

# Comparing Cooccurrence Probabilities and Markov Random Fields for Texture Analysis of SAR Sea Ice Imagery

David A. Clausi, *Senior Member, IEEE*, and Bing Yue

**Abstract**—This paper compares the discrimination ability of two texture analysis methods: Markov random fields (MRFs) and gray-level cooccurrence probabilities (GLCPs). There exists limited published research comparing different texture methods, especially with regard to segmenting remotely sensed imagery. The role of window size in texture feature consistency and separability as well as the role in handling of multiple textures within a window are investigated. Necessary testing is performed on samples of synthetic (MRF generated), Brodatz, and synthetic aperture radar (SAR) sea ice imagery. GLCPs are demonstrated to have improved discrimination ability relative to MRFs with decreasing window size, which is important when performing image segmentation. On the other hand, GLCPs are more sensitive to texture boundary confusion than MRFs given their respective segmentation procedures.

**Index Terms**—Comparison, discrimination ability, gray-level cooccurrence probability (GLCP), Markov random fields (MRFs), remote sensing, segmentation, texture features.

## I. INTRODUCTION

**T**EXTURE, a representation of the spatial relationship of gray levels in an image, is an important characteristic for the automated or semiautomated interpretation of digital images. Given the existing and expected volume of remote sensing imagery to be processed, the use of computer-assisted interpretation methods is important. Texture is recognized to be an integral component of such interpretation processes. Since synthetic aperture radar (SAR) imagery contains spatially dependent class characteristics, texture extraction methods have been commonly used for class discrimination. For example, in the case of SAR sea ice imagery, texture has been recognized to provide better characterization than using gray levels alone [1]–[3]. Sea ice segmentation has been a challenging research topic for years, and there are few publications about SAR sea ice image segmentation using texture methods, especially using unsupervised approaches. Many different variables may influence the appearance of sea ice in a SAR image (e.g.,

Manuscript received July 24, 2002; revised April 23, 2003. This work was supported in part by the Meteorological Service of Canada under the CRYosphere SYStem in Canada (CRYSYS) Project and in part by the Natural Sciences and Engineering Research Council. The Radarsat data in this paper are copyright by the Canadian Space Agency (CSA).

D. A. Clausi is with the Department of Systems Design Engineering, University of Waterloo, Waterloo, ON, Canada N2L 3G1 (e-mail: dclausi@eng-mail.uwaterloo.ca).

B. Yue is with Noetix Research Inc., Ottawa, ON Canada N2L 3G1 and also with the Canada Centre for Remote Sensing, Ottawa, ON, Canada K1A 0Y7 (e-mail: Bing.Yue@CCRS.NRCan.gc.ca).

Digital Object Identifier 10.1109/TGRS.2003.817218

5.2	4.3	3.2	4.2	5.1
4.4	2.2	1.2	2.1	4.1
3.1	1.1	c	1.1	3.1
4.1	2.1	1.2	2.2	4.4
5.1	4.2	3.2	4.3	5.2

Fig. 1. GMRF parameters used to generate the synthetic textures [30].

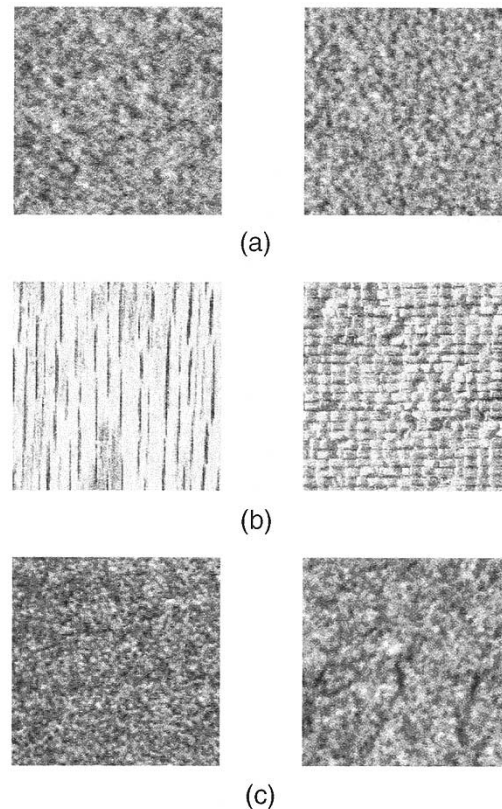


Fig. 2. Test image pairs for the second research question. (a) Synthetic GMRF textures (models A and B in Table I) ( $1024 \times 1024$ ). (b) Brodatz textures (wood grain and raffia) ( $1024 \times 1024$ ). (c) SAR sea ice textures (L: first-year ice R: multiyear ice) ( $768 \times 768$ ).

tone, texture, season, geographical location, weather conditions, etc.). The same ice type generally varies in appearance from image to image as well as within the same image. Thus, the selected training samples in a supervised algorithm may not be sufficient to include all the class variability throughout the image.

