ELSEVIER

Preserving boundaries for image texture segmentation using grey level co-occurring probabilities

Rishi Jobanputra, David A. Clausi*

Department of Systems Design Engineering, 200 University Avenue West, Waterloo, ON, Canada, N2L 3G1

Abstract

Texture analysis has been used extensively in the computer-assisted interpretation of digital imagery. A popular texture feature extraction approach is the grey level co-occurrence probability (GLCP) method. Most investigations consider the use of the GLCP texture features for classification purposes only, and do not address segmentation performance. Specifically, for segmentation, the pixels in an image located near texture boundaries have a tendency to be misclassified. Boundary preservation when using the GLCP texture features for image segmentation is important. An advancement which exploits spatial relationships has been implemented. The generated features are referred to as weighted GLCP (WGLCP) texture features. In addition, an investigation for selecting suitable GLCP parameters for improved boundary preservation is presented. From the tests, WGLCP features provide improved boundary preservation and segmentation accuracy at a computational cost. As well, the GLCP correlation statistical parameter should not be used when segmenting images with high contrast texture boundaries.

© 2005 Pattern Recognition Society. Published by Elsevier Ltd. All rights reserved.

1. Introduction

Texture is a significant property in digital imagery. It has an important role in human visual perception and offers information for recognition and interpretation. Consequentially, considerable research effort has been focussed on texture analysis in recent years. In general, unsupervised image texture segmentation involves (1) extracting each pixel's texture features and (2) clustering like pixels based on their features. Selecting texture features which are independent and discriminable will aid in the segmentation process. Tuceryan and Jain categorize the various texture analysis methods into four groups: statistical, geometrical, model-based and signal processing-based [1]. Here, a particular statistical method is studied for its ability to perform texture image segmentation.

The grey level co-occurrence probability (GLCP) method is a second-order statistical texture analysis approach [2]. The GLCPs describe the probability of any grey level occurring spatially relative to any other grey level, within a given image window. From this distribution, many parameters are applied to generate features.

Texture analysis studies generally focus on image classification [3–7], rather than segmentation. Classification studies are concerned with identifying regions of pure texture samples within an image, and do not focus on interpreting the image in its entirety. As a result, classification studies do not address performance issues that occur in regions with a mixture of classes. In addition, most applications of computer vision are concerned with identifying an entire scene; hence, classification studies are not truly indicative of real-world problems.

Texture features tend to misclassify and erode texture boundaries during segmentation [8]. "When the input image contains a large number of textures, the boundary

* Corresponding author. Tel.: +1 519 888 4567x2604;
fax: +1 519 746 4791.

E-mail address: dclausi@engmail.uwaterloo.ca (D.A. Clausi).

regions between textures are more likely to form new classes during the clustering procedure. Furthermore, many texture categories together with the boundary classes may overlap in the feature space [9, p. 179]". Boundary misclassification becomes especially problematic when using larger window sizes and for images with irregularly shaped texture boundaries [10]. However, when using the GLCP method, large window sizes are necessary to gather sufficient data to characterize local texture regions; small window sizes will result in poorly sampled co-occurring probabilities and will produce incoherent features. This paper will present an innovation to the GLCP technique which will improve boundary accuracy for image segmentation applications. This method is referred to as the weighted GLCP (WGLCP) method.

Research involving the selection of proper parameters for the co-occurrence method when applied to image segmentation is unknown. Many studies consider parameter selection based on classification of pure samples [3–7], but recommendations from these papers may not be suitable for the segmentation case. In order to properly use the co-occurrence method for image segmentation, care must be made in terms of proper parameter selection. Results from the classification studies can be used as a starting point for such an investigation.

This paper is arranged in the following manner. Section 2 gives an overview on how to generate the GLCP and WGLCP texture features. Section 3 investigates the individual texture statistics' suitability for boundary response performance. Section 4 is a comparative segmentation study and conclusions are presented in Section 5.

2. Texture feature generation

2.1. Grey level co-occurrence texture features

The co-occurrence method provides a second-order approach for generating texture features [2]. Given a spatial window within the image, the co-occurrence method finds the conditional joint probabilities, C_{ij}, of all pairwise combinations of grey levels given the inter-pixel displacement vector (δ_x, δ_y), which represents the separation of the pixel pairs in the x- and y-directions respectively. The set of grey level co-occurring probabilities (GLCP) can be defined as:

$$C_{ij} = \frac{P_{ij}}{\sum_{i,j=0}^{G-1} P_{ij}}, \tag{1}$$

where P_{ij} represents the frequency of occurrence between two grey levels, i and j, for a given displacement vector (δ_x, δ_y), for the specified window size. G is the number of quantized grey levels. Traditionally, the probabilities are stored in a grey level co-occurrence matrix (GLCM), where index (i, j) in the matrix represents the probability C_{ij}. Statistics are applied to the GLCM to generate texture features which are assigned to the center pixel of the image

window. Three commonly used grey level shift invariant statistics are:

$$Entropy(ENT) \quad -\sum_{i,j=0}^{G-1} C_{ij} \log C_{ij}, \tag{2}$$

$$Contrast(CON) \quad \sum_{i,j=0}^{G-1} C_{ij}(i-j)^2, \tag{3}$$

$$Correlation(COR) \quad \sum_{i,j=0}^{G-1} \frac{(i-\mu_x)(j-\mu_y)C_{ij}}{\sigma_x \sigma_y}. \tag{4}$$

Extraction of GLCP texture features from an image requires expert knowledge of the parameters involved. The parameters used by the GLCP method are: (1) image quantization (G), (2) displacement vector (δ_x, δ_y), (3) window size (n_x, n_y) and (4) statistic selection. For G, Soh and Tsatsoulis [7] indicate that setting G to 64 is sufficient. Clausi [5] recommends setting G to 24, however, larger values of G (> 64) are deemed excessive. Selection of interpixel displacement vector (δ_x, δ_y) is dependent on the nature of the textures to be segmented [11]. The displacement vector can be represented by polar co-ordinates: displacement (δ) and orientation (θ). Typically, four common orientations ($0°, 45°, 90°$ and $135°$) are used as they are easy to calculate. Small window sizes will cause poorly sampled co-occurring probabilities in the region of interest. This leads to an inconsistent estimate of the individual texture features, thus poorer segmentation [12]. However, if the window size is too large, it is more likely to overlap multiple classes in the image, thus eroding class boundaries.

Many of the statistics suggested by Haralick [2] produce highly correlated texture features [13,3,6] which is not desirable [14]. Baraldi and Parmiggiani [13] recommend that energy (or entropy) and contrast are the most significant texture statistics. Barber and LeDrew [3] indicate the best texture discrimination is achieved for a set of three texture statistics considered simultaneously. Clausi [5] studied the relationship of the statistical parameters and concludes that entropy, contrast and correlation compose a preferred set of statistical parameters.

