Texture description using Dual Tree Complex Wavelet Packets

M. Liedlgruber[1], M. Häfner[2], J. Hämmerle-Uhl[1], and A. Uhl[1]

[1] Visual Computing and Security Lab (VISEL)
Department of Computer Sciences, University of Salzburg, Austria
[2] Department for Internal Medicine
St. Elisabeth Hospital, Vienna, Austria
andreas.uhl@sbg.ac.at

Abstract. In this work we extend several DWT-based wavelet and wavelet packet feature extraction methods to use the dual-tree complex wavelet transform. This way we aim at alleviating shortcomings of the different algorithms which stem from the use of the underlying DWT.
We show that, while some methods benefit significantly from extending them to be based in the dual-tree complex wavelet transform domain (and also provide the best overall results), for other methods there is almost no impact of this extension.

1 Introduction

In the past, various different wavelet-based feature extraction methods have already been successfully applied to the problem of texture classification. However, since most of the approaches are based on the discrete wavelet transform (DWT) or the discrete wavelet packets transform (DWPT), they also inherit two major shortcomings inherent to the DWT when used for image processing and classification. First, the DWT is not able to capture directional information. Second, the DWT lacks shift-invariance.

An extension to the pyramidal DWT, which aims at coping with these problems, is the dual-tree complex wavelet transform (DT-CWT) [1]. In the past it has already been shown that features based on the DT-CWT are able to deliver superior classification results as compared to the pyramidal DWT and other feature extraction methods [2].

To exploit the benefits of the DT-CWT we extend a set of wavelet-packet-based feature extraction approaches originally defined to be used with the DWT to use the DT-CWT (resulting in a DT-CWPT). We then use these methods to extract features from two different image databases and investigate the classification performances as compared to the original (i.e. non-complex) version of the algorithms.

The remaining part of this work is organized as follows: in Section 2 we provide a rough overview of the DT-CWT and the complex wavelet packets transform (DT-CWPT). We then describe the methods evaluated and their extension to the complex domain in Section 3. Details on the experimental setup

used are given in Section 4, followed by the results obtained in Section 5. We conclude this work in Section 6.

2 Background

2.1 Dual-Tree Complex Wavelet Transform

To overcome the limitations of the DWT the original DT-CWT uses 2^D pyramidal DWTs for a D-dimensional transform (i.e. for a 2-D DT-CWT four DWTs are needed). The outcomes of these transforms are then combined to obtain six complex-valued subbands in the 2-D case. These subbands capture image details at $\pm 15°, \pm 45°$, and $\pm 75°$. In addition, the DT-CWT is approximately shift-invariant.

The complex-valued wavelet, which forms the basis for the D-dimensional DT-CWT, can be expressed as

$$\psi_c(t) = \psi_h(t) + j\psi_g(t), \tag{1}$$

where $\psi_h(t)$ and $j\psi_g(t)$ are the real and imaginary part of the complex wavelet, respectively. The complex scaling function can be defined analogously as

$$\phi_c(t) = \phi_h(t) + j\phi_g(t). \tag{2}$$

Based on the separable implementation of the 2D-DWT the complex equivalent can be written as:

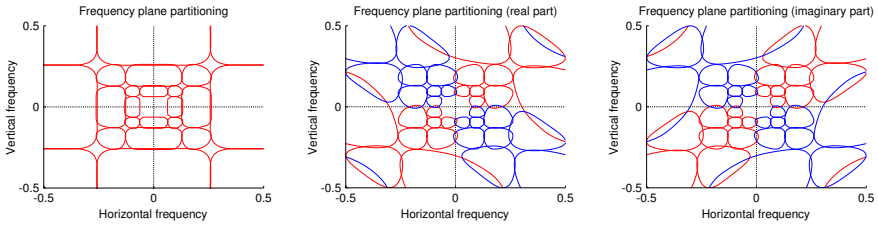
$$\psi^{(-)}_{\mathrm{LH}}(x,y) = \phi_c(x)\psi_c(y) \tag{3}$$
$$\psi^{(-)}_{\mathrm{HL}}(x,y) = \psi_c(x)\phi_c(y) \tag{4}$$
$$\psi^{(-)}_{\mathrm{HH}}(x,y) = \psi_c(x)\psi_c(y) \tag{5}$$
$$\psi^{(+)}_{\mathrm{LH}}(x,y) = \phi_c(x)\overline{\psi_c(y)} \tag{6}$$
$$\psi^{(+)}_{\mathrm{HL}}(x,y) = \psi_c(x)\overline{\phi_c(y)} \tag{7}$$
$$\psi^{(+)}_{\mathrm{HH}}(x,y) = \psi_c(x)\overline{\psi_c(y)}. \tag{8}$$