Many texture feature extraction methods exist. Tuceryan and Jain [4] identify four major categories for texture fea-

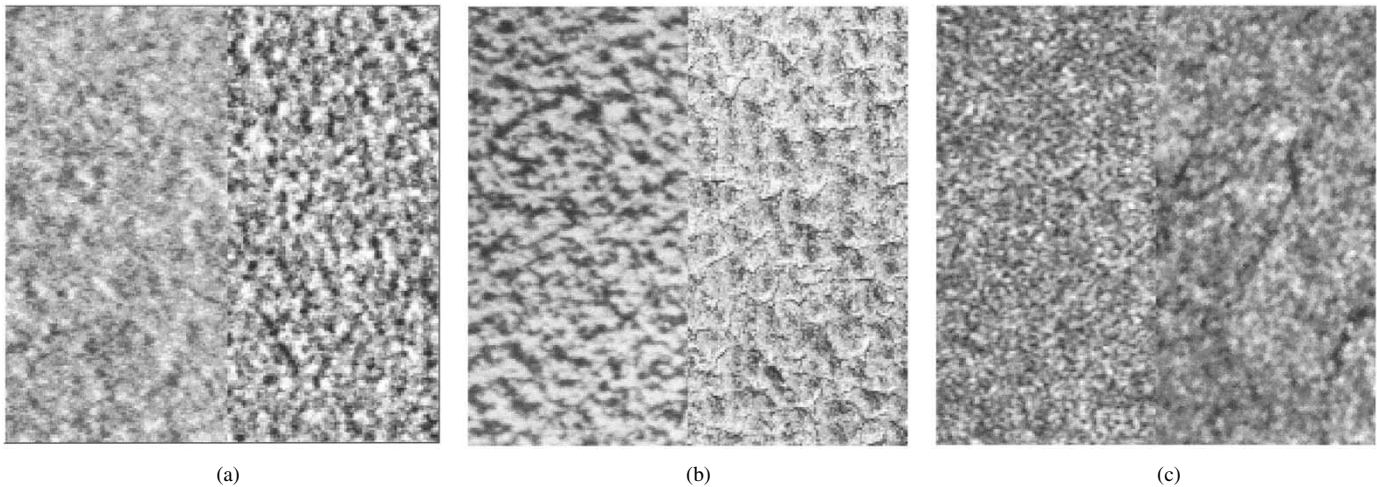


Fig. 3. Three texture images separated by straight boundary. (a) Synthetic texture. (b) Brodatz textures (paper and pigskin). (c) SAR sea ice textures.

tures: statistical (such as gray-level cooccurrence probabilities (GLCPs) [5]), geometrical (including structural), model-based (such as Markov random fields (MRFs) [6]), and signal processing (such as Gabor filters [7], [8]). Limited research has been conducted to compare methods between these major categories or to compare specific methods within each category. Also, limited research is known that compares methods in operational situations, such as the computer-assisted tactical use of remote sensing imagery. Any such comparative studies are summarized here.

Comparative texture studies not using remote sensing imagery include the following. Randen and Husoy [9] extensively compare only signal processing texture methods for segmentation purposes. Pichler *et al.* [10], [11] investigate wavelet segmentation methods applied to artificial and Brodatz imagery. Ojala *et al.* [12] performed a classification-only study comparing four second-order statistical operators (not including GLCPs). Ohanian and Dubes [13] compared MRFs, Gabor features, GLCPs, and fractal methods for classification purposes. Connors and Harlow [14] perform a theoretical classification comparison of GLCP, gray-level run length, gray-level difference, and power spectrum methods.

