Journal of Multimedia Processing and Technologies

Print ISSN: 0976-4127
Online ISSN: 0976-4135

DLINE JOURNALS

JMPT 2024; 15 (1)
https://doi.org/10.6025/jmpt/2024/15/1/9-16

Most Commonly used Directional Transforms in Image Coding

Ivo R. Draganov
Ivo R. Draganov is with the Faculty of Telecommunications
8 Kliment Ohridski Blvd., 1000 Sofia, Bulgaria
idraganov@tu-sofia.bg

ABSTRACT

In this paper, we attempt to review some of the most commonly used directional transforms in terms of their use in image coding, as well as some recommendations for improving their integration into full compression algorithms in the future. We compare DDCTs, DWHTs, DDWTs and others to each other based on image quality and execution time as well as memory efficiency when used for image compression. These transforms are even more efficient when used with some spectral coefficient rearrangement, as suggested for further discussion.

Received: 25 September 2023
Revised: 19 November 2023
Accepted: 29 November 2023

Keywords: Image Coding, Directional Transform, DDCT, DWHT, DDWT

1. Introduction

One of the most popular and widely used idea in image coding (compression) is the decorrelation of the intensity (color) values and possibly removing non-significant portions of the transformed data for which the human eye is less sensitive. Such classical techniques are the Discrete Cosine Transform (DCT) used in JPEG compressors and Discrete Wavelet Transform (DWT) in JPEG2000 algorithm [1]. Due to their separability along horizontal and vertical direction it is possible to apply them by using 1D masks reducing the amount of memory needed in one computational pass and reducing the complexity of the code structure, e.g. by using simpler look-up tables, etc.

Nevertheless of these enhancements there are large amount of cases where inside the coded image exist periodic structures with one and the same pattern alternating along an arbitrary direction different from horizontals and verticals, e.g. stripes, slopes, etc. In these cases it is appropriate to apply a transform in the same direction in which the dominant pattern spreads, that is a directional transform.

Two large groups can be defined for the existing directional transforms – the group of the linear orthogonal transforms [7-15, 17-22] and that of the directional wavelet transforms [2-6, 16]. Here a brief review is made for both groups revealing their basic principle, advantages and range of applicability in image coding. The rest of this paper is organized as the following – in section 2 a description of some of the most popular directional linear orthogonal transforms is given followed by directional wavelet transforms in section 3 and then in section 4 –conclusion is made.

2. Directional Linear Orthogonal Transforms

One of the simplest 2D-transforms is considered to be the Hadamard transform which is generalized in [10] by using the following jacket matrix:

$$K = \begin{bmatrix} 1 & 1 & \dots & 1 & 1 \\ 1 & * & \dots & * & \pm 1 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & * & \dots & * & \pm 1 \\ 1 & \pm 1 & \dots & \pm 1 & \pm 1 \end{bmatrix}, \quad (1)$$

which later is used by Monadjemi and Moallem [17] for texture classification where the case of sequency-ordered matrix of rank=3 (8x8 size) is applied to extract features from the image. They achieved classification accuracy of 90.5 % against 90.0 % for approach using the Gabor filter and only 77.5 % for the ordinary Walsh-Hadamard transform-like features. More than that, the execution time for the Directional Walsh-Hadamard Transform (DWHT) is more than 10 times less than for the Gabor filter which is considerable difference.

The more advanced transform – DCT – has been extended to its directional form by Zeng and Fu [8, 15] by modifying the weighting factors according to:

$$\hat{\alpha}(i) = \begin{cases} 1 / N_k, \; i = 0 \\ \sqrt{2 / N_k}, \; i \neq 0 \end{cases}, \; k = 0, 1, ..., 2N - 2, \quad (2)$$

where N is the size of the transform vector and k – the current number of the direction the transform is being applied. They also introduce a DC correction given in [15] by:

$$\Delta DC = \sum_{k=0}^{2N-2} \sqrt{N_k}\, B^Q(0, k) \Big/ \sum_{k=0}^{2N-2} \sqrt{N_k}, \quad (3)$$

where B denotes a column-vector with the spectral coefficients along the current direction.