2.2. Weighted grey level co-occurrence probability (WGLCP) texture features

2.2.1. Motivation

Large windows used for feature extraction are problematic for images with irregularly shaped texture boundaries and in regions of high *boundary density*. Boundary density is defined as the percentage of boundaries found in a given image window. A boundary differs from an edge in that it defines the border between two classes in an image; whereas an edge is any high contrast delineated region in an image. Mathematically, the boundary density for a given
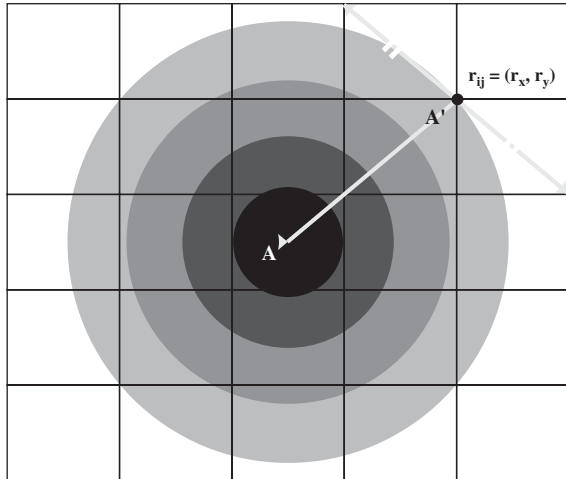
Fig. 1. Measuring the pixel pair distance to the center of the image window.

rectangular window (W) in an image is defined as

$$\frac{\sum_{\forall x,y} B(x, y)}{n_x n_y},\qquad(5)$$

where n_x, n_y are the number of columns and rows in W respectively, and x, y are the co-ordinates of W. $B(x, y) = 1$ if the pixel is a boundary and 0 otherwise.

Although features generated from the GLCM are assigned to the center pixel of the image window, all pixel pairs in the window are given a uniform probability weighting. This implies that pixel pairs far from the center of the image window will have the same impact to the feature measurement as those close to the center. A Gaussian weighting scheme is proposed which will be based on the pixel pair's distance from the center of the window. This technique will be referred to as the weighted GLCP (WGLCP) method. As illustrated in Fig. 1, the location of the pixel pair to the center of the image window determines the relative probability of co-occurrence for the weighted GLCM. The greater the length of A-A′ in the figure, the lower the probability of occurrence. By employing this weighting scheme, larger window sizes are permitted which is expected to lower the variability of the texture features. As well, the weighting scheme will provide better neighborhood approximations of the feature values, thus preserving the texture boundaries.

2.2.2. Generation of weighted co-occurrence texture features

Assume for the image window, W, that n_x and n_y are odd numbers. Then the image window can be indexed as follows:

$$W_x \in \left\{ -\left\lfloor \frac{n_x}{2} \right\rfloor, -\left(\left\lfloor \frac{n_x}{2} \right\rfloor - 1\right), \ldots, \left(\left\lfloor \frac{n_x}{2} \right\rfloor - 1\right), \left\lfloor \frac{n_x}{2} \right\rfloor \right\},$$

$$W_y \in \left\{ -\left\lfloor \frac{n_y}{2} \right\rfloor, -\left(\left\lfloor \frac{n_y}{2} \right\rfloor - 1\right), \ldots, \left(\left\lfloor \frac{n_y}{2} \right\rfloor - 1\right), \left\lfloor \frac{n_y}{2} \right\rfloor \right\},\qquad(6)$$

where $W_x \times W_y$ is the set of pixels in the image window indexed by their x–y (i.e. column–row) designations. This indexing scheme is selected so that the origin (0,0) of the image window will reside at the center pixel location, which is convenient for calculating the WGLCPs. The grey levels in the image window can be represented as a function of the index as follows:

$$W(x, y) = k \quad \text{where } k \in \{0, 1, \ldots, G - 1\}.\qquad(7)$$

Given two pixels in the image window that are separated by (δ_x, δ_y), one with grey level i, the other with grey level j, the pixel pair location is simply the mean location of $W(x_1, y_1)$ and $W(x_2, y_2)$, and is defined as follows:

$$
\begin{aligned}
r_{xy|ij} &= \left(\frac{x_1 + x_2}{2}, \frac{y_1 + y_2}{2} \right) \\
&= (r_x, r_y) | W(x_1, y_1) = i, W(x_2, y_2) = j, \\
&\quad (\delta_x, \delta_y) = (x_1 - x_2, y_1 - y_2).
\end{aligned}\qquad(8)
$$

In Section 2.1, P_{ij} is defined as the co-occurring frequency between two grey levels. However, in the weighted formulation, P_{ij} is the non-normalized weighted co-occurring probability and is a Gaussian function of $r_{xy|ij}$. Formally, for a fixed interpixel displacement vector (δ_x, δ_y), the non-normalized weighted co-occurring probability is

$$
\begin{aligned}
P_{ij}\left(r_{xy|ij}\right) = \sum_{\forall r_{xy|ij}} \frac{1}{2\pi\sigma_x\sigma_y} \\
\times \exp\left\{ -\frac{1}{2}\left[\left(\frac{r_x}{\sigma_x}\right)^2 + \left(\frac{r_y}{\sigma_y}\right)^2 \right] \right\},
\end{aligned}\qquad(9)
$$

where σ_x and σ_y represent the standard deviation of the two-dimensional Gaussian probability density function (pdf). The GLCM data structure is used to store these weighted probabilities, hence, Eq. (1) can be used to normalize P_{ij} to generate C_{ij}. As before, the statistics entropy, contrast, and correlation are applied to generate texture features.

2.2.3. Selecting standard deviation

The WGLCP requires determination of σ_x and σ_y. Since square image windows are used, $\sigma = \sigma_x = \sigma_y$. Smaller standard deviations (σ) indicate that pixel pairs further away from the center of the image window should be given less significance, whereas larger values for standard deviation provide a more uniform weighting scheme. Since the size of the image window determines the maximum Euclidean distance that a pixel pair will be from the center, the standard deviation should be selected accordingly.

Choice of an appropriate window size is influenced by the boundary density of an image. For circumstances with high boundary density, a smaller window size is recommended so that there is less texture boundary overlap. Once an adequate window size is selected for a given image, the standard deviation can be selected. To maximize the energy of the Gaussian pdf within the effective image window, the standard deviations are estimated to be $\frac{1}{4}$ of the window size.

