Hill, P. R., Bull, D., & Canagarajah, C. N. (2000). Rotationally invariant texture features using the dual-tree complex wavelet transform. 901 - 904. 10.1109/ICIP.2000.899602

Link to published version (if available):
10.1109/ICIP.2000.899602

ROTATIONALLY INVARIANT TEXTURE FEATURES USING THE DUAL-TREE COMPLEX WAVELET TRANSFORM

P. R. Hill, D.R. Bull and C.N. Canagarajah
Image Communications Group, Centre for Communications Research, University of Bristol,
Merchant Ventures Building, Woodland Road, Bristol, BS8 1UB, UK
Tel: +44 117 9545125, Fax: +44 117 9545206
Email: Paul.Hill@bristol.ac.uk

ABSTRACT

New rotationally invariant texture feature extraction methods are introduced that utilise the dual tree complex wavelet transform (DT-CWT). The complex wavelet transform is a new technique that uses a dual tree of wavelet filters to obtain the real and imaginary parts of complex wavelet coefficients. When applied in two dimensions the DT-CWT produces shift invariant orientated subbands. Both isotropic and anisotropic rotationally invariant features can be extracted from the energies of these subbands. Using simple minimum distance classifiers, the classification performance of the proposed feature extraction methods were tested with rotated sample textures. The anisotropic features gave the best classification results for the rotated texture tests, outperforming a similar method using a real wavelet decomposition.

1. INTRODUCTION

Efficient content based retrieval of images and video is ultimately dependent on the features used for data annotation. Recently developed texture based features have proved to be one of the effective descriptions of content. Spatial-frequency analysis techniques using Gabor filters and wavelets have provided good characterisation of textures in controlled environments. However, in order to better characterise textures, extracted features must capture the nature of the texture invariant to rotational, shift and scale transformations.

The focus of the work presented here is rotational invariance of texture features. This can be classified into two types: isotropic and anisotropic. Features extracted with isotropic rotational invariance represent averaged measures from annular frequency regions. Anisotropic rotational invariance features also contain measures from annular frequency regions but also represent the angular distribution of frequency content.

Initial attempts to produce isotropic and anisotropic rotationally invariant texture features from wavelet decompositions used the steerable pyramid transform [2]. This transform has the disadvantage of being considerably overcomplete with the amount of overcompletenes increasing with the number of analysed orientations. Classic dyadic wavelet decompositions have also been used to produce rotationally invariant features. Such features have been extracted from simple combinations of subband measures [3] as well as from hidden Markov models used to model rotation variations in the wavelet output [4]. The use of such decompositions has the disadvantage of lacking directional selectivity. However, Wu and Wei [5] used a spiral resampling lattice before using a similar dyadic wavelet packet decomposition to produce rotationally invariant features.

Non-separable wavelets have been implemented to obtain more flexible isotropic and anisotropic rotationally invariant features [6]. Similar features have been extracted using Gabor filters (the G-let) implemented in the Fourier domain [3] to produce isotropic rotationally invariant features. The disadvantage with these methods is the complexity associated with using a frequency decomposition (FFT or similar) before analysis.

The dual tree complex wavelet transform has already been shown to provide good results for unrotated texture classification using a wavelet packet type decomposition [9]. In this paper we use this decomposition to extract efficient isotropic and anisotropic rotationally invariant features. This is made possible because the DT-CWT decomposition gives good directional selectivity whilst remaining computationally efficient and does not require a resampling lattice.

2. DUAL TREE COMPLEX WAVELET TRANSFORM

The DT-CWT is a spatial frequency transform that uses spatial filters to decompose an image or image region into dyadic subbands similarly to the classic dyadic wavelet transform.

0-7803-6297-7/00/$10.00 © 2000 IEEE

Shift invariance can be achieved in a dyadic wavelet transform by doubling the sampling rate. This is effected in the DT-CWT by eliminating the down sampling by 2 after the first level of filtering. Two parallel fully decimated trees are constructed by placing the downsampled outputs of first level filters of tree one sample offset from the outputs of the other. To get uniform intervals between the two trees' samples, the subsequent filters in one tree must have delays that are displaced by one half sample. For linear phase, this is enforced if the filters in one tree are of even length and the filters in the other are of odd length. Additionally, better symmetry is achieved if each tree uses odd and even filters alternatively from level to level. The filters are chosen from a perfect reconstruction biorthogonal set and the impulse responses can be considered as the real and imaginary parts of a complex wavelet [1].