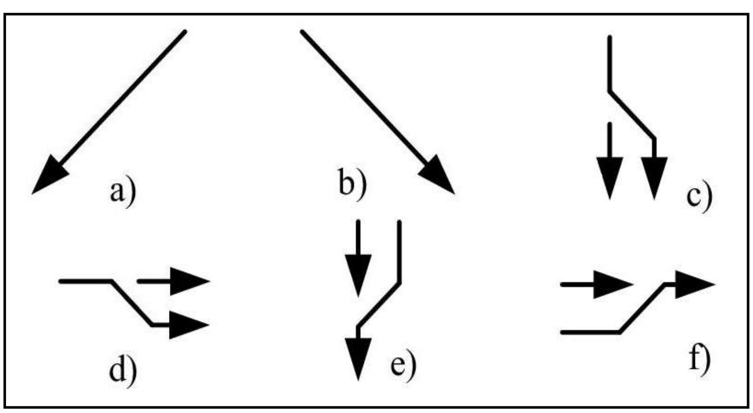

Figure 1. Different DDCT modes: a) Diagonal Down-Left, b) Diagonal Down-Right, c) Vertical-Right, d) Horizontal-Down, e) Vertical-Left, f) Horizontal-Up

Journal of Multimedia Processing and Technologies Volume 15 Number 1 March 2024

Depending on the direction type, e.g. diagonal down-left, diagonal down-right, etc. a certain amount of modes is defined, usually 8. In Figure 1 some of the possible modes are given by their basic direction. Combining the proper ones inside different images Zeng and Fu [15] obtain difference for the Peak Signal-to-Noise Ratio (PSNR) from 0.5 to 1.5 dB compared to that produce by the JPEG coder. This is true for static images compressed inside the range from 0.1 to 2 bpp where the actual PSNR changes between 31 and 42 dB and the transform block size is fixed to 8x8 pixels. Similar experiment is done for motion pictures where the comparison is made between the H.263 codec and its modified version using the DDCT. The compression ratios achieved are between 0.1 and 6 bpp for the different videos where PSNR varies between 28 and 45 dB. The advantage of using DDCT is obvious along all the selected range where the PSNR dominance is from 0.1 to 1.5 dB.

Some simplified form of the DDCT using only diagonal directions is fully described in [7] showing that it is actually optimal ortho normal transform because of the minimized object function maximizing smoothness. Also it is proven that no separability can be achieved here and thus it has no explicit functional form and any fast algorithm.

More productive attempts to speed up the DDCT are developed using lighting-based schemes. In [11] Xu et al. use factorization of 8-point DCT into 35 primary operations formalized in:

$$\vec{Y} = DCT(\vec{X}) = O_{35} \circ O_{34} \circ ... \circ O_2 \circ O_1 (\vec{X}), \qquad (4)$$

where $\vec{X}$ is the input vector and $\vec{Y}$ – the transformed one. There are two types primary operations – the first represents direct connection between two pixels along a direction of processing and is defined according [14] by:

$$O(X[n_i], X[n_i], \alpha) = \left\{ \vec{X} \leftarrow \begin{bmatrix} 1 & & & & & & & \\ & 1 & & & & & & \\ & & 1 & & & & & \\ & & & \beta & & & & \\ & & & & 1 & & & \\ & \gamma & & & & 1 & & \\ & & & & & & 1 & \\ & & & & & & & 1 \end{bmatrix} \vec{X} \right\} \qquad (5)$$

where β = 1 and γ = α ; the second type is when direct path could not be selected and scale parameter α is equal to β while γ = 0 representing the crossings over mid-lying pixels (Figure 2).

Experimental results based on that approach are presented in [11, 12, 14] concerning the quality comparison with the JPEG algorithm in the range from 0.5 to 2.5 bpp for lossy compression. The PSNR for static images changes generally from 28 to 42 dB. In virtually all the cases DDCT produces PSNR positive difference from 0.2 to 2.0 dB.