The four DWTs needed for the DT-CWT can now be easily developed by substituting Equs. 1 and 2 into Equs. 3 to 5 and computing the real and imaginary parts of the results. This way we obtain the DWTs needed for the negative orientations (i.e. $-15°, -45°$, and $-75°$). To obtain the DWTs for the positive angles (i.e. $+15°, +45°$, and $+75°$) the same computations must be carried out based on Equs. 6 to 8.

For the DT-CWT we need two filter banks for the first stage and two filter banks for the the remaining stages. These filter banks are then combined in all possible ways to obtain the filter banks for the four DWTs needed (i.e. different filters are used for the row- and column-wise transform of the separable 2-D transform).

We refer to the first stage real part filters as h_0 (low-pass) and h_1 (high-pass). The imaginary part filters for the first stage are denoted by g_0 and g_1. The remaining stages filters are denoted by $h_0^\bullet, h_1^\bullet, g_0^\bullet$, and $g_1^\bullet$. To get an analytic wavelet (i.e. ability to capture directional information), h_1 and g_1 must form a

Hilbert transform pair (the same applies to $h_1^\bullet$ and $g_1^\bullet$). In addition, to obtain a shift-invariant transform, h_0 and g_0 must meet the requirement of a one-sample shift between them. The remaining stages low-pass filters $h_0^\bullet$ and $g_0^\bullet$ must have a half-sample shift between them. Fig. 1 shows the different 2-D frequency partitionings produced by the DWT and the DT-CWT. We notice that the DWT affects all four quadrants equally. Contrasting, the DT-CWT produces differently oriented parts which enables the DT-CWT to differentiate between more directions of details (i.e. two neighboring quadrants are colored differently).



(a) DWT (b) DT-CWT

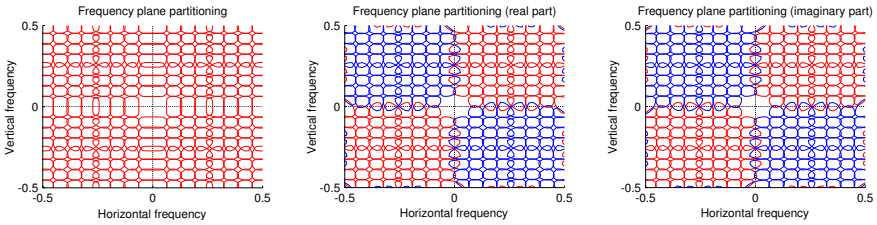
Fig. 1. A comparison of the 2-D frequency plane partitioning between (a) the DWT and (b) the real and imaginary part of the DT-CWT. The plots show the 70%-peak magnitude as contour lines for the first three levels of decomposition (different directions, i.e. positive and negative angles in case of the DT-CWT, are shown in red and blue, respectively).

2.2 Dual-Tree Complex Wavet Packets Transform

In order to extend methods which rely on a DWPT, we need a full wavelet packet transform based on the DT-CWT. Unfortunately, analyticity gets lost for deeper decomposition levels if the DT-CWT is just extended to decompose the high-frequency subbands too [3]. But it has already been shown that a solution to this problem can be obtained fairly easily [3, 4]. We decided to use the method proposed in [3] due to its simplicity when it comes to integrate it into an existing DT-CWT implementation.

As already pointed out above, in the pyramidal case of the DT-CWT four different filter combinations are used among the four DWTs. The solution proposed in [3] is quite simple: to retain analyticity even for deeper levels of the transform, the filters used must remain the same across the different DWTs for most nodes in the decomposition tree for decomposition levels greater two. For more details on this extension we refer the reader to [3].