Application to images is achieved by separable complex filtering in two dimensions. The 2:1 redundancy in one dimension translates to 4:1 redundancy in two dimensions with the output from each filter and its conjugate forming six orientated subbands at each scale. Complex wavelets are able to separate positive and negative frequencies thus differentiating and splitting the subbands of a dyadic decomposition into subbands orientated at ±15°, ±45°, ±75° as shown in figure 1.

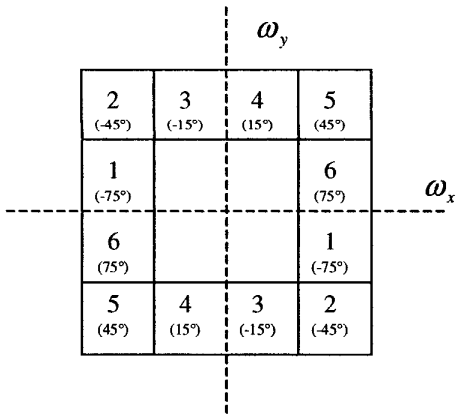

Figure 1: *Frequency plane showing 6 orientated subbands of the complex wavelet output*

3. EXTRACTION OF ISOTROPIC AND ANISOTROPIC ROTATIONALLY INVARIANT FEATURES FROM THE DT-CWT

In the subsequently described experiments, the texture image regions are decomposed into six bandpass orientated subbands at each scale with the low-low subbands being recursively decomposed for three levels. This gives 6×3 orientated subbands with two residual low-low pass images. i.e. 20 subbands in all. Channel energies were extracted from each subband using the L_1 norm:

$$e_k = \frac{1}{N^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left|x_k(i,j)\right| \qquad (1)$$

where e_k is the energy for the k^{th} subband of dimension $N\times N$ with coefficients $x_k(i,j)$.

3.1. Isotropic Rotationally Invariant Features

In a similar scheme to that produced by Porter for the DWT [3], isotropic rotationally invariant features are produced by summing the energies from each of the subbands at each scale. As the subbands at ±45° were judged to be at significantly different radial frequencies than the rest, an alternative feature set was constructed that omitted them from the summations. Both cases gave feature vectors of length 4.

3.2. Anisotropic Rotationally Invariant Features

At each scale anisotropic features were extracted by using the discrete Fourier transform. If f_q represents the 6 orientated channel energy values at a particular scale then the DFT is given by:

$$\hat{f}_k = \sum_{q=0}^{5} f_q e^{-i\pi qk/3} \quad for\ k = 0,1,2,3,4,5 \qquad (2)$$

The zeroth harmonic, $\hat{f}_0$, is just the DC summation. The magnitudes of $\hat{f}_1$, $\hat{f}_2$ and $\hat{f}_3$ (the first, second and third harmonic) can be used as anisotropic features at each scale or can be combined into single features. Coefficients above the third harmonic ($\hat{f}_4$, $\hat{f}_5$) are above the Nyquist limit and therefore not useful. Greenspan et al. [7] have used a similar analysis technique with the steerable pyramid and rotated filters. Additional anisotropic invariant features using autocorrelation measures of these subband energies have been developed by Hill et al. [8].

4. EXPERIMENTAL RESULTS

Sixteen textures were taken from the Brodatz texture album to test the classification performance of the developed features. These textures are shown in figure 2 and were chosen to represent textures that contained a range of periodic, stochastic and directional elements. The textures were scanned as eight-bit raw grey level images of size 256×256 pixels.