This captures approximately 95 percent of the pdf's energy (area) in the effective window. Formally, the standard deviation can be written as a function of the window size as follows:

$$\sigma_x = \frac{n_x}{4}, \quad \sigma_y = \frac{n_y}{4}. \tag{10}$$

3. Boundary response performance

A boundary transect is a profile view of a texture feature as the image window moves across a texture boundary. This section uses boundary transects to compare different features. The first test will compare the relative performance of the GLCP and WGLCP features, whereas the second test will investigate the suitability of the texture statistics.

3.1. Explanation

Fig. 2 illustrates a bi-partite image with a vertical boundary. The left side of the image is texture "A" and the right side is texture "B". Using the parameter set outlined in Table 1 with a window size of 19×19, the boundary transects are calculated as follows:

(1) Randomly select a row in the image.
(2) For each column in the given row, calculate the relevant GLCP and WGLCP statistics for the different displacement vectors.
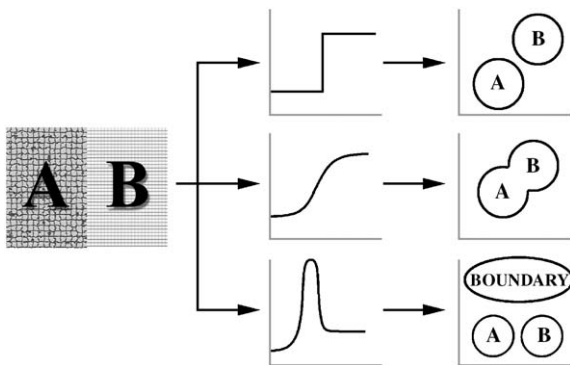

Fig. 2. Visual interpretation of edge transects. Illustrated are ideal (step-like), over- and under-damped texture boundary transitions for the edge transects. Each type of response will have impact on the feature space separability.

Table 1
GLCP and WGLCP parameters used in feature extraction

(W)GLCP parameters	Values
Displacement vector	$(1, 0), (1, 1), (0, 1), (-1, 1)$
Quantization	32 levels
Statistics	entropy, contrast, correlation
Standard deviation	$n/4 = 4.75$ (for WGLCP only)

(3) As the selected textures are isotropic in nature, average each statistic over the four displacement vectors used.
(4) To increase the signal-to-noise ratio, repeat steps 1 to 3 twenty times and average over the twenty samples.
(5) Normalize each feature dimension to the range [0, 1] to resolve the scales between the different features.
(6) For each feature, plot the feature value versus image column position.

Ideally, the value of a feature will have a step-like transition as it crosses the texture boundary, which results in well-defined separable classes in the feature space. In theory, the transition can also be over- or under-damped [15]. Clausi and Yue [10] indicate the boundary response is expected to be over-damped for most scenarios. The under-damped scenario is unanticipated as they expect there to be a relatively linear weighting between the texture features across the boundary. These three scenarios are outlined in Fig. 2. An over-damped response causes a higher within-class variance and lower between-class separability, resulting in poorer cluster discrimination relative to the step response. When the response is severely under-damped (impulse-like), the boundary response will have significant overshoot as it crosses the texture boundary and cause a confounding feature space. Essentially, features which are close to the boundary can form a distinct and separable third class (Fig. 2) when there is high boundary density. This will cause clustering algorithms to separate boundary versus non-boundary pixels rather than different textures.

3.2. Comparing GLCP and WGLCP features

3.2.1. Test data

The first test image (Fig. 3a) consists of synthetic aperture radar (SAR) ice (left-side) and water (right-side) textures. These textures are from the Beaufort Sea (June 1998) acquired from the Radarsat-1 platform, ScanSAR wide beam mode. This mode operates in C-band frequency, HH polarization, and is 50×50 m resolution that is 2×2 block averaged. The second test image (Fig. 4a) is comprised of cork (left-side) and wool (right-side) textures from the Brodatz photo album [16]. The textures for this image are adjusted to have the same mean grey level (DC shift with no information loss).

3.2.2. Methods of analysis

To determine which feature set provides better boundary preservation, a comparison of the gradient for each feature as it crosses the texture boundary is computed. For a window size of 19×19, approximately 10 samples of the gradient are taken across the boundary and averaged; this is defined as the average gradient for the texture boundary. This number of samples is deemed to be sufficient to represent the slope of the boundary transition based on the effective window size. By taking the WGLCP:GLCP ratio of the average gradient for each feature, the methods can be compared. If
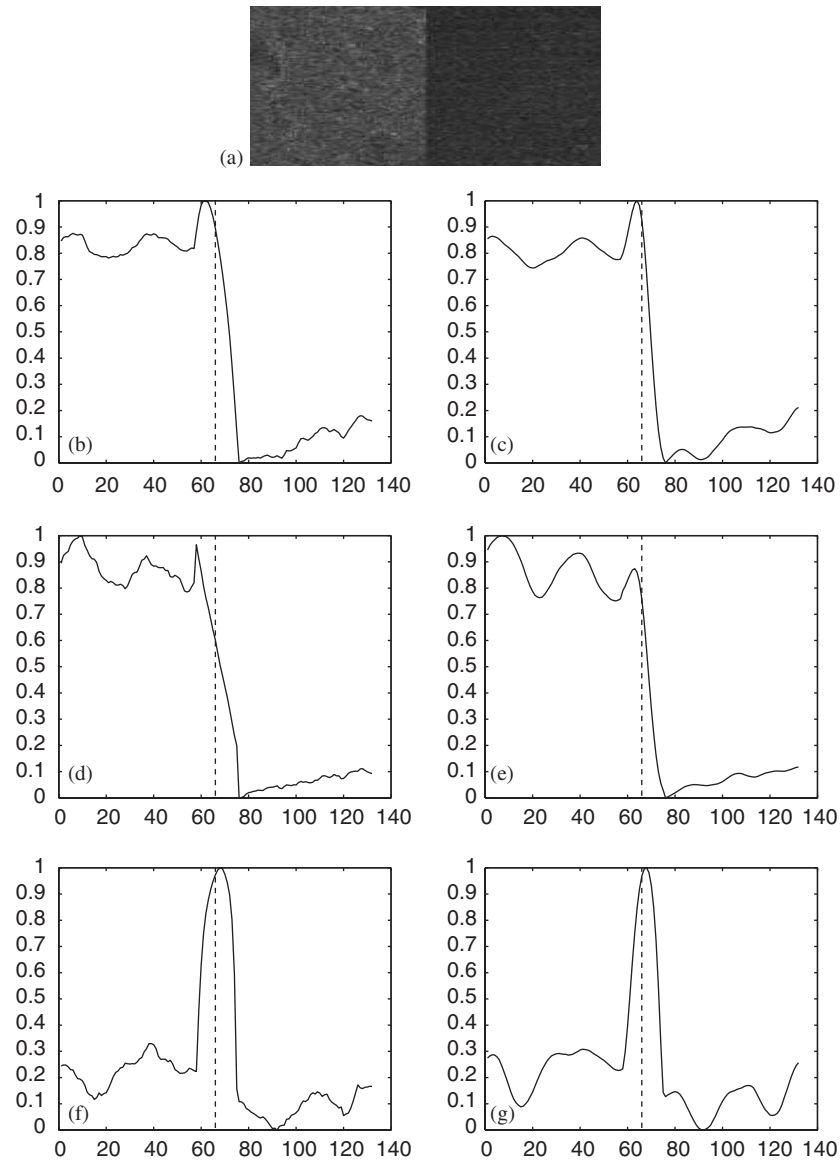
Fig. 3. Edge transects for SAR ice-water bi-partite image for selected feature statistics. Original test image included. (a) SAR ice-water image, (b) GLCP: entropy results, (c) WGLCP: entropy results, (d) GLCP: contrast results, (e) WGLCP: contrast results, (f) GLCP: correlation results and (g) WGLCP: correlation results.