Fig. 2 shows the different 2-D frequency partitionings produced by the DWPT and the DT-CWPT. Similar to the DWT, the DWPT affects all four quadrants equally. Contrasting, the DT-CWPT produces differently oriented parts.



(a) DWPT (b) DT-CWPT

Fig. 2. A comparison of the 2-D frequency plane partitioning between (a) the DWPT and (b) the real and imaginary part of the DT-CWPT. The plots show the 70%-peak magnitude as contour lines for the first three levels of decomposition.

3 Evaluated Feature Extraction Methods and Their Extensions

3.1 DWT-based and DWPT-based Methods

The following methods are either based on the real DWT or the real DWPT:

WPC: Each image is decomposed using the DWT. The features to be classified are then extracted based on coefficients in the resulting high-frequency subbands.

WT-BB [5]: Using the DWPT and the Best-basis algorithm [6], each image is decomposed into an optimal basis with respect to a cost function (based on the coefficients in the resulting subbands). The features are then extracted from all resulting subbands, ignoring the approximation subband.

Since it is very likely that the optimal bases differ among different images, the features would not be comparable directly in a meaningful way. As a consequence, a feature vector is filled with zeros at positions which correspond to subbands which are not present in that feature vector but which are present in at least one image decomposition structure in the remaining images from the image set. This way we end up with feature vectors which are comparable since each position in a feature vector then corresponds to a certain subband.

WT-BBCB [5]: This method also relies on the DWPT and the Best-basis algorithm. Hence, each image is decomposed into an optimal basis. Considering the decomposition trees for the resulting bases, the decomposition tree which on average is most similar to all other decomposition trees is searched for (we call this tree the centroid). To compute the similarity between two decomposition trees we employ the quadtree distance metric used in [5].

Once the centroid has been found, all images are decomposed into the respective basis. The features are then extracted from the resulting subbands, ignoring the approximation subband.

WT-LDB [7]: Using the Local discriminant bases algorithm [8], which is based on the DWPT, an optimal basis with respect to the discriminant power

of subbands among different image classes is computed. The discriminant power is similar to the cost function used in the Best-basis algorithm.

Once the optimal basis has been found, all images are decomposed into the respective basis and the features are extracted from the resulting subbands, except for the approximation subband.

3.2 Extended Complex Methods

In case of WPC, extending the method to employ the DT-CWT is straightforward, since the originally proposed DT-CWT yields a pyramidal transform already. There are just two differences:

- the DT-CWT yields complex coefficients. To extract features from the complex subbands we simply compute the coefficient magnitudes and compute the features from these.
- we obtain six high-frequency subbands instead of three, which doubles the feature vector lengths.

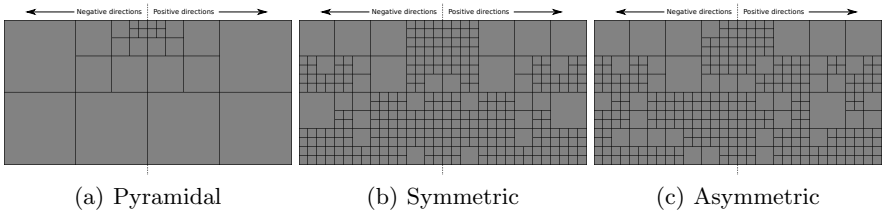
For methods, which rely on the DWPT, we can use the complex variant DT-CWPT, outlined in Section 2.2. Similar to WPC, we again extract the features based on the magnitudes of the complex coefficients. However, in the complex case we obtain subbands for features oriented at negative and positive angles, which can be considered to be two separate decomposition trees (these are the same in case of the DT-CWT). As a consequence we evaluate two different ways to extract features in the complex case when dealing with the DT-CWPT:

- **Symmetric case**: For the positive and negative directions the same decomposition structures are used. For each subband position considered for feature extraction features are extracted for both directions and stored in an interleaved fashion.
  In case an adaptive method is used to find an optimal basis, the pruning process inherent to the basis finding methods computes the cost for a decomposition tree node based on the sum of costs for the two directions.
- **Asymmetric case**: The decompositions for positive and negative directions are allowed to be different. As a consequence the two decomposition trees are pruned separately. The features are then extracted separately for both decomposition trees and concatenated to obtain the final feature vectors.
  Due to the fact that each subband in a complex decomposition depends on all four DWPTs involved, the trees must be kept synchronous to be able to perform an inverse transform. As a consequence we always perform a full DWPT on an image in the asymmetric case and keep the coefficients for all nodes at each decomposition stage. The pruning process is then not a traditional pruning based on an inverse transform but we merely mark nodes which should be included in the final basis for each direction (i.e. we end up with two different bases for the different directions).