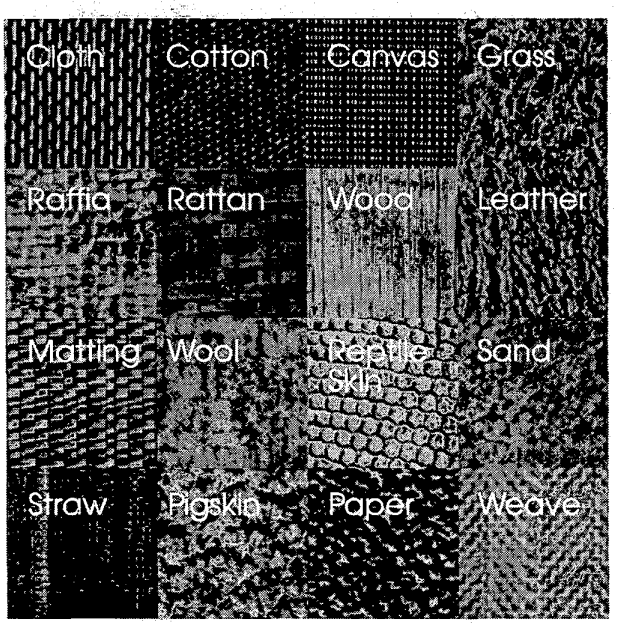

Figure 2: 16 Brodatz textures used in texture classification experiments

Many different configurations of feature vectors are possible from the DT-CWT and the rotational Fourier analysis described above. The following feature vectors were tested in the experiments resulting in tables 1 and 2.

1. Average of the 6 subband channel energies at each scale. **[4 features]**
2. Average of 4 subbands channel energies (i.e. no ±45° subbands) at each scale. **[4 features]**
3. Magnitudes of $\hat{f}_0$ and $\hat{f}_1$ for each scale. **[7 features]**
4. Magnitudes of $\hat{f}_0$ and $\hat{f}_2$ for each scale. **[7 features]**
5. Magnitudes of $\hat{f}_0$ and $\hat{f}_3$ for each scale. **[7 features]**
6. Magnitudes of $\hat{f}_0$ for each scale. Each of $\hat{f}_1$, $\hat{f}_2$ and $\hat{f}_3$ averaged over all the scales. **[7 features]**
7. Magnitudes of $\hat{f}_0$, $\hat{f}_1$ and $\hat{f}_2$ at each scale.**[10 features]**
8. Magnitudes of $\hat{f}_0$, $\hat{f}_1$, $\hat{f}_2$ and $\hat{f}_3$ for each scale. **[13 features]**

Of course all configurations include an energy measure for the residual low-low channel values.

One version of each texture class was used for training at angles of 0°, 30°, 45° and 60°. Seven different versions of each texture were used for classification and presented at angles 20°, 70°, 90°, 120°, 135° and 150°. This gave 42 classifications per texture and 672 in all. Each training texture was tiled into 16×16 squares with feature values being extracted from the complex wavelet decomposition of each tile leading to a mean vector and covariance matrix for each texture class. The four angles of training were used to enable the mean feature vector and the covariance matrix to be properly estimated under texture rotation. Similarly, the mean feature vectors were extracted from the test textures from the complex wavelet transform of tiled 16×16 squares. Textures were classified using a minimum Mahalanobis distance classifier.

A 9-7 biorthogonal wavelet pair was used as the odd filters and a 6-2 biorthogonal wavelet pair was used for the even filters. These were chosen because they are linear phase, approximately matched to give shift invariance and more spatially localised than the filters (13-19 & 12-16) used by Kingsbury in [1]. Good spatial localisation can be important for texture analysis and when using such small analysis areas (16×16) can minimise edge effects. No better results were obtained with the larger filters developed by Kingsbury in [1] even when decomposing using the entire texture image.

Table 1 shows the correct classification results of the best feature sets using decompositions on areas of 16×16 pixels. Inclusion of the 45° subbands gave the best results for the isotropic rotationally invariant feature vectors (i.e. feature vector 1). The best results for the anisotropic rotationally invariant feature vectors was achieved with the full 13-length feature vector (i.e. feature vector 8). For comparison, the best correct classification rate with the same data for the isotropic rotationally invariant features extracted from a DWT as developed by Porter [10] are shown. Table 2 shows the correct classification results of the best feature vectors using wavelet decompositions over the entire images for both training and classification.