this number is greater than 1, it indicates that the WGLCP feature has a steeper texture boundary response and is more desirable.

3.2.3. Results

Figs. 3 and 4 show the edge transect plots for the SAR and Brodatz bi-partite images, respectively. For these figures, the statistical features used are entropy, contrast and correlation as recommended in [5]. The dashed vertical line in these figures represents the true texture boundary. For both images, it is visually apparent that the WGLCP features have a sharper transect response as they cross the texture boundary, indicating better boundary preservation. For both test images, the WGLCP:GLCP average gradient ratio is greater than 1 (Table 2) for the statistics (excluding correlation),

verifying that the WGLCP features provide better boundary preservation during segmentation compared to the GLCP features.

Observing the correlation plot for the SAR image (Fig. 3), the transect is under-damped, and causes one to question the true suitability of using the correlation statistic for these high contrast textures. Section 3.3 will investigate the GLCP correlation statistic further to provide an explanation for this phenomenon.

3.3. Investigating statistical parameters

3.3.1. Test data

The test data consists of two synthetic bi-partite textured images as shown in Fig. 5a and b. The textures were
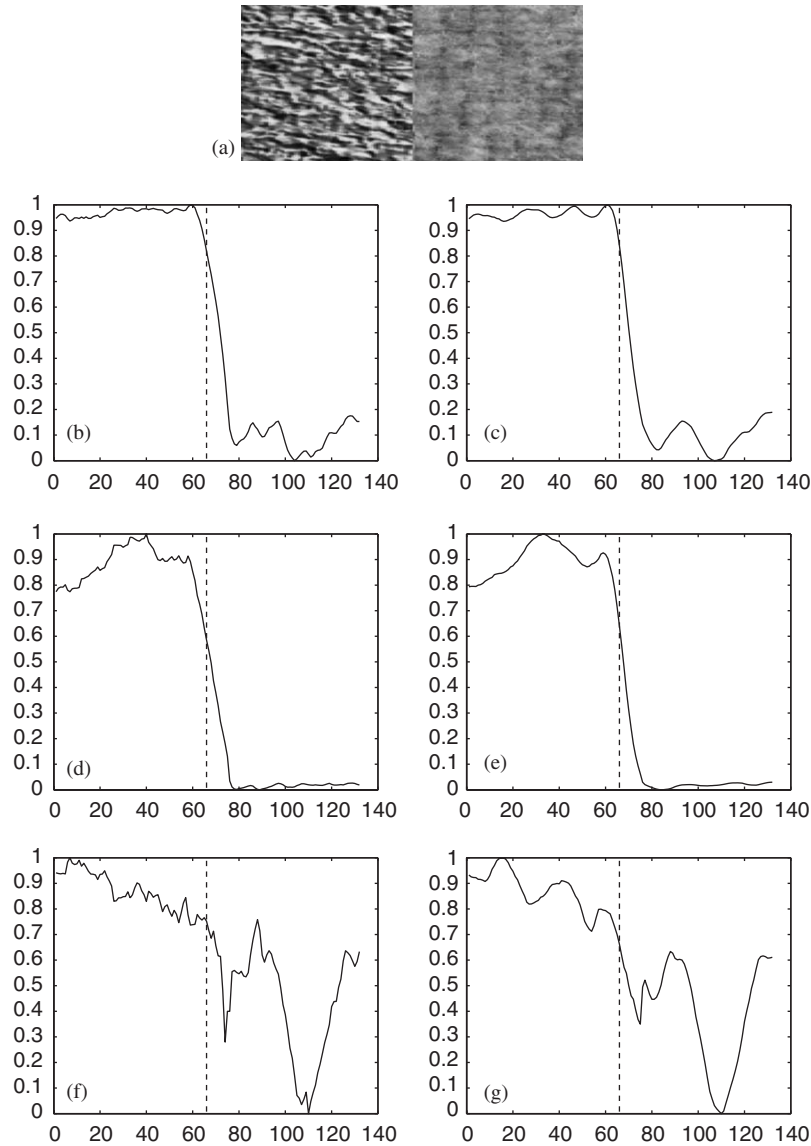
Fig. 4. Edge transects for Brodatz cork-wool bi-partite image for selected feature statistics. Original test image included. (a) Brodatz cork-wool image, (b) GLCP: entropy results, (c) WGLCP: entropy results, (d) GLCP: contrast results, (e) WGLCP: contrast results, (f) GLCP: correlation results and (g) WGLCP: correlation results.

Table 2
WGLCP:GLCP average gradient ratio for selected statistical features

	SAR image	Brodatz image
Entropy	1.30	1.22
Contrast	2.12	1.48
Correlation	N/A	N/A

synthetically created from sinusoidal functions with a specified frequency, amplitude and DC gain. By varying these three parameters, different textural effects can be synthesized. As well, white Gaussian noise ($\sigma = 1.5$) was added to each image in the test set to make the results more realistic.

Each test image contains different texture pairs as follows:

(1) *High contrast boundary*: The first texture is dark (i.e. low DC gain), has low variance (i.e. amplitude) and is of low frequency. The second texture is brighter, has higher variance and higher frequency.
(2) *Low contrast boundary*: Both textures in this image have the same mean grey level. The first texture has a low variance and frequency. The second texture has a high variance and frequency.