Drem eau et al. propose in [13] a new compression algorithm based on extended DDCT to rectangular bases and then use bin tree segmentation along with dynamic programming for optimal bases selection according to a rate distortion criterion. They compare their approach with the

JPEG and JPEG 2000 coders. The test images compression ratios achieved are between 0.01 and 2 bpp while the PSNR is changing between 15 and 44 dB. All along this range their approach proves to be better than both the other with between 1 and 5 dB.*

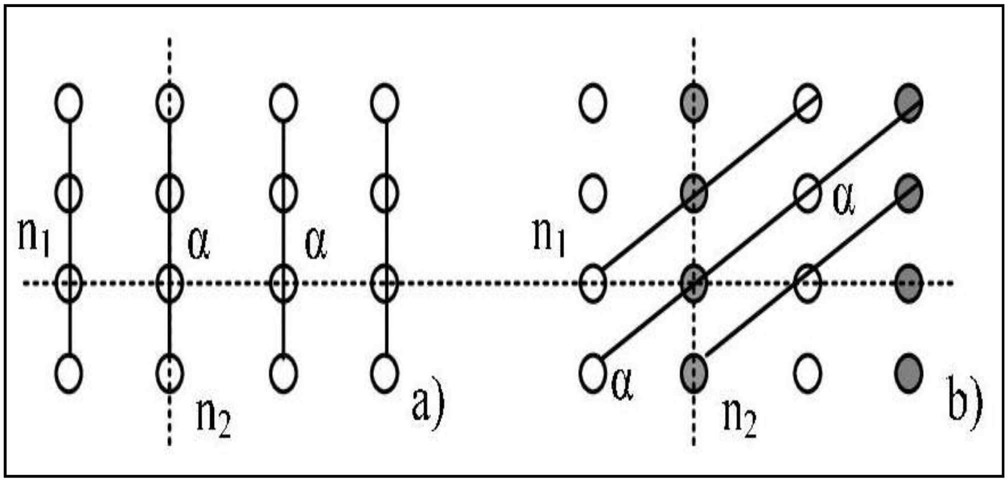

Figure 2. Two types of pixel connections determining the primary type operations in the Fast DDCT: a) direct and b) non-direct

Another approach employing the directional linear transforms is described in [9] for predicting visual residuals in a large number of cases – moving vehicles, people and other objects in city environment, landscapes, etc. The approach is tested with videos for both intra and inter coding and the results show enhancement as for the total PSNR with 1 to 7 dB in comparison to AVC coder. The direction in which the transform is going to be applied is found adaptively in each frame.

Some considerable expanding of the directional transforms is done when considering the challenges met in high resolution image coders design. In coders such as HD Photo overlapping is introduced. In [18] similar technique is undertaken in combination with DDCT and the achieved results are very promising. In a range from 0.1 to 7 bpp the PSNR benefit here is from 1 to 20 dB over the HD Photo itself which is remarkable result.

In [19] a hierarchical class structure is introduced for the I and B-frames from video processed by 1D directional unified transform along with bidirectional intra prediction. There are 5 classes – from A to E. The total rate gain only for the bidirectional prediction is 3.72 %, for the directional unified transform – it is 5.64 % and the cumulative effect for both is 8.76 %. This scheme is currently used by the Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T.

Some further development of the spatial prediction inside a video frame is done in [20] where such an optimal prediction is looked for along with an adaptive transform by Han et al. Here a hybrid transform is constructed alternating between sinusoidal transform regarding the frequencies and phases of the harmonics which are precisely defined by the boundary effects between adjacent blocks inside the frame. Inter-block correlations are exploited in an effective manner which is proven by the fact that PSNR is higher by 1 to 5 dB over the classical approaches using DCT for video in a range of 0.3 to 1.8 bpp compression ratios.

Some comparable results are presented in [21] and in [22] where sparse ortho normal transforms and direction-adaptive partitioned block transform are introduced.

3. Directional Wavelet Transforms

The second large group of directional transforms includes the wavelet ones. All of them are based on the classical wavelet transform realized by the lifting implementation (Figure 3) [6].