Feature Extraction Technique	No. of features	Correct Classification Rate (%)
Complex Wavelet 1: Sum of 6 channels at each scale	4	91.30
Complex Wavelet 8: $\hat{f}_0$, $\hat{f}_1$, $\hat{f}_2$ and $\hat{f}_3$ for each scale	13	93.75
Real wavelet [10] Summation of channel energies at each scale	4	87.35

***Table 1:** Classification performance of wavelet features on rotated images: decomposition on 16×16 areas*

Feature Extraction Technique	No. of features	Correct Classification Rate (%)
Complex Wavelet 1: Sum of 6 channels at each scale	4	81.99
Complex Wavelet 8: $\hat{f}_0$, $\hat{f}_1$, $\hat{f}_2$ and $\hat{f}_3$ for each scale	13	90.33
Real wavelet [10] Summation of channel energies at each scale	4	73.21

Table 2: *Classification performance of wavelet features on rotated images: decomposition on whole images*

The features extracted using 16×16 areas provided better classification results. This is likely to be because the variation in covariance distribution of the features was better estimated and therefore the Mahalanobis distance a better measure. However, in cases where the smallest repeating element was larger than this area, larger or entire image decompositions would be preferred.

5. CONCLUSION

The ability of the DT-CWT to distinguish between positive and negative frequencies results in six orientated subbands at each scale when it is applied in two dimensions. A discrete Fourier transform of these subband energies results in a harmonic representation of the angular frequency content. This is not only rotationally invariant but characterises the angular frequency distribution i.e. anisotropic rotational invariance.

The Classification performance in the conducted tests of a feature vector formed from rotational harmonics extracted from a DT-CWT decomposition was over 5% better than a similar method based on a real wavelet transform. Although less well matched to produce shift invariance the adopted filters produced identical or better classification results as those used by Kingsbury in [1] whilst being more spatially localised.

This method is considerably less complex than previous attempts at producing anisotropic rotationally invariant features [7]. The complexity of the method is roughly equivalent to four single two dimensional wavelet transforms. Although this is a significant increase in complexity over the normal DWT it still represents less complexity than a 2D-FFT decomposition for the same size of image.

ACKNOWLEDGEMENTS

This work was supported by the Virtual Centre of Excellence in Digital Broadcast and Multimedia Technology. The authors acknowledge the support and information provided by Dr N.G. Kingsbury of Cambridge University.

REFERENCES

[1] N.G. Kingsbury, "The dual-tree complex wavelet transform: a new technique for shift invariance and directional filters", *IEEE Digital Signal Processing Workshop:86, DSP 98, Bryce Canyon,* August 1998

[2] E. P. Simoncelli, W.T. Freeman, "The Steerable Pyramid: A Flexible Architecture for Multi-Scale Derivative Computation", *IEEE International Conference on Image Processing,* October 1995

[3] R. Porter, "Texture Classification and Segmentation", *PhD Thesis, University of Bristol,* November 1997

[4] J-L Chen and A. Kundu, "Rotation and Gray Scale Transform Invariant Identification Using Wavelet Decomposition and Hidden Markov Model", *PAMI*, Vol. 16, No. 2, February 1994

[5] W-R Wu and S-C Wei, "Rotation and Gray-Scale Transform-Invariant Texture Classification Using Spiral Resampling, Subband Decomposition, and Hidden Markov Model", *IEEE trans. on image processing*, Vol 5, No. 10, October 1996

[6] P. Vautrot, G. Van De Wouwer, P. Scheunders, S. Livens and D. Van Dyck, "Non-Separable Wavelets for Rotation-Invariant Texture Classification and Segmentation", *PAMI*, Vol 8, No 4, pp.472-481, July 1998

[7] H. Greenspan, S. Belongie, R. Goodman and P. Perone, "Overcomplete Steerable Pyramid Filters and Rotation Invariance" *PAMI, pp.* 222-228, June 1994

[8] P.R. Hill, C.N. Canagarajah and D.R. Bull, "Rotationally Invariant Texture Classification", *IEE Seminar on Time-Scale, Time-Freq Analysis and Applications*, pp 20/1-20/5, February 2000

[9] S. Hatipoglu, S.K. Mitra and N.G. Kingsbury, "Texture Classification using Dual-Tree Complex Wavelet Transform", *IEE 7th Intl Conf. Image Processing and it's Applications,* 1999.

[10] R. Porter and C.N. Canagarajah, "Rotation Invariant Texture Classification Schemes Using GMRFs and Wavelets," *Proceedings of the International Workshop on Image and Signal Processing*, pp.183-186, November 1996