A number of studies have compared texture operators using remote sensing imagery. Weszka *et al.* [15] perform a remote sensing classification study using the same statistics as Connors and Harlow [14], with comparable results. Weszka *et al.* found that the power spectrum features generally performed poorly relative to the other feature sets. Carr and Miranda [16] compare, in a classification context, a statistical method (the semivariogram) with GLCPs. They utilize a variety of remote sensing imagery for testing. Their results indicate that the semivariogram produced higher classification accuracies than the GLCP method for microwave imagery and lower classification accuracies than the GLCP method for optical imagery. Whether or not one method is significantly outperforming the other is unknown, since no statistical testing was performed. Gong *et al.* [17] compare the GLCP method to a pair of alternative statistical methods for land-use classification applied to a Satellite Pour l'Observation de la Terre High Resolution Visible (SPOT HRV) band 3 land class image. They demonstrate

that the GLCP and simple statistical transformation methods can largely improve the classification accuracies using only the spectral images. Clausi [18] comprehensively compares GLCPs, MRFs, and Gabor filters for classification using nine classes of validated SAR sea ice imagery. Classification and significance level testing demonstrate that GLCPs classify the data with the highest classification rate, with Gabor filters a close second. MRF results significantly lag Gabor and cooccurrence results. However, the MRF features are uncorrelated with respect to cooccurrence and Gabor features. The fused cooccurrence/MRF feature set achieves higher classification performance.

The background review yields limited research comparing the two common texture approaches: model-based and statistical texture categories. Comprehensive, recognized supervised or unsupervised segmentation comparative studies using remote sensing imagery are not known to the authors. The cited comparison texture papers often only consider the classification problem (using homogeneous texture window samples with known class labels for training and testing), without considering full image segmentation, noted to be more challenging. In the case of image segmentation, windows will sometimes contain more than one texture class, making the feature extraction and class assignment decisions far more challenging, especially when unsupervised segmentation is employed. This paper compares the unsupervised segmentation capabilities of two popular texture methods: GLCPs (a statistical method) and MRFs (a model-based method). Most notably, the paper emphasizes the role of window size selection when determining and utilizing GLCP and MRF texture features for unsupervised segmentation. Test data includes synthetic (MRF generated), Brodatz, and SAR sea ice imagery. Synthetic and Brodatz imagery give the opportunity to test algorithms, where the class assignments are known for certain; the same cannot be said for the SAR sea ice imagery. Even though professionally created ice charts are used for validation, the potential for ice type mixing exists and is unavoidable.

The paper proceeds in the following fashion. Section II discusses the two feature extraction and segmentation schemes. Section III motivates and poses a number of research questions.

TABLE I  
GMRf PARAMETERS USED TO GENERATE  
THE SYNTHETIC TEXTURES [25]

Model Parameters	Model A	Model B
(-1,0)	0.520252	0.468389
(0,-1)	0.0934154	0.308257
(-1,-1)	0.0303413	-0.0755398
(-1,1)	0.0180476	-0.0755797
(-2,0)	-0.148331	-0.100678
(0,-2)	-0.0216434	-0.0407557

This is followed by Section IV, which describes the experiments and results used to investigate the research questions. A summary (Section V) concludes the paper.

## II. TEXTURE ANALYSIS METHODS

This section briefly describes the two texture analysis methods: GLCPs and MRFs.

### A. GLCPs

GLCPs, first published by Haralick *et al.* [5], are the most common method for texture feature extraction in the remote sensing literature. The cooccurrence probabilities are the conditional joint probabilities of all pairwise combinations of gray levels ( $i, j$ ) in the fixed-size spatial window given two parameters: interpixel distance ( $\delta$ ) and orientation ( $\theta$ ). To generate texture features based on the cooccurrence probabilities, statistics are applied to the probabilities. Generally, these statistics identify some structural aspect of the arrangement of probabilities stored within a matrix indexed on  $i$  and  $j$ , which in turn reflects some qualitative characteristic of the local image texture (e.g., smoothness or roughness). There are many statistics that can be used; however, due to the redundancy amongst these statistics, only three statistics are advocated for shift-invariant applications, since these should generate preferred discrimination [3]. The selected statistics are dissimilarity (D), entropy (E), and correlation (C). These have been used in this study, since each tends to produce independent features with respect to features produced by other cooccurrence statistics, and each represents a different characteristic of the cooccurring probabilities [19].