Journal of Multimedia Processing and Technologies Volume 15 Number 1 March 2024

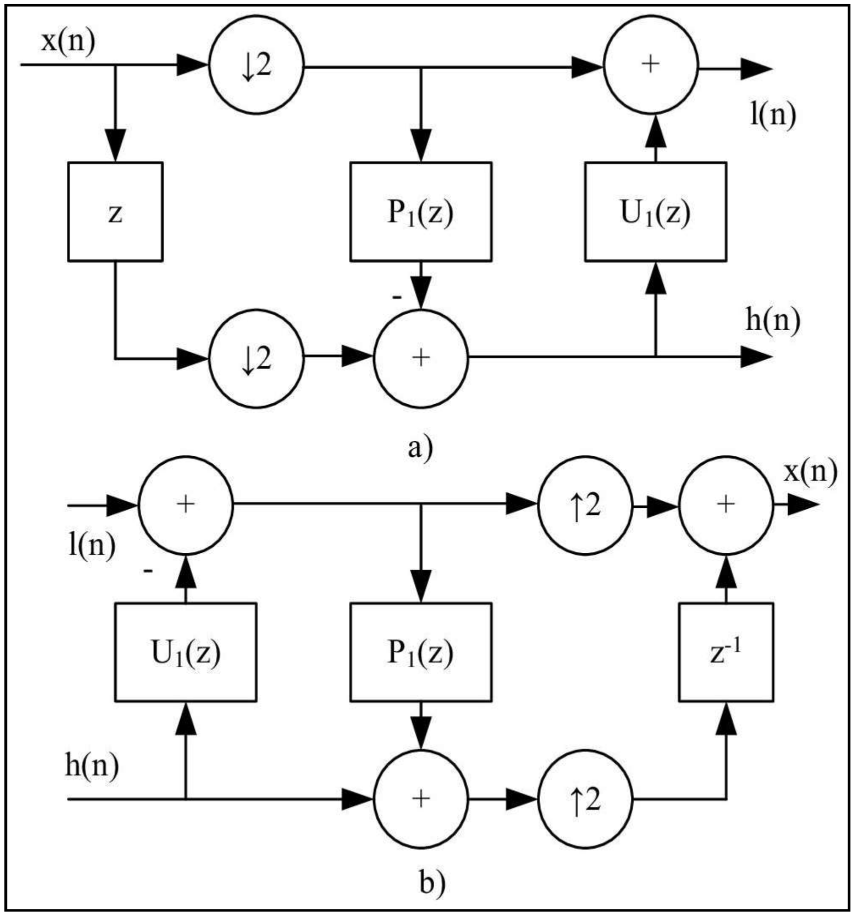

Figure 3. One-pass link for the lifting scheme of the wavelet transform: a) analysis filter bank and b) synthesis filter bank

The prediction P(z) and update U(z) filters are at the base of the analysis and synthesis filter banks of the DWT. One of the most often filters used in this process are the 9/7 biorthogonal wavelet filters described in [6] by:

$$\left|\begin{array}{l} P_1 = +1.58613(1+z^{-1}) \\ U_1(z) = -0.05298(1+z^{+1}) \\ P_2 = -0.88291(1+z^{-1}) \\ U_2(z) = +0.44350(1+z^{+1}) \\ s_1 = 1.23017 \quad s_2 = 1/s_1 \end{array}\right. \qquad (6)$$

Here the odd samples are predicted by two neighbouring even pixels which are first averaged and then the result is scaled while the even ones are predicted by averaging of two neighbouring odd pixels of the prediction residual.

In [2] an efficient embedded coding is suggested for medical image compression using the contour let transform. It is an extensive scheme which incorporates discrete contour let transform, laplacian pyramid, directional filter bank, some post noise removal steps, optionally extraction of region of interest (ROI), modified fuzzy C means segmentation, ROI based modified EZW algorithm ending *with modified Huffman coding. The authors managed to raise the compression ratio from almost 2 times at relatively constant PSNR around 33 dB for the test MRI images in comparison to the classic EZW approach and the SPIHT algorithm.*

Another approach is the robust adaptive directional lifting wavelet transform used for image denoising in [3] by Wang et al. The adaptive directional lifting is at the base of this method combining the directional spatial prediction and the conventional lifting scheme which removes the spatial redundancy leaved by the directional attributes. Additional novelty here is the classification at pixel level and the interscale correlation which assure more robustness of the orientation estimation algorithm. The transform itself is applied at pixel level and affecting only those pixels which belong to texture regions of interest. The PSNR for filtered images is increased by 7 dB for some typical for the practice cases.

