IMAGE SEGMENTATION USING INVARIANT TEXTURE FEATURES FROM THE DOUBLE DYADIC DUAL-TREE COMPLEX WAVELET TRANSFORM

Edward H. S. Lo[†]*, Mark R. Pickering, Michael R. Frater and John F. Arnold*

School of Information Technology and Electrical Engineering
University College, The University of New South Wales
Australian Defence Force Academy, Canberra, Australia
{edward.lo, m.pickering, m.frater, j.arnold}@adfa.edu.au

ABSTRACT

In this paper we propose a new texture segmentation technique that produces segmentation results which more closely match the manual segmentation that would be performed by a human operator. To perform this type of segmentation, we propose a new texture feature based on the Double Dyadic Dual-Tree Complex Wavelet Transform (D^3T-CWT) which provides the ability to analyse a signal at and between dyadic scales. This new texture feature is invariant to shift, rotation and scale and hence can group the texture features in a single object (which may have different sizes and orientations) into a single more meaningful segment. When compared with other texture segmentation approaches, the proposed approach provides segmentation results which more closely match the semantically meaningful objects in the scene.

Index Terms— Complex wavelets, texture segmentation, scale invariance, rotation invariance

1. INTRODUCTION

When searching for and retrieving images from a digital database, a query is often performed using low level texture descriptors obtained from a query image. We argue that it would be more meaningful to use descriptors that can capture the entire texture in a semantically meaningful object. In this way the query engine could perform the search based on the actual objects segmented from the query image rather than low level textures which may not represent the semantic content of the image.

In this paper we propose a new texture segmentation technique that produces segmentation results which more closely matches the manual segmentation that would be performed by a human operator. A human operator is more likely to segment an image into semantically meaningful objects rather than image regions that are composed of identical texture information. Semantically meaningful objects are often composed of textures at different scales and rotations e.g. zebra stripes shown in Fig 2(a). Standard texture segmentation typically involves matching a texture exactly and is evaluated using texture patches that contain a single texture at the same scale and rotation. This type of segmentation would over-segment the zebra into regions containing stripes at different rotations and scales. We argue that, in order to segment the zebras from the background, it is necessary to use feature vectors that can capture the concept of “stripes” and is able to discriminate between stripes and other textures. Because the zebra stripes can be of any width and at any orientation we argue that the feature vector must be invariant to scale and rotation.

Research into the human visual system suggests that vision is based on responses of cells with a characteristic similar to Gabor wavelets [1, 2]. As such, Gabor wavelets have played an important role in image processing as a representation for texture in the area of image segmentation [3, 4], representation for texture images in digital libraries [5, 6] and are considered as a texture descriptor for MPEG-7 [7]. Manthalkar, Biswas and Chatterji describe a way of generating scale and rotation invariant 2D Gabor based features by combining the Discrete Fourier Transform (DFT) with Gabor wavelets [8]. In this paper we propose a similar methodology but base our technique on the Dual-Tree Complex Wavelet Transform.

2. THE DOUBLE DYADIC DUAL-TREE COMPLEX WAVELET TRANSFORM

The Dual-Tree Complex Wavelet Transform (DT-CWT), developed by Kingsbury, efficiently implements a complex wavelet transform using dual trees of filters that independently generate the real and imaginary responses [9-11]. In 2D, the DT-CWT has a Gabor-like response with six directional subbands spaced 30° apart. In this paper, we will denote the DT-CWT coefficients for image analysis as $\zeta_l(x, y, d)$ for level $l = 1, 2, \ldots, L$, subband $d = 1, \ldots, 6$ and

[†] Employed by Defence Science & Technology Organisation, Department of Defence, Canberra, ACT, 2600, Australia

1-4244-0728-1/07/$20.00 ©2007 IEEE

decimated location (x, y). The energy of the DT-CWT coefficients is defined as $|\zeta_l(x, y, d)|^2$ and abbreviated as $|\zeta|^2$.

The Double Dyadic DT-CWT (D^3T-CWT) was first proposed by the authors to extend the DT-CWT to allow it to analyse a signal at and between dyadic scales [12]. The ability to analyse a signal between dyadic scales is achieved by applying the same DT-CWT on a resized version of the original signal. This novel approach tunes the resulting subband frequencies to lie midway between dyadic scales using a signal resize factor of $2^{-0.5}$.