Since in case of the WPC method the decomposition structures are always pyramidal ones, asymmetric trees are not used for this method. For all other methods hope is raised that using asymmetric decomposition trees allows the decomposition to adapt better to the characteristics of an image.

To extend the WT-LDB method to a complex one, the so-called time frequency energy map (TFEM) used in the Local discriminant bases algorithm must also be extended. Simply spoken the TFEM contains the mean energy at a certain coefficient position across all images of a class. Hence, we obtain one TFEM for each image class. To extend the TFEM for our needs, we simply compute two different TFEMs for each class, one for each direction. Then, depending on which decomposition tree is pruned, the right TFEM is used (in the symmetric case both TFEMs are used to compute the sum of discriminant powers for each node).

Fig. 3 shows the difference between the pyramidal DT-CWT, the symmetric case, and the asymmetric case evaluated for feature extraction.



(a) Pyramidal (b) Symmetric (c) Asymmetric

Fig. 3. A comparison of the decomposition trees between (a) the pyramidal DT-CWT, (b) an example symmetric decomposition structure, and (c) an example asymmetric decomposition structure (in both cases a decomposition depth of 4 has been used).

4 Experimental Setup

4.1 Image Databases Used

Kylberg Database (KB-DB [9]): The original Kylberg database consists of 28 image classes, each containing 160 grayscale images with a size of 576×576 pixels. The balanced nature of this database (i.e. same number of images in each class) and the high number of total images (4480) allowed us to split the database into two separate, equally-sized sets for training and validation (each containing 2240 images). The database split has been done by using 80 images from each class for training (Kylberg sample names *c* and *d*) and the remaining 80 images for validation (Kylberg sample names *a* and *b*). In order to reduce the computation time, especially in case of time-consuming feature extraction methods, we modified the image set by extracting center patches of size 128×128 pixels from the original images. The cropped images are then used for the experiments.

High-magnification Colonic Polyp Database (HM-DB) [2]: This image database is based on 327 endoscopic color images (either of size 624×533 pixels or 586×502 pixels) acquired between the years 2005 and 2009 at the Department of Gastroenterology and Hepatology (Medical University of Vienna) using a zoom-colonoscope (Olympus Evis Exera CF-Q160ZI/L) with a magnification factor of 150. In order to acquire the images 40 patients underwent colonoscopy. To obtain a larger set of images we manually extracted subimages (regions of interest) with a size of 256×256 pixels from the original images. This resulted in an extended image set containing 716 images in total.

Lesions found during colonoscopy have been examined after application of dye-spraying with indigocarmine, as routinely performed in colonoscopy. Biopsies or mucosal resection have been performed in order to get a histopathological diagnosis.

Details on the endoscopic image database used are provided in Table 1. In these tables the columns $\mathbf{N_O}$, $\mathbf{N_E}$, and $\mathbf{N_P}$ denote the number of original images (i.e. the source images for patch extraction), the number of extracted patches, and the number of patients, respectively. From table 1 we notice that the total number of patients given is slightly higher as compared to the number of patients who underwent endoscopy for the respective databases. The reason for this is that in case of some patients different types of pathologies showed up across the patient images. As a consequence a patient may be contained in more than one class.

In case of the colonic polyp database we distinguish between a 2-classes case and a 3-classes case. In the former we simply distinguish between normal mucosa (non-neoplastic) and mucosal changes which need a medical intervention (neoplastic). A more fine-grained classification was proposed in [10]. In this classification scheme the images are divided into three classes: normal lesions, non-invasive lesions, and invasive lesions. This classification scheme is of particular importance since normal mucosa needs not to be removed, non-invasive lesions must be removed endoscopically, and invasive lesions must not be removed endoscopically.