There is no known rigorous optimal method for selecting  $\delta$  and  $\theta$  *a priori*. In this paper, given no other information concerning the texture of interest, the preferred parameters will be as follows. Adjacent neighbors should be utilized, motivating the use of  $\theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ$ , and  $\delta = 1$ . The standard approach is to use a symmetrical cooccurrence matrix as determined by capturing  $0^\circ$  as well as  $180^\circ, 45^\circ$  as well as  $225^\circ$ , etc. SAR texture is generally characterized over spatial scales on the order of the sensor resolution, so setting  $\delta = 1$  pixel for the interpixel spacing is appropriate. To account for anisotropic behavior, individual orientations of  $0^\circ, 45^\circ, 90^\circ$ , and  $135^\circ$  are advocated. This combination of offset and orientation has characterized SAR texture well and has been identified for being preferred for SAR sea ice applications [3], [18]. Geomatica (a commercial remote sensing software package created by PCI Geomatica Inc.) is used to generate the cooccurrence texture features on a pixel-by-pixel basis.

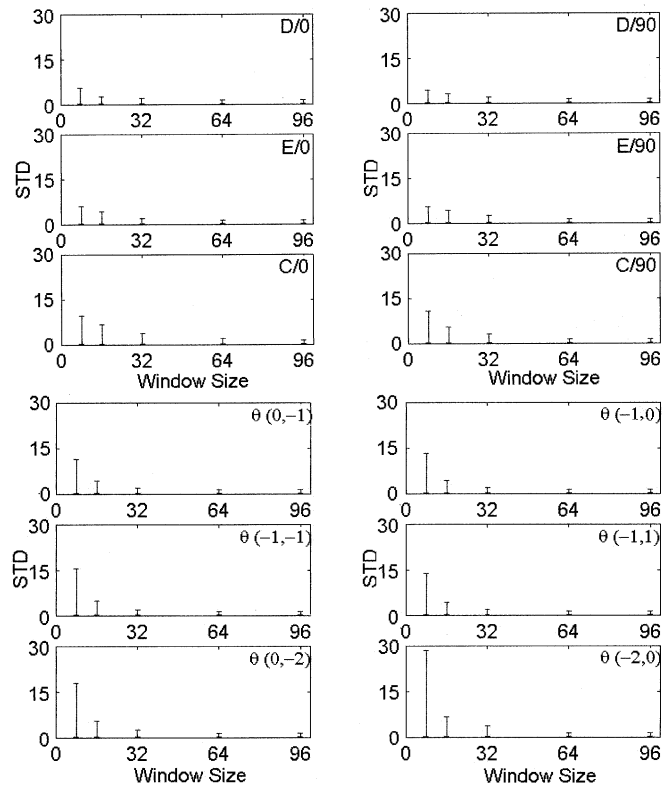


Fig. 4. Change of standard deviation of the estimated GLCP texture features and the GMRf model parameters with different window size. (a) GLCP texture features (D: dissimilarity, E: entropy, C: correlation,  $\delta = 1$ ). (b) GMRf model parameters.

TABLE II  
RATIO OF THE STANDARD DEVIATIONS FOR EACH WINDOW SIZE (64, 32, 16, 8)  
WITH RESPECT TO WINDOW SIZE 96 FOR BOTH GLCP AND GMRf FEATURES

GLCP texture features				
	8 to 96	16 to 96	32 to 96	64 to 96
Synthetic	9.33	5.45	2.93	1.55
Brodatz	6.85	4.22	2.33	1.24
Sea ice	2.51	2.04	1.34	0.99
GMRf model parameters				
	8 to 96	16 to 96	32 to 96	64 to 96
Synthetic	28.32	8.50	3.95	1.66
Brodatz	16.44	4.85	2.23	1.24
Sea ice	21.38	7.35	3.19	1.75

This software uses a fixed gray-level quantization level of 16. The role of varying the gray-level quantization with respect to GLCP statistics is presented in [19].

There are a variety of clustering approaches that could be used to assign class labels to the GLCP feature vectors. For simplicity, a standard  $k$ -means clustering routine [20] is used for the GLCP segmentation examples in this paper.