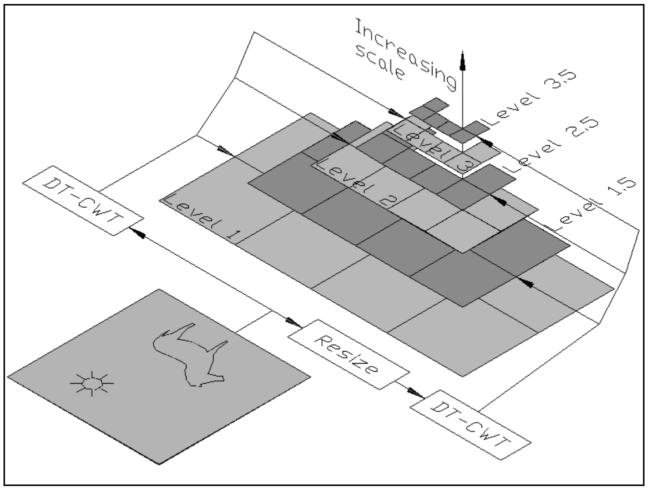

Fig 1: Decomposing an image using the D^3T-CWT.

In this paper, the coefficients from the D^3T-CWT is denoted as $\Psi_l(x, y, d)$ for level $l = 1, 1.5, \ldots, L+0.5$, subband $d = 1, \ldots, 6$ and decimated location (x, y). The energy of the D^3T-CWT coefficients is defined as $|\Psi_l(x, y, d)|^2$ and abbreviated as $|\Psi|^2$. The coefficients in Ψ are formed by interleaving the results of the DT-CWT decompositions from both execution branches based on increasing scale. Fig 1 describes this process for an image.

3. SCALE AND ROTATION INVARIANT TEXTURE FEATURE

Kingsbury's DT-CWT has proven to be a useful tool for analysing signals particularly in image processing. It is shift invariant and multi-scale texture features derived from the DT-CWT were shown by Kam to be useful for image segmentation [13]. His implementation involved:

1. Generate $|\zeta|^2$ from DT-CWT decomposition of image
2. Obtain the logarithm of all coefficients in $|\zeta|^2$

In this approach, the compact representation of the pyramid structure of the DT-CWT provides an advantage over the less efficient representation of texture using Gabor wavelets.

3.1. Rotation Invariant Texture Feature

Hill, Bull and Canagarajah describe how to produce a rotation invariant texture feature from the DT-CWT [6]. They achieve this by applying the Discrete Fourier Transform (DFT) around rotation dimension on the subband energies from the 2D DT-CWT coefficients and taking its magnitude.

We extend this idea to the individual coefficient at each level in $|\zeta|^2$ and $|\Psi|^2$ to generate spatially distinct rotation invariant texture features that relate to pixels in an image. For each level l and decimated location (x, y), there are six coefficients, u_d (where $d = 1, \ldots, 6$) from each subband in either $|\zeta|^2$ or $|\Psi|^2$. The DFT generates U_m (for $m = 1, \ldots, 6$) from u_d via,

$$U_m = \sum_{d=1}^{6} u_d e^{-i2\pi(d-1)(m-1)/6}.$$

The set of values, $f_{RI} = \{|U_1|, \ldots, |U_6|\}$ provides a rotation invariant texture feature. Since $U_2 = U_6^*$ and $U_3 = U_5^*$ (complex conjugates), f_{RI} at each level l and location (x, y) can be compacted down to $\{|U_1|, \ldots, |U_4|\}$. The results generate pyramid F_{RI}, consisting of rotation invariant features f_{RI}.

3.2. Scale Invariant Texture Feature

In its original form, the DT-CWT does not provide a good basis for generating scale invariant texture features because the response profile in scale dimension varies with scale. The proposed D^3T-CWT overcomes this problem by using additional overlapping filters between dyadic scales. The extra coefficients provide additional information to the response profile over the scale dimension. The new response profile then has the desirable property that a change in scale in the input signal corresponds to a shift in scale dimension of the $|\Psi|^2$ response profile.