3.3.2. Results

Fig. 5 contains the boundary transect results for the high and low contrast images. The true texture boundary is indicated by the dashed vertical line. The plots on the left-hand
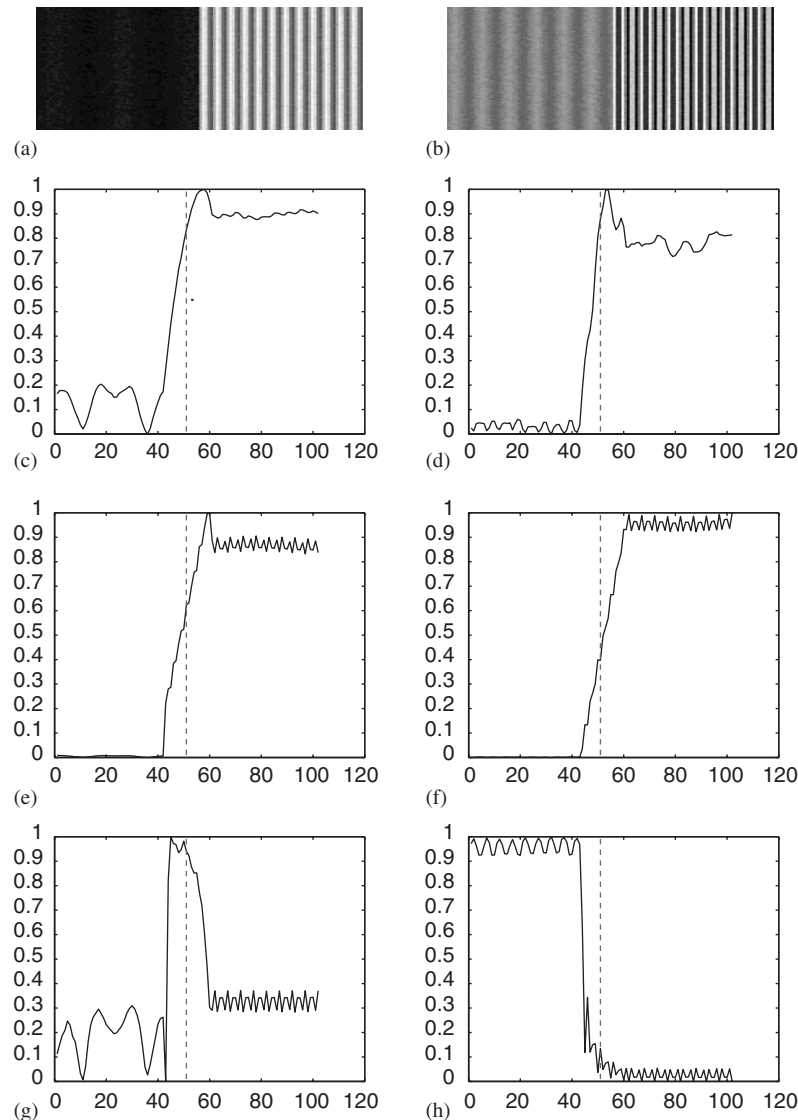
Fig. 5. Edge transects for synthetic bi-partite imagery. The images on the left are the feature transects for high contrast image. The images on the right are the feature transects for low contrast image. The true texture boundary is indicated by the dashed black line. Original test images included. (a) Textures with different DC gain, different amplitude and different frequency. (b) Textures with the same DC gain, different amplitude and different frequency. (c) High contrast image: entropy results. (d) Low contrast image: entropy results. (e) High contrast image: contrast results. (f) Low contrast image: contrast results. (g) High contrast image: correlation results. (h) Low contrast image: correlation results.

side are transects for the high contrast texture boundary, whereas the right-hand side contains transects of the low contrast texture boundary. The entropy and contrast statistical features provide a step-like response across the texture boundary for both images.

The GLCP correlation feature across only the high contrast boundary shows an under-damped response. This can be explained through analyzing the probability distributions within the GLCM (Fig. 6). During feature extraction, the distribution of non-zero elements in the GLCM was monitored within and across the texture boundary. For the dark texture, there is a high distribution of non-zero elements concentrated in the upper-left quadrant of the GLCM. For the

higher frequency bright texture, the bottom-right quadrant of the GLCM is more highly populated. During the transition from dark-to-light textures, the probability distribution is distributed across the top-left quadrant to the bottom-right quadrant; hence, the co-occurrence probability distribution becomes observably more correlated.

For the image where both textures have the same mean grey level, the correlation parameter produces the expected over-damped response. This indicates that it will make an appropriate feature for segmentation under such circumstances. However, if there are high contrast differences at the boundary between the two textures, the undesirable under-damped response results.
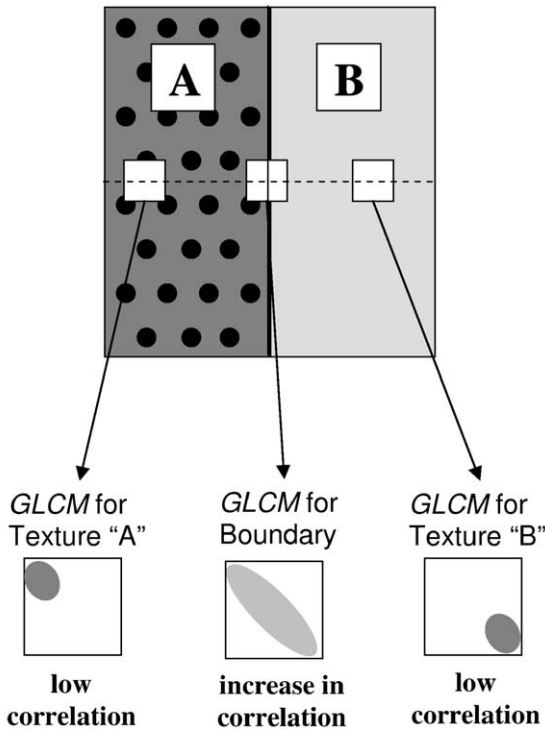
Fig. 6. Visualization of how the probability distribution in the GLCM changes across high contrast texture boundaries. Texture "A" is has a low mean grey level and texture "B" has a high mean grey level. The probability distribution in the GLCM becomes more correlated as it crosses the texture boundary.

4. Segmentation performance

This section focuses on investigating the segmentation performance of the texture features. The first test will measure the cluster separability of the GLCP features. The intention is to verify that the correlation parameter produces poor texture features for high contrast imagery. The second test will compare various feature sets in terms of segmentation performance. The final test will provide a computational comparison for both methods in terms of image segmentation time.