In [4] some optimization is done for the directional lifting with reduced complexity. While the major disadvantage of the direction-adaptive discrete wavelet transform is the need of exhaustive search for the optimal prediction direction which makes it too complex in contrast to the classical DWT, here lowering of this complexity is aimed. Prediction of the optimal direction is done using gradient-based technique over a formal model of the prediction errors generated by the directional lifting of input wedge image. Stevens et al. [4] proved practically that the prediction step remains very simple and fast and the total complexity reduction has a factor of 11/4 preserving the prediction accuracy. The difference with the original test images and the coded ones with exhaustive search and with the optimized algorithm is about or less than 0.5 dB.

Similar algorithm to that described in [3] is presented in [5] by Chang and Girod. They use local adaptation of the filtering directions to the image content based on directional lifting. The advantage is that energy compaction is more for sharp image features. Additionally anisotropic statistical image model is created for quantifying the gain achieved by adapting the filtering directions. In such a way the authors claim that this algorithm is even more effective than similar ones developed earlier and gain of up to 2.5 dB for the PSNR is achieved. No loss of image structure is reported in the processed pictures.

Kamisli and Lim propose in [6] the directional wavelet transform to be used for prediction residuals in the video coding process. They clearly distinguish the coding of prediction residuals of frame intensities such as the motion compensation residual and the resolution enhancement residual. Special attention is dedicated to the specific characteristics of the different prediction residuals and how they differ from those of the image (frame) itself. Adapting the model for the directional transforms used then produce better results than to unify one and the same algorithm for both. The experiments carried out by Kamisli and Lim indicate that coefficient savings over the classical DWT are between 1 and 40 % with an average close to 30% and considerably more in some cases when DDWT model is adapted to the specific type of residual being processed.

4. Conclusion

In this paper a brief review is presented of the directional transforms used in image coding. They can be divided in two large groups consisting of linear orthogonal transforms and wavelet approaches respectively. A lot of combinations exist with other popular in practice techniques such as local and holistic decompositions, lifting schemes, groupings based on spatial and time correlation, etc. Depending on the application being developed all of them have their place in practice revealing even broader opportunities for future study and addendum. Some especially perspective approach is considered the hierarchical spectral multistage decomposition of the wavelet spectrum of an image where classical linear orthogonal and wavelet transforms are combined together. Substituting transforms represented of fixed matrix coefficients with adaptive ones such as Karhunen-Loev transform will produce even more efficient algorithms as for the quality and compression ratios of the coded images.

Acknowledgement

This paper was supported by the National Science Fund of the Bulgarian Ministry of Education, Youth and Science (Contract – DDVU 02/13 – "Public and Private Multimedia Network Throughput Increase by Creating Methods for Assessment, Control and Traffic Optimization").

References

[1] Rao, K. R., Yip, P. (2001). *The Transform and Data Compression Handbook*. CRC Press.

[2] Tamilarasi, M., Palanisamy, V. (2011). An Efficient Embedded Coding For Medical Image Compression Using Contourlet Transform. *European Journal of Scientific Research*, 49(3), 442-454.

[3] Wang, X., Shi, G., Niu, Y., Zhang, L. (2011). Robust adaptive directional lifting wavelet transform for image denoising. *IET Image Processing*, 5(3), 249-260.

[4] Stevens, R., Munteanu, A., Cornelis, J., Schelkens, P. (2008). Optimized Directional Lifting with Reduced Complexity. In *Proceedings of the 16th European Signal Processing Conference* (EUSIPCO 2008), Lausanne, Switzerland, pp. 1-5.

[5] Chang, C.-L., Girod, B. (2007). Direction-Adaptive Discrete Wavelet Transform for Image Compression. *IEEE Transactions on Image Processing*, 16(5), 1289-1302.