We propose that scale invariant features can be approximated from the D^3T-CWT by applying the DFT across scale dimension of $|\Psi|^2$ and then taking the magnitude of the result. For each subband d (where $d = 1, \ldots, 6$) and decimated location (x, y) that spatially relates to pixel p in the image, we can extract the coefficients, v_l (where $l = 1, \ldots, 2L$) from each level in $|\Psi|^2$. The DFT generates V_n (for $n = 1, \ldots, 2L$) from v_l via,

$$V_n = \sum_{l=1}^{2L} v_l e^{-i2\pi(l-1)(n-1)/6}.$$

The values, $f_{SI} = \{|V_1|, \ldots, |V_{2L}|\}$ provides a texture feature that is approximately invariant to scale. The DFT assumes that the input signal repeats itself forever yet there is no reason why the fine and coarse ends of the response profiles should connect together. However if the response profile is close to zero at both ends, the approach provides good scale invariance. Due to complex conjugates, f_{SI} at each subband d and location (x, y) that relates to pixel p can be compacted down to $\{|V_1|, \ldots, |V_{L+1}|\}$. The results generate matrix F_{SI}, consisting of scale invariant features, f_{SI}.

3.3. Scale and Rotation Invariant Texture Feature

In this paper, we generate a scale and rotation invariant texture feature by merging the individual processes of generating rotation invariant features and approximately scale invariant features, described above. The details of the algorithm are described in pseudo code as follows:

1. Decompose image with the D^{3}T-CWT and form $|\Psi|^2$
2. Generate rotation invariant texture features f_{RI} from $|\Psi|^2$ (described above) and store in F_{RI}
3. Apply the process to generate scale invariant features over F_{RI} (instead of $|\Psi|^2$) to generate scale and rotation invariant texture f_{SIRI} and store results in matrix F_{SIRI}

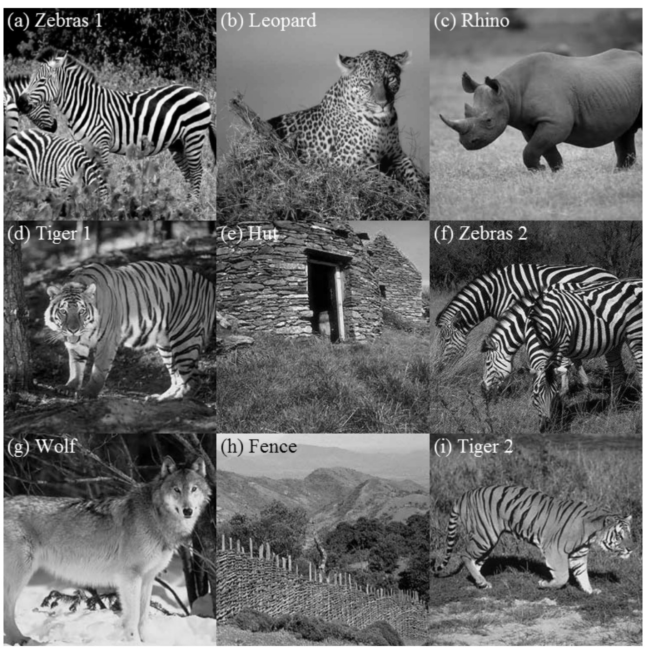

Fig 2: Test images used in this experiment.

4. EXPERIMENTAL RESULTS

Our experiment involves analysing the performance of our scale and rotation invariant texture feature derived from the D^{3}T-CWT in comparison to other popular measures of texture. The test images used in this paper are shown in Fig 2 while the texture features used are:

1. Energy of Gabor wavelet coefficients with just touching dyadic spaced filters, by Manjunath [5]
2. Texture features from $|\zeta|^2$ (DT-CWT), by Kam [13]
3. Rotation invariant texture features, f_{RI} from $|\zeta|^2$ (DT-CWT), extended from Hill [6]
4. Scale & rotation invariant texture features f_{SIRI} from $|\Psi|^2$ (D^{3}T-CWT) proposed in this paper [12]

The traditional approach for considering multi-scale texture information in image segmentation has been to use a multi-scale approach to classification such as with the hierarchical Markov Random Field. In our experiment we purposely use a non-hierarchical based classification technique to demonstrate the benefit of deriving a texture feature invariant to differences in scale and rotation. The segmentation process used in this paper is described as:

1. Logarithm of texture features extracted from image
2. Dimension reduction of the feature from part 1 into two dimensions using Principal Component Analysis
3. Add spatial component to form spatial-texture feature
4. Automatic classification of normalized feature vectors using the Mean Shift Procedure (MSP)
5. Post processing using mode based filtering

The first stage involves extracting texture features from the image and taking the logarithm of the results. The second stage of our process involves using Principal Component Analysis (PCA) to dimensionally reduce the result from the first stage [14]. The third stage generates the spatial-texture feature from the second stage result. The Mean Shift Procedure (MSP) is an unsupervised classification technique that seeks modes in the feature space and considers them to be cluster centres. Modes are defined as locations in feature space consisting of high feature density and modes are iteratively found by moving in the direction of slope specified by the density gradient estimate [15]. The version of MSP used in this paper is based on the technique proposed by Comaniciu and Meer [16]. To fairly compare our approach, we directly use the values of the parameters as chosen by Kam for the MSP [17]. The final post processing stage involving mode based filtering is used to remove small segments.

TABLE I AUTOMATIC IMAGE SEGMENTATION RESULTS COMPARING OUR TEXTURE FEATURE WITH OTHER TECHNIQUES.

	Gabor	**Kam**	**RI**	**SIRI**
Zebras 1	69%	64%	73%	**88%**
Leopard	**83%**	73%	72%	66%
Rhino	62%	**68%**	61%	**68%**
Tiger 1	54%	48%	**62%**	61%
Hut	76%	88%	**92%**	90%
Zebras 2	47%	55%	65%	**82%**
Wolf	**46%**	44%	40%	42%
Fence	52%	**74%**	67%	64%
Tiger 2	**67%**	64%	65%	**67%**
Average	62%	64%	66%	**70%**

% correct pixels compared to ground truth & best result highlighted in bold.

Table I shows the results of automatic image segmentation conducted over the input images. These results are presented as the percentage of pixels labelled correctly when compared to the manually segmented ground truth data. From Table I we can see that the results of segmentation using the scale and rotation invariant features from the D^3T-CWT (in the final column) compare well against the compared features.

The segmentation of Zebras 1 is shown in Fig 3. These results demonstrate the capability of the scale and rotation invariant features from the D^3T-CWT to extract the entire region of zebra stripe texture as a single segment.

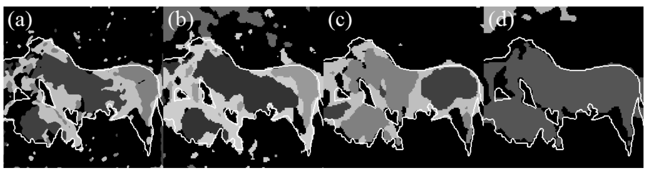

Fig 3: Segmentation of Zebras1 using (a) Gabor, (b) Kam, (c) rotation invariant, (d) scale & rotation invariant texture feature.

The segmentation of Tiger 2 is shown in Fig 4. Similar performance is achieved by all methods in extracting the tiger from the foreground and background. While it is true for most cases, in some, a scaled texture may not necessarily be part of the same object. It is interesting to note that in Fig 4(d) the shadow below the tiger has been associated with tiger stripes.

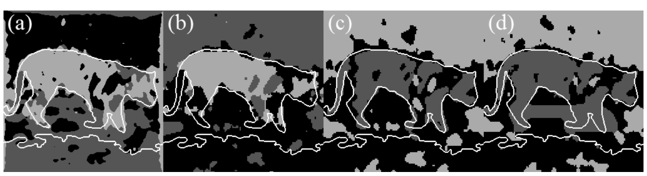

Fig 4: Segmentation of Tiger 2 using (a) Gabor, (b) Kam, (c) rotation invariant, (d) scale & rotation invariant texture feature.

5. CONCLUSIONS

Overall, using the scale and rotation invariant texture features from the energy of the D^3T-CWT coefficients produces segmentation results that more closely match the manual segmentation than when using other texture features. In instances where our technique performs poorly, all the other techniques perform just as inaccurately. This would indicate that another feature, such as colour, is required to segment those images well.