4.1. Cluster separability analysis

4.1.1. Explanation

Studies suggest that pixels near boundary regions have a tendency to form their own clusters in the feature space [9]. Using a labelled feature space, one can perform Fisher analysis [17] to determine if these boundary pixels are forming separate clusters in the feature space. This test intends on extracting labelled GLCP texture features from a SAR sea ice image according to the parameters defined in Table 1 with a window size of 15×15. One of three labels will be assigned to each feature: ice, water or boundary. The interclass separability is evaluated by the Fisher criterion for all pairs of classes (i.e. 'ice versus water', 'ice versus boundary', and 'water versus boundary').
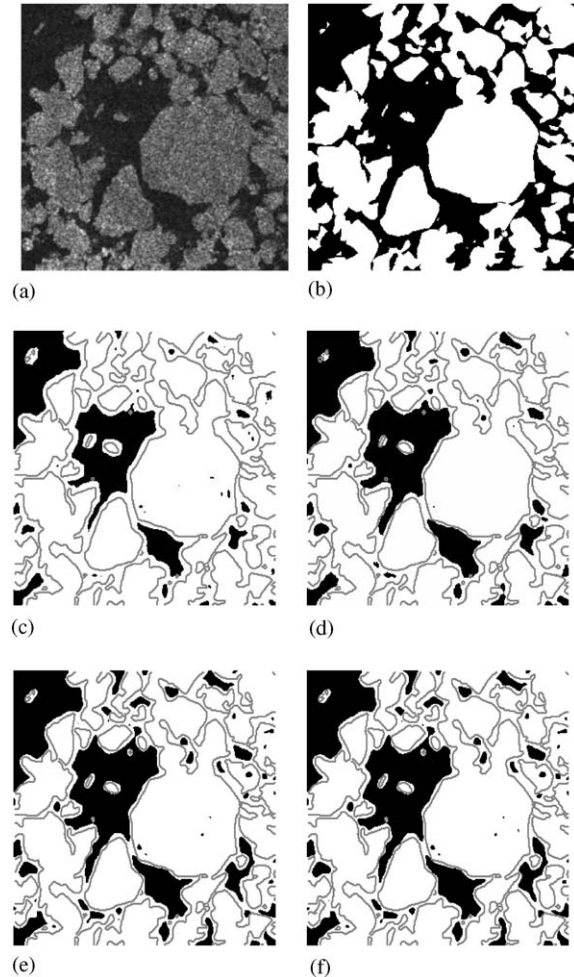

Fig. 7. RADARSAT SAR sea ice image of the Beaufort Sea. Courtesy of Canadian Space Agency©. Included are the segmentation composites for tests in Section 4.2. Grey lines represent true boundaries as obtained from the ground truth image. (a) Original, (b) Manual segmentation, (c) GLCP using entropy, contrast, correlation, (d) GLCP using entropy and contrast, (e) WGLCP using entropy, contrast, correlation and (f) WGLCP using entropy and contrast.

Using the ground truth segmentation, class labels are assigned. However, due to the smearing effects of large window sizes required by the GLCP features, a methodology for labelling 'boundary' pixels is determined. A pixel is labelled boundary if its distance is at least $\lceil n/4 \rceil$ pixels from a true boundary; where n is the size of the image window, and the true boundary is determined from the ground truth image.

4.1.2. Test data

Fig. 7a is a SAR sea ice image of Beaufort Sea acquired from the RADARSAT-1 platform. The radar platform parameters are identical to the parameters for Fig. 3a (Section 3.2). This image is selected for testing as it has many high-contrast, irregularly shaped texture boundaries. Included is a manual segmentation of the image which is used as ground truth (Fig. 7b).

4.1.3. Methods of analysis

The distance between two clusters is a common measurement in determining cluster separability. Using the Fisher criterion, a weighted distance between the projected clusters is measured [14]. Essentially, to determine the Fisher criterion for two labelled classes (C_1 and C_2) in the feature space, the ratio of the between-class scatter matrix (S_B) and within-class scatter matrix (S_W) is used as follows [17, p. 448]:

$$J = tr\left(S_W^{-1} S_B\right), \tag{11}$$

where *tr* denotes trace. Larger values of J indicate greater cluster separability.

4.1.4. Results

Table 3 summarizes the results of Fisher analysis on different feature spaces. According to the Fisher criterion, when the feature space comprises only the entropy statistic, the ice-water separability is substantially higher (4.21) than the other pairs (0.08 for ice-boundary and 2.57 for water-boundary), indicating that the ice and water pixels are more distinct in the feature space. In this case, there is a stronger tendency to segment into the two desired classes of ice and water. The ice-boundary separability is low (0.08) relative to the water-boundary separability (2.57), indicating that there is a bias which will tend to group ice and boundary pixels together, even in situations where the boundary pixel should be classified as water. Table 3 also shows that contrast features alone will produce similar results as entropy features in terms of inter-class separability. Using entropy, the separability of ice and water is 2.74 while the separability of ice and boundary (0.39) and water and boundary (1.00) are substantially lower. Thus, in the cases of using entropy or contrast features alone, proper segmentation is expected to be performed.

On the other hand, the Fisher results (Table 3) verify that the correlation statistic is not suitable for segmentation. Using correlation, both ice versus boundary (1.82) and ice versus water (0.71) have higher separability than ice versus water (0.50), indicating the boundary is easily identified as a separate class. Given the high contrast nature of the ice and water boundary, this result parallels the result obtained in Section 3.3.

When all three statistics (entropy, contrast, correlation) are combined, the Fisher criterion shows the best separability is between ice and water classes (12.55). Also, for this scenario, the numbers show that there is still a slight bias to over-classify boundary pixels as ice (1.83 versus 2.87). Finally, when the correlation statistical feature is removed from the feature space (leaving only contrast and entropy features), ice-water separability is increased (15.01 versus 12.55) and the boundary biases are reduced (1.83 down to 0.45 and 2.87 down to 2.72). Under this testing, contrast and entropy represent the preferred statistics for image

segmentation when high contrast boundaries could be present in the image.

4.2. Determining a preferred feature set

4.2.1. Explanation

WGLCP texture features have been introduced and are demonstrated to provide better texture boundary preservation compared to the GLCP features. Also, the correlation parameter does not make a suitable feature when segmenting images with high contrast boundaries. Combining these ideas, four feature spaces for segmenting a variety of imagery are compared:

(1) GLCP texture feature extraction using entropy, contrast and correlation statistics;
(2) GLCP texture feature extraction using entropy and contrast (no correlation) statistics;
(3) WGLCP texture feature extraction using entropy, contrast and correlation statistics; and
(4) WGLCP texture feature extraction using entropy and contrast (no correlation) statistics.

Aside from the statistics, the remaining (W)GLCP user-defined variables are listed in Table 1. After feature extraction, the feature space is normalized and k-means clustering is applied [14]. Based on the experiments conducted in the previous sections, the hypothesis is that feature set (4) (i.e. WGLCP using entropy and contrast statistics) should provide the preferred segmentation results for the high contrast images used in this section. Note that adding the grey tone to the feature space for high contrast images would improve segmentation; however, these tests focus on comparing only the texture feature sets.