### B. MRFs

MRFs are recognized for being effective for texture analysis [21], [22]. Since the method allows a model of the grayscale pattern to be created, algorithms are developed based on fundamental principles (as opposed to ad hoc heuristics) [23]. The model parameters can be used to characterize the texture as well as generate facsimiles of the texture [21]. The basic premise

TABLE III  
BHATTACHARYYA ERROR BOUNDS AND FISHER CRITERIA  
OF THE TEXTURE PAIRS IN FIG. 2

Bhattacharyya error bounds (BEB) and Fisher criterion ( $J$ )				
8 × 8 window size				
	GMRF		GLCP	
	BEB	$J$	BEB	$J$
synthetic	$3.5 \times 10^{-1}$	0.24	$2.9 \times 10^{-1}$	1.68
Brodatz	$3.7 \times 10^{-1}$	0.24	$2.0 \times 10^{-1}$	1.52
sea ice	$2.5 \times 10^{-1}$	0.55	$4.8 \times 10^{-2}$	6.58
16 × 16 window size				
	GMRF		GLCP	
	BEB	$J$	BEB	$J$
synthetic	$1.7 \times 10^{-1}$	4.19	$1.3 \times 10^{-1}$	4.98
Brodatz	$1.3 \times 10^{-1}$	2.21	$4.4 \times 10^{-2}$	5.64
sea ice	$7.9 \times 10^{-13}$	101.03	$1.1 \times 10^{-3}$	21.07
32 × 32 window size				
	GMRF		GLCP	
	BEB	$J$	BEB	$J$
synthetic	$4.0 \times 10^{-4}$	26.33	$3.0 \times 10^{-3}$	19.75
Brodatz	$3.2 \times 10^{-3}$	21.85	$3.5 \times 10^{-3}$	21.54
sea ice	$1.5 \times 10^{-51}$	462.21	$6.7 \times 10^{-8}$	58.74

is that neighborhood pixels are expected to have similar characteristics. This method sets itself apart from the standard pattern recognition methods that do not inherently account for the image spatial interactions. As a result, these texture features and the corresponding segmentation algorithm are quite different in behavior than GLCP texture features combined with  $k$ -means clustering.

Let  $X$  represent an image. Then,  $X(i, j)$  represents a random variable at a site  $(i, j)$  on the  $M \times M$  lattice system  $S$ . For convenience,  $X(i, j)$  can be indexed as  $X_s$ ,  $s = 1, 2, \dots, M^2$ . A label represents a site distinction in the image defined on  $S$ . Let  $L$  be a set of discrete labels. For image segmentation, the challenge is to assign  $X_s$  an appropriate label from the set  $L$ .

For a discrete label set  $L$ , the probability that random variable  $X_s$  takes the value  $x_s$  is denoted  $P(X_s = x_s)$ , abbreviated  $P(x_s)$ , and the joint probability is denoted  $P(X = x) = P(X_1 = x_1, \dots, X_{M^2} = x_{M^2})$  and abbreviated  $P(x)$ .  $X$  is said to be an MRF on  $S$  with respect to a neighborhood system  $N$  if and only if  $P(x) > 0$  and  $P(x_s)$  is only dependent on its neighbors. The latter condition indicates that only neighboring sites have direct interactions with each other.

Under the assumption of a Gaussian MRF (GMRF), the following model is produced [21]:

$$x_s = \sum_{r \in N_s} \theta_r (x_{s+r} + x_{s-r}) + e_s \quad (1)$$

where  $x_s$  is a real number representing the center pixel of the neighborhood,  $x_{s+r}$  and  $x_{s-r}$  are a pair of pixels centered around  $x_s$ ,  $e_s$  is a zero-mean Gaussian noise, and  $\theta_r$  represents the MRF parameters, i.e., the parameters that describe the model and characterize the texture. The summation is over some neighborhood  $N_s$ , as defined by the order of the MRF model. The order of the MRF model is based on the distance from the center pixel. For example, as shown in Fig. 1, the MRF parameters weight the sum of a symmetrical pair of pixels.