When compared with other texture segmentation approaches, the proposed approach can provide segmentation results which more closely match the semantically meaningful objects in the scene.

6. REFERENCES

[1] D. Gabor, "Theory of Communication," *J. Inst. Electrical Engineers,* vol. 93, pp. 429-441, November 1946.

[2] T. S. Lee, "Image Representation using 2D Gabor Wavelets," *IEEE Trans. Pattern Analysis & Machine Intelligence,* vol. 18, pp. 959-971, October 1996.

[3] A. K. Jain and F. Farrokhnia, "Unsupervised Texture Segmentation using Gabor Filters," *Pattern Recognition,* vol. 24, pp. 1167-1186, 1991.

[4] N. Zhang and W. K. Leow, "Perceptually Consistent Segmentation of Texture using Multiple Channel Filter," in *Proc. Asian Conf. Comp. Vision* Hong Kong, China: Hong Kong University of Science & Technology, 1998, pp. 17-24.

[5] B. S. Manjunath and W. Y. Ma, "Texture Features for Browsing and Retrieval of Image Data," *IEEE Trans. Pattern Analysis & Machine Intelligence,* vol. 18, pp. 837-842, August 1996.

[6] P. R. Hill, D. R. Bull, and C. N. Canagarajah, "Rotationally Invariant Texture Features using the Dual-Tree Complex Wavelet Transform," in *Proc. Int'l Conf. Image Processing*. vol. 3 Vancouver, Canada: IEEE, 2000, pp. 901-904.

[7] B. S. Manjunath, J. R. Ohm, V. V. Vasudevan, and A. Yamada, "Colour and Texture Descriptors," *IEEE Trans. Circuits & Systems for Video Technology,* vol. 11, June 2001.

[8] R. Manthalkar, P. K. Biswas, and B. N. Chatterji, "Rotation and Scale Invariant Texture Classification using Gabor Wavelets," in *Proc. Int'l w/s Texture Analysis & Synthesis* Copenhagen, Denmark: Heriot-Watt Uni., 2002, pp. 87-90.

[9] N. Kingsbury, "Image Processing with Complex Wavelets," *Phil. Trans. R. Soc. A,* vol. 357, pp. 2543-2560, 1999.

[10] N. Kingsbury, "Complex Wavelets for Shift Invariant Analysis and Filtering of Signals," *Applied & Computational Harmonic Analysis,* vol. 10, pp. 234-253, 2001.

[11] I. W. Selesnick, R. G. Baraniuk, and N. G. Kingsbury, "The Dual-Tree Complex Wavelet Transform," *IEEE Signal Processing Magazine,* vol. 22, pp. 123-151, November 2005.

[12] E. H. S. Lo, M. Pickering, M. Frater, and J. Arnold, "Scale and Rotation Invariant Texture Features from the Dual-Tree Complex Wavelet Transform," in *Proc. Int'l Conf. Image Processing*. vol. 1 Singapore: IEEE, 2004, pp. 227-230.

[13] A. H. Kam, T. T. Ng, N. G. Kingsbury, and W. J. Fitzgerald, "Content Based Image Retrival Through Object Extraction and Querying," in *Proc. IEEE Workshop Content-Based Access of Image & Video Libraries* Hilton Head Island, USA: IEEE, 2000, pp. 91-95.

[14] A. K. Jain, R. P. W. Duin, and J. Mao, "Statistical Pattern Recognition: A Review," *IEEE Trans. Pattern Analysis & Machine Intelligence,* vol. 22, pp. 4-37, January 2000.

[15] K. Fukunaga and L. D. Hostetler, "The Estimation of the Gradient of a Density Function, with Applications in Pattern Recognition," *IEEE Trans. Information Theory,* vol. 21, pp. 32-40, 1975.

[16] D. Comaniciu and P. Meer, "Distribution Free Decomposition of Multivariate Data," *Pattern Analysis & Applications,* vol. 2, pp. 22-30, 1999.

[17] A. H. Kam, "A General Multiscale Scheme for Unsupervised Image Segmentation," in *Dept. Engineering* Cambridge, England: University of Cambridge, 2000.