Table 1. The detailed ground truth information for HM-DB.

Image Class	**3 classes**			**2 classes**		
	$\mathbf{N_O}$	$\mathbf{N_E}$	$\mathbf{N_P}$	$\mathbf{N_O}$	$\mathbf{N_E}$	$\mathbf{N_P}$
Normal	72	198	14	72	198	14
Non-Invasive	212	420	27	255	518	32
Invasive	43	98	6			
Total	**327**	**716**	**47**	**327**	**716**	**46**

4.2 Wavelet-Transform Setup

In case of the complex methods we use Kingsbury's Q-Shift (14,14)-tap filters (for decomposition stages ≥ 2) in combination with (13,19)-tap near-orthogonal filters (for the first decomposition stage). For the DT-CWPT we use the Q-Shift filters for the decomposition nodes needing special treatment (using the methodology proposed in [3]). To make the real and complex methods more comparable, we use the Q-Shift filters also for the real methods (i.e. just one of the filter banks).

For methods which are based on the Best-basis algorithm we use the entropy as cost function. For the Local discriminant bases algorithm we use the l^2-norm as discriminant measure.

4.3 Feature Extraction and Classification

The feature we use for all methods is the entropy which is computed from the coefficients (magnitude) in the high-frequency subbands.

To reduce the dimensionality of the feature vectors and to improve the comparability of the techniques we perform a principal component analysis (PCA). Prior to applying the PCA to the features, we center the training feature vectors by subtracting the feature-wise mean from each feature. Then, after computing the eigenvalues and eigenvectors for a given set of training feature vectors, the eigenvalues and eigenvectors are sorted in descending manner with respect to the eigenvalues. This is followed by computing the number of components p to retain from the cumulative sum of the eigenvectors, such that the cumulative sum for the first p largest eigenvalues is above 0.99.

Once the validation features have been extracted and centered (using the means from the original training features), the feature projection computed from the training features is also applied to the validation features.

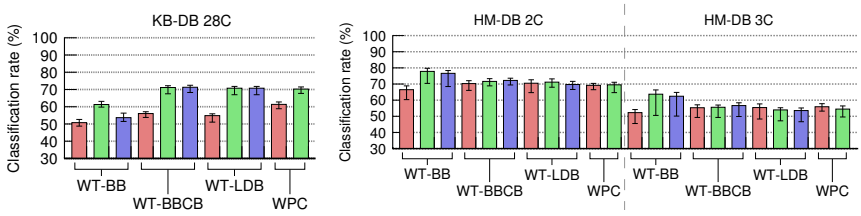
For the classification we use the k-Nearest neighbors (k-NN) classifier using the l^1-norm to compute the distances between feature vectors. This rather weak classifier has been chosen to emphasize more on the effect of extending the feature extraction methods to complex ones. We carried out experiments with different values for k (i.e. $k = 1, \ldots, 25$) and present the average results.

To estimate the classification accuracies in case of the colonic polyp database we use the leave-one-patient-out cross-validation (LOPO-CV).

5 Results

Fig. 4 provides an overview of the overall classification rates we obtained in our experiments. The red, green, and blue bars denote the mean overall rates for real, complex, and complex asymmetric versions of the respective methods (over all choices for the k-value from the k-NN classifier). On top of each bar the we also indicate the range of classification rates over all choices for k.

As can be seen from Fig. 4(a), switching to the complex domain consistently improves the mean overall classification rates by up to 10% in case of KB-DB.



(a) KB-DB (b) HM-DB

Fig. 4. An overview of the results from our experiments.

Only WT-BBCB is able to slightly benefit from switching from the complex symmetric mode to the complex asymmetric mode. WT-BB delivers the lowest rates for KB-DB when compared to the other methods.

We notice from Fig. 4(b) that switching to the complex domain only improves the mean overall classification rates for WT-BB in case of HM-DB. In the 3-classes case we even observe a slight results drop slightly for WT-LDB and WPC. However, we also notice that, while the result range over all choices for k is rather small for all other methods, it is wider in case of WT-BB. When comparing the symmetric and asymmetric modes we again notice (similar to the results of the KB-DB) that only in case of WT-BBCB an improvement can be observed (although it is only a minor one). In all other cases switching to the asymmetric mode leads to a small result drop.