[6] Kamisli, F., Lim, J. (2009). Directional Wavelet Transforms for Prediction Residuals in Video Coding. In *Proceedings of the 16th IEEE International Conference on Image Processing (ICIP), Cairo*, Egypt, pp. 613-616.

[7] Selesnick, I., Guleryuz, O. (2011). A Diagonally-Oriented DCT-Like 2D Block Transform. *Proc. SPIE*, 8138.

[8] Fu, J., Zeng, B. (2007). Directional Discrete Cosine Transforms: A Theoretical Analysis. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2007)*, Honolulu, HI, USA, pp. I-1105 – I-1108.

[9] Cohen, R., Klomp, S., Vetro, A., Sun, H. (2010). Direction-Adaptive Transforms for Coding Prediction Residuals. In *Proceedings of the 17th IEEE International Conference on Image Processing*, Hong Kong, PRC, pp. 185-188.

[10] Horadam, K. (2005). A Generalised Hadamard Transform. In *Proceedings of the IEEE International Symposium on Information Theory*, Adelaide, Australia, pp. 1006-1008.

[11] Xu, H., Xu, J., Wu, F. (2007). Lifting-Based Directional DCT-Like Transform for Image Coding. In *Proceedings of the IEEE International Conference on Image Processing (ICIP 2007)*, Vol. 3, San Antonio, TX, USA, pp. III-185 – III-188.

[12] Chen, B., Wang, H., Cheng, L. (2010). Fast Directional Discrete Cosine Transform for Image Compression. *Optical Engineering*, 49(2), 020501-1 – 020501-3.

[13] Dr´emeau, A., Herzet, C., Guillemot, C., Fuchs, J.-J. (2010). Sparse Optimization with Directional DCT Bases for Image Compression. In *Proceedings of the IEEE Int'l Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Dallas, TX, USA, pp. 1290-1293.

[14] Xu, H., Xu, J., Wu, F. (2007). Lifting-Based Directional DCT-Like Transform for Image Coding. *IEEE Transactions on Circuits and Systems for Video Technology*, 17(10), 1325-1335.

[15] Zeng, B., Fu, J. (2006). Directional Discrete Cosine Transforms for Image Coding. In *Proceedings of the IEEE International Conference on Multimedia and Expo (ICME 2006)*, Toronto, Ontario, Canada, pp. 721-724.

[16] Monadjemi, A. (2004). *Towards Efficient Texture Classification and Abnormality Detection*. PhD Thesis, University of Bristol, UK.

[17] Monadjemi, S., Moallem, P. (2006). Texture Classification Using a Novel Walsh/Hadamard Transform. In *Proceedings of the 10th WSEAS International Conference on COMPUTERS*, Vouliagmeni, Athens, Greece, pp. 1002-1007.

[18] Xu, J., Wu, F., Liang, J., Zhang, W. (2010). Directional Lapped Transforms for Image Coding. *IEEE Transactions on Image Processing*, 19(1), 85-97.

[19] Tanizawa, A., Yamaguchi, J., Shiodera, T., Chujoh, T., Yamakage, T. (2010). Improvement of Intra Coding by Bidirectional Intra Prediction and 1 Dimensional Directional Unified Transform. Input Document to JCT-VC, JCTVC-B042, 2nd Meeting: Geneva, CH.

[20] Han, J., Saxena, A., Rose, K. (2010). Towards Jointly Optimal Spatial Prediction and Adaptive Transform in Video/Image Coding. In *Proceedings of the IEEE Int'l Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Dallas, TX, USA, pp. 726-729.

[21] Sezer, O., Harmanciy, O., Guleryuz, O. (2008). Sparse Orthonormal Transforms for Image Compression. In *Proceedings of the 15th IEEE International Conference on Image Processing (ICIP 2008)*, San Diego, CA, USA, pp. 149-152.

[22] Chang, C.-L., Makar, M., Tsai, S., Girod, B. (2010). Direction-Adaptive Partitioned Block Transform for Color Image Coding. *IEEE Transactions on Image Processing*, 19(7), 1740-1755.

Journal of Multimedia Processing and Technologies Volume 15 Number 1 March 2024