4.2.2. Test data

This test data consists of two test images. The first image is the SAR sea ice image (Fig. 7a) used in Section 4.1. This image was selected as it represents natural imagery. A window size of 15×15 was used for feature extraction. The second image (Fig. 8a) is a Brodatz mosaic image and its corresponding ground truth image (Fig. 8b). This image is adapted from [18] and contains four distinct (and labelled) textured regions from the Brodatz photo album [16]. The textures used in the image are cork (D4), cotton canvas (D77), wool (D92) and straw matting (D55). This image has many irregularly-shaped texture boundaries, which makes segmentation more difficult. As well, Brodatz textures are extensively used in the research literature. A window size of 19×19 is used for feature extraction.

4.2.3. Methods of analysis

Using the ground truth image, the cumulative correct and incorrect classifications for each test set are determined and stored in a confusion matrix. The kappa statistic ($\hat{\kappa}$) and confidence interval ($\hat{\sigma}$) are commonly used to

Table 3
Fisher measures for various feature spaces of the Beaufort image

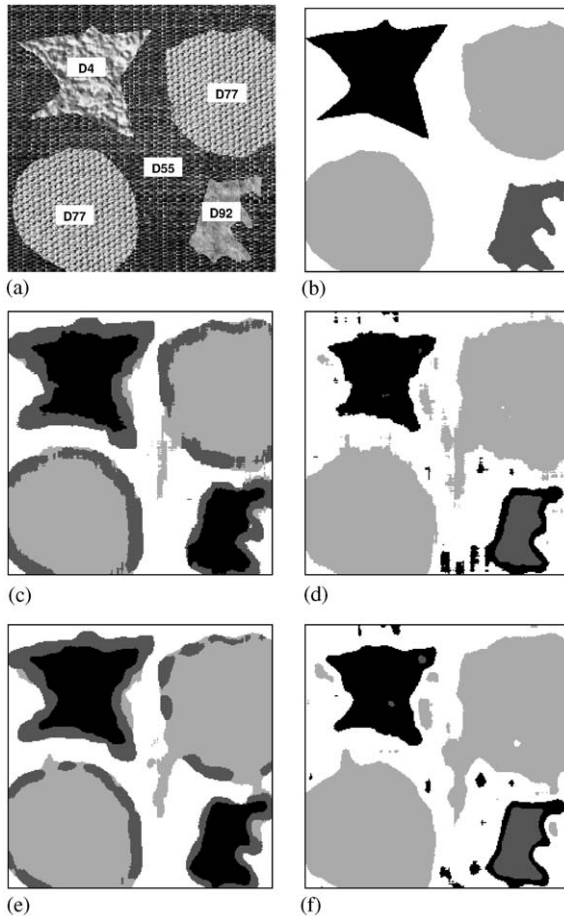|	Statistics used in feature space				
	Entropy	Contrast	Correlation	Entropy, contrast, correlation	Entropy, contrast
Ice vs. water	4.21	2.74	0.50	12.55	15.01
Ice vs. boundary	0.08	0.39	1.82	1.83	0.45
Water vs. boundary	2.57	1.00	0.71	2.87	2.72

Fig. 8. Brodatz mosaic "star" image. Contains four different textures: cork (D4), cotton canvas (D77), wool (D92) and straw matting (D55). Segmentation results included. (a) Original, (b) Manual segmentation, (c) GLCP using entropy, contrast, correlation, (d) GLCP using entropy and contrast, (e) WGLCP using entropy, contrast, correlation and (f) WGLCP using entropy and contrast.

Table 4
Comparing segmentation performance of different feature sets

	Beaufort image		Brodatz "Star" image	
	Overall accuracy (%)	Kappa statistic	Overall accuracy (%)	Kappa statistic
Feature set 1	72.9	0.39	70.9	0.59
Feature set 2	77.0	0.49	85.2	0.77
Feature set 3	79.1	0.54	74.2	0.63
Feature set 4	79.3	0.54	86.3	0.78

4.2.4. Results

Fig. 7 shows the segmentation results using the four feature sets for the Beaufort image. The grey lines overlayed in the image represent the true boundaries. Percentage accuracies and kappa coefficients are presented in Table 4. With respect to the GLCP results, the segmentation accuracy increases from 72.9% (using contrast, entropy and correlation statistics) to 77.0% (when the correlation statistic is removed). Performing pairwise statistical significance testing on all six possible feature set pairs, all pairs produce significantly different results except for the WGLCP (with correlation) and WGLCP (without correlation) pair. Note that there is a percentage increase when removing the correlation statistic in the WGLCP case: the segmentation accuracy increases from 79.1% to 79.3%.

Fig. 8 shows the segmentation results using the four feature sets applied to the Brodatz "star" image. For both GLCP and WGLCP methods, it is very apparent that including correlation feature leads to poorer segmentation. The texture boundaries between the four objects and the background form a separate class, which causes the cork and wool textures to be clustered together. Once the correlation feature is removed from the feature space (Fig. 8(c) vs. (d) and Fig. 8(e) vs. (f)), the boundary errors are mostly resolved. However, there is some boundary misclassification around the wool texture (lower-right). The boundary transition from wool to straw matting mimics the cork texture features, resulting in boundary misclassification. As indicated in Table 4, the results omitting correlation for the GLCP and WGLCP methods lead to improvements of 85.2% from 70.9% (for GLCP) and 86.3% from 74.2% (for WGLCP). Table 4 also indicates that the WGLCP method is more accurate than the GLCP method, given the same feature set.

evaluate the confusion matrix [19]. When two error matrices are compared, the following test statistic can be used to determine a significance value (using a significance level of 5%) [19]:

$$Z \sim \frac{\hat{\kappa}_1 - \hat{\kappa}_2}{\hat{\sigma}_1 - \hat{\sigma}_2}. \qquad (12)$$

All pairwise comparisons generate statistically significant differences except for the pair for GLCP (all statistics) versus WGLCP (all statistics). In this case, the correlation statistic leads to poor boundary approximation in each case leading to results that are not significantly different. Note that there is a percentage increase from 70.9% to 74.2% in this case.

Given these results for the two images, it is demonstrated that the WGLCP is an improvement over the GLCP and that removing the correlation statistic improves the segmentation.