It is interesting to note that for the regular textures of the KB-DB the comlex version of the WT-BBCB approach gives the best results (which employs a single decomposition structure to all images subject to classification – still this is adaptively chosen), while for the less regular (only texture-like) images of the endoscopical HM-DB the complex version of the WT-BB technique provides superior results, for which each image is potentially decomposed into a different decomposition structure. It seems that for the less homogeneous imagery the higher adaptivity potential of the WT-BB approach is beneficial.

6 Conclusion

The results obtained show that at least one method (WT-BB) is able to consistently improve the classification rates when using the complex version of it. For all other methods there is a dependency on the data set: For the KB-DB, the complex version improves the results for all considered feature extraction technique, partially significantly so. For the endoscopic HM-DB, we see improvments for WT-BB only, thus confirming results on earlier investigations on high-definition endoscopic data [11].

However, it also turned out that, at least for most of the methods evaluated, it is sufficient to use complex symmetric decomposition trees. Only WT-BBCB is able to consistently improve the classification rates when using the complex asymmetric version.

7 Acknowledgments

This work has been supported by the Austrian Science Fund (FWF) under Project No. TRP-206.

References

[1] Selesnick, I.W., Baraniuk, R.G., Kingsbury, N.G.: The dual-tree complex wavelet transform - a coherent framework for multiscale signal and image processing. IEEE Signal Processing Magazine **22**(6) (November 2005) 123–151

[2] Häfner, M., Kwitt, R., Uhl, A., Gangl, A., Wrba, F., Vécsei, A.: Feature-extraction from multi-directional multi-resolution image transformations for the classification of zoom-endoscopy images. Pattern Analysis and Applications **12**(4) (December 2009) 407–413

[3] Bayram, İ., Selesnick, I.W.: On the dual-tree complex wavelet packet and m-band transforms. IEEE Transactions on Signal Processing **56**(6) (jun 2008)

[4] Weickert, T., Kiencke, U.: Analytic wavelet packets - combining the dual-tree approach with wavelet packets for signal analysis and filtering. IEEE Transactions on Signal Processing **57**(2) (feb 2009)

[5] Liedlgruber, M., Uhl, A.: Statistical and structural wavelet packet features for pit pattern classification in zoom-endoscopic colon images. In Dondon, P., Mladenov, V., Impedovo, S., Cepisca, S., eds.: Proceedings of the 7th WSEAS International Conference on Wavelet Analysis & Multirate Systems (WAMUS'07), Arcachon, France (October 2007) 147–152

[6] Coifman, R.R., Wickerhauser, M.V.: Entropy based methods for best basis selection. IEEE Transactions on Information Theory **38**(2) (March 1992) 719–746

[7] Häfner, M., Liedlgruber, M., Wrba, F., Gangl, A., Vécsei, A., Uhl, A.: Pit pattern classification of zoom-endoscopic colon images using wavelet texture features. In Sandham, W., Hamilton, D., James, C., eds.: Proceedings of the International Conference on Advances in Medical Signal and Image Processing (MEDSIP'06), Glasgow, Scotland, UK (July 2006) 1–4

[8] Saito, N., Coifman, R.R.: Local discriminant bases. In: SPIE's 1994 International Symposium on Optics, Imaging, and Instrumentation, International Society for Optics and Photonics (1994) 2–14

[9] Kylberg, G.: The kylberg texture dataset v. 1.0. External report (Blue series) 35, Centre for Image Analysis, Swedish University of Agricultural Sciences and Uppsala University, Uppsala, Sweden (September 2011)

[10] Kato, S., Fu, K.I., Sano, Y., Fujii, T., Saito, Y., Matsuda, T., Koba, I., Yoshida, S., Fujimori, T.: Magnifying colonoscopy as a non-biopsy technique for differential diagnosis of non-neoplastic and neoplastic lesions. World Journal of Gastroenterology **12**(9) (March 2006) 1416–1420

[11] Häfner, M., Liedlgruber, M., Uhl, A.: Colonic polyp classification in high- definition video using complex wavelet-packets. In: Proceedings of Bildverarbeitung für die Medizin 2015 (BVM'15). (March 2015) 365–370