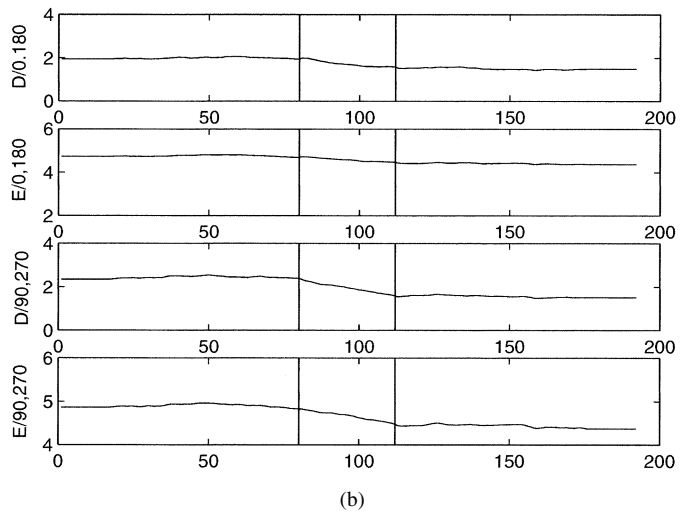
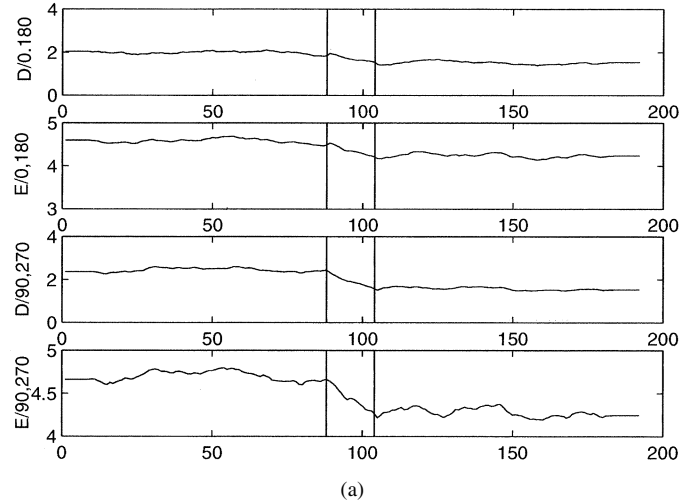


Fig. 5. Averaged GLCP texture features over 50 arbitrarily selected rows from Fig. 3(c) (between the two vertical lines are the boundary regions). (a)  $n = 16$ . (b)  $n = 32$ .

The two first-order parameters will weight the symmetrical neighbors at 1.x in Fig. 1. A second-order system would include all the neighbors marked by 1.x and 2.x, generating a total of four MRF parameters. A third-order system (all of 1.x, 2.x, and 3.x pairs) would require six MRF parameters, and so on.

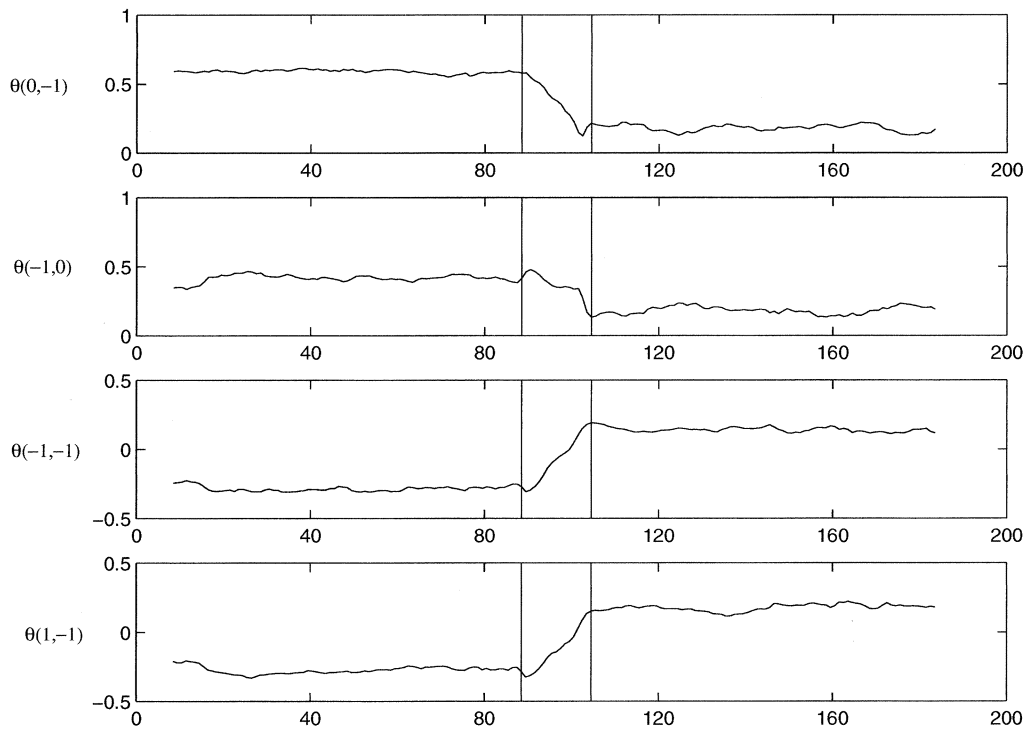
Selecting the model order for a given dataset is an important consideration. Here, the method developed by Kashyap and Chellappa [24] is used to estimate the model order for each dataset [25]. Using this method, the third-order GMRF model is used for synthetic texture, and the fourth-order GMRF model is used for Brodatz and sea ice textures.