4.3. Computation time

For this research, the experiments are performed on a Pentium IV, 1.6 GHz computer running Windows 2000© and running Matlab version 6.1© Release 12.1 with 512 Mb RAM. The results indicate the WGLCP method is an order of magnitude slower than the GLCP method, where the GLCP features are calculated using a modification of the GLCHS [20] method, which is found in the PCI Geomatica™ software. For a 256×256 image, the WGLCP method takes approximately 9.5 hours to calculate the texture features according to the parameter set outlined in Table 1 with a window size of 19×19. Nevertheless, with the advancements of computer hardware architecture, this computational burden will continue to be reduced. As well, this time can be reduced significantly by implementing the algorithm in C code (versus Matlab©), as C code is far more efficient at performing iterations. By using a coarser quantization level, the computational time can be further reduced.

5. Conclusions

The goal of this paper is to improve image texture segmentation with focus on the preservation of texture boundaries. From these tests, several conclusions are reached:

- The WGLCP is a novel technique that generates texture features that are better able to identify class boundaries than the standard GLCP approach. By weighting pixel pairs in the center of the window higher than pixel pairs at the window boundary, improved features are generated.
- In images where there is high contrast difference across class boundaries, the correlation statistic is not suitable for segmentation. In such cases, the correlation statistic should be avoided. An explanation for this phenomenon has been provided. Since natural texture images often have high contrast boundaries, avoiding the correlation statistic for segmentation problems is strongly advocated.
- For general image segmentation, this paper advocates analyzing edge transects across class boundaries prior to performing segmentation. By analyzing the profile of a single feature measurement across a class boundary, the feature suitability is determined. If the feature shows an impulse-like (or under-damped) response across the class boundary, this indicates it will perform poorly at separating classes in the image. For optimal feature separability, a step-like response is desired across the class boundary. In essence, the extra initial time required to analyze edge transects will provide a more intelligent method in feature selection and will save time an aid in the overall segmentation process.

References

[1] M. Tuceryan, A.K. Jain, Handbook of Pattern Recognition and Computer Vision, Chapter 2: Texture Analysis, World Scientific, Singapore, 1993.

[2] R.M. Haralick, K. Shanmugam, I. Dinstein, Textural features for image classification, IEEE Trans. Syst. Man Cyber. 3 (6) (1973) 610–621.

[3] D.G. Barber, E.F. LeDrew, SAR sea ice discrimination using texture statistics: a multivariate approach, Photogrammetric Engineering and Remote Sensing 57 (4) (1991) 385–395.

[4] D.A. Clausi, Comparison and fusion of co-occurrence, Gabor, and MRF texture features for classification of SAR sea ice imagery, Atmosphere Oceans 39 (4) (2001) 183–194.

[5] D.A. Clausi, An analysis of co-occurrence texture statistics as a function of grey level quantization, Canad. J. Remote Sens. 28 (1) (2002) 45–62.

[6] M.E. Shokr, Evaluation of second-order texture parameters for sea ice classification from radar images, J. Geophy. Res. 96 (C6) (1991) 625–640.

[7] L.K. Soh, C. Tsatsoulis, Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices, IEEE Trans. Geosci. Remote Sens. 37 (2) (1999) 780–794.

[8] R. Jobanputra, D.A. Clausi, Texture analysis using Gaussian weighted grey level co-occurrence probabilities, in: 1st Canadian Conference on Computer and Robot Vision, London, Ontario, Canada, May 17–19, 2004.

[9] J. Mao, A.K. Jain, Texture classification and segmentation using multiresolution simultaneous autoregressive models, Pattern Recognition 25 (2) (1992) 173–188.

[10] D.A. Clausi, B. Yue, Comparing cooccurrence probabilities and Markov random fields for texture analysis of SAR sea ice imagery, IEEE Trans. Geosci. Remote Sens. 42 (1) (2004) 215–228.

[11] S. Zucker, D. Terzopoulos, Finding structure in co-occurrence matrices for texture analysis, Comput. Graph. Image Process. 12 (1980) 286–308.

[12] D.A. Clausi, B. Yue, Comparing co-occurrence probabilities and Markov random fields for texture analysis of SAR sea ice imagery, IEEE Trans. Geosci. Remote Sens. 42 (1) (2004) 215–228.

[13] A. Baraldi, F. Parmiggiani, An investigation of the textural characteristics associated with gray level coocurrence matrix statistical parameters, IEEE Trans. Geosci. Remote Sens. 33 (2) (March 1995) 293–303.

[14] R.O. Duda, P.E. Hart, D.G. Stork, Pattern Classification, second ed., Wiley, New York, 2001.

[15] R.C. Dorf, R.H. Bishop, Modern Control Systems, ninth ed., Prentice-Hall, Englewood Cliffs, NJ, 2001.

[16] P. Brodatz, Textures: A Photographic Album for Artists and Designers, Dover, New York, 1966.

[17] K. Fukunaga, Introduction to Statistical Pattern Recognition, Academic Press, New York, 1990.

[18] D.A. Clausi, H. Deng, Design-based fusion of Gabor filter and co-occurrence probability features for texture recognition, IEEE Trans. Image Process., July 2004, accepted.

[19] Y.M. Bishop, S.E. Fienberg, P.W. Holland, Discrete Multivariate Analysis: Theory and Practise, The MIT Press, Cambridge, MA, 1977.

[20] D.A. Clausi, Y. Zhao, Rapid extraction of image texture by co-occurrence using a hybrid data structure, Comput. Geosci. 28 (2002) 763–774.

About the Author—RISHI JOBANPUTRA earned his B.A.Sc. (2002) and M.A.Sc. (2004) in Systems Design Engineering at the University of Waterloo (Waterloo, Ontario, Canada). He is currently working in a global management consulting firm. During his academic career, Rishi has published refereed journal and conference papers ranging from aerospace to computer vision. His research interests are in pattern recognition, computer vision, and machine learning. He has received several scholarships and conference paper awards.

About the Author—DAVID A. CLAUSI earned his B.A.Sc. (1990), M.A.Sc. (1992), and Ph.D. (1996) in Systems Design Engineering at the University of Waterloo (Waterloo, Ontario, Canada). After completing his doctorate, Prof. Clausi worked in the medical imaging field at Mitra Imaging Inc. (Waterloo). He started his academic career in 1997 as an Assistant Professor in Geomatics Engineering at the University of Calgary, Alberta, Canada. In 1999, he returned to his alma mater and was awarded tenure and promotion to Associate Professor in 2003. Prof. Clausi is an active interdisciplinary and multidisciplinary researcher. He has an extensive publication record, publishing refereed journal and conference papers in the diverse fields of remote sensing, computer vision, algorithm design, and biomechanics. His primary research interest is the automated interpretation of synthetic aperture radar (SAR) sea ice imagery, in support of operational activities of the Canadian Ice Service. The research results have led to successfully commercial implementations. He has received numerous graduate scholarships, conference paper awards, two Teaching Excellence Awards.