38,16 dB	37,58 dB
2 bpp	43,70 dB	43,80 dB	43,66 dB	43,58 dB

Table 1. PSNR values for test image Barbara, 4 resolution levels.

The results are reported in Table 1. In these experiments the DADWT is applied only for the first decomposition level, for lower levels the classical lifting transform being used. The experimental results demonstrate significant improvements over the DWT employing the classical lifting [2], similar to those observed in the case of DADWT [5]. Furthermore, the experiments show that PSNR improvements of up to 0.35 dB over the block-based DADWT [5] are achieved with SD DADWT, if we do not consider the cost for coding the contour image C_0'. This indicates that adapting the bases both spatially and in terms of orientation might be a better option than adapting them only in terms of orientation. This also shows that SD DADWT is a better option than DADWT in applications that require performing both image segmentation tasks and compression.

Accounting for the contour coding costs affects the compression performance, particularly at low rates. However, the visual quality in the reconstructed image (Fig. 5) is still not affected. On the contrary, despite of the PSNR differences, some regions in the reconstructed image obtained via a segmentation-driven DADWT have a better visual quality that those obtained from the block-based DADWT (notice the highlighted regions in Fig. 5).

One can safely conclude that the proposed coding approach retains the advantage of encoding a multiresolution image segmentation, having potential applications in object-based coding and image analysis, and in the same time preserves significant performance advantages over an equivalent system employing the classical DWT.

6. CONCLUSIONS

In this paper we propose a multi-resolution image representation based on DADWT and segmentation of the image in non-rectangular regions. A scalable coding system employing the proposed SD DADWT is designed, and promising preliminary experimental results are provided.

It is important to observe that in contrast to the segmented image coding (SIC) approaches developed in the past [7], [8], synthesizing the image as a union of adjacent segments, the proposed SD DADWT is a global transform, making use of segmentation only in order *to drive* a global transformation process. This ensures that a coding system based on SD DADWT does not generate disturbing visual artefacts around the segment borders, as those observed in SIC coding at very low rates [7], [8].

One concludes also that the proposed SD DADWT offers a multiresolution representation of the geometric features in the image, via a multiresolution image segmentation, and in the same time provides a sparse image representation via directional lifting. These features make it beneficial in applications that require performing both image analysis and compression.

a) Block-based DADWT

b) Proposed SD DADWT

Figure 5. Encoded image at 0.3 bpp. The ellipses highlight the visual differences between (a) block-based and (b) SD DADWT.

7. ACKNOWLEDGMENTS

The authors would like to thank V. Christopoulos for the software development of the Edgmentation algorithm. P. Schelkens and A. Munteanu have postdoctoral fellowships with the FWO.

8. REFERENCES

[1] E. Le Pennec and S. Mallat, "Sparse Geometric Image Representations with Bandelets," *IEEE Trans. Image Proc.*, vol. 14, no. 4, pp. 423-438, April 2005.

[2] W. Sweldens, ‘The lifting scheme: A construction of second generation wavelets,” *SIAM J. Math. Analysis*, vol. 29, no. 2, pp. 511-546, 1998.

[3] W. Ding, F. Wu, and S. Li, “Lifting-based wavelet transform with directionally spatial prediction,” *Picture Coding Symposium*, pp. 483–488, 2004.

[4] Ö. N. Gerek and A. E. Çetin, “A 2-D orientation-adaptive prediction filter in lifting structures for image coding,” *IEEE Trans. Image Proc.*, vol. 15, pp. 106-111, 2006.

[5] C.L. Chang and B. Girod, “Direction-Adaptive discrete wavelet transform via directional lifting and bandeletization,” *Int. Conf. Image Proc.*, pp. 1149-52, 2006.

[6] R. Deklerck, J. Cornelis, M. Bister, “Segmentation of medical images”, *Image and Vision Computing*, vol. 11, no. 8, pp. 486-502, Oct. 1993.

[7] C. Christopoulos, W. Philips, A. Skodras, J. Cornelis, "Segmented Image Coding: techniques and experimental results", *Signal Processing: Image Communication*, vol. 11, pp.63-80, Nov. 1997.

[8] V. Christopoulos, "Segmented still image and video coding," PhD thesis, Vrije Universiteit Brussel, 1998.

[9] P. Schelkens, A. Munteanu, J. Barbarien, M. Galca, X. Giro-Nieto, J. Cornelis, "Wavelet coding of volumetric medical datasets," *IEEE Trans. Medical Imaging*, vol. 22, no. 3, pp. 441-458, March 2003.