For a given region, there are a number of ways to determine the GMRF parameters. Here, a least squares approach is employed. In matrix notation, (1) is indicated by

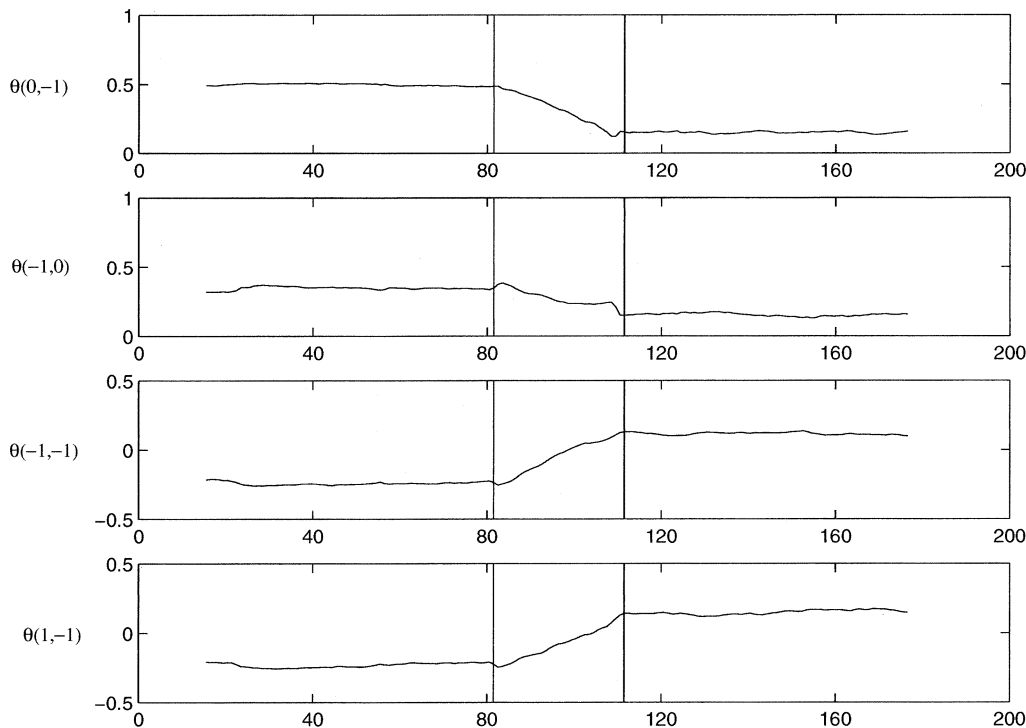
$$x_s = \theta^T Q_s + e_s \quad (2)$$

where  $\theta$  is a vector composed of all  $\theta_r$ , and  $Q_s$  is a vector defined by

$$Q_s = [x_{s+r_1} + x_{s-r_1}, \dots, x_{s+r_n} + x_{s-r_n}]^T. \quad (3)$$



(a)



(b)

Fig. 6. Averaged GMRF model parameters over 50 arbitrarily selected rows from Fig. 3(c) (between the two vertical lines are the boundary regions). (a)  $n = 16$ . (b)  $n = 32$ .

The least square estimation of the parameters is evaluated by

$$\hat{\theta} = \left[ \sum_{N_s} Q_s Q_s^T \right]^{-1} \left[ \sum_{N_s} Q_s x_s \right]. \quad (4)$$

For every pixel in a given window defined by  $N_s$ , a  $Q_s$  is determined. Then, for every window, a  $\theta$  is estimated that provides the GMRF features for that window.

The approach of model-based image segmentation is to articulate the regularities between neighboring pixels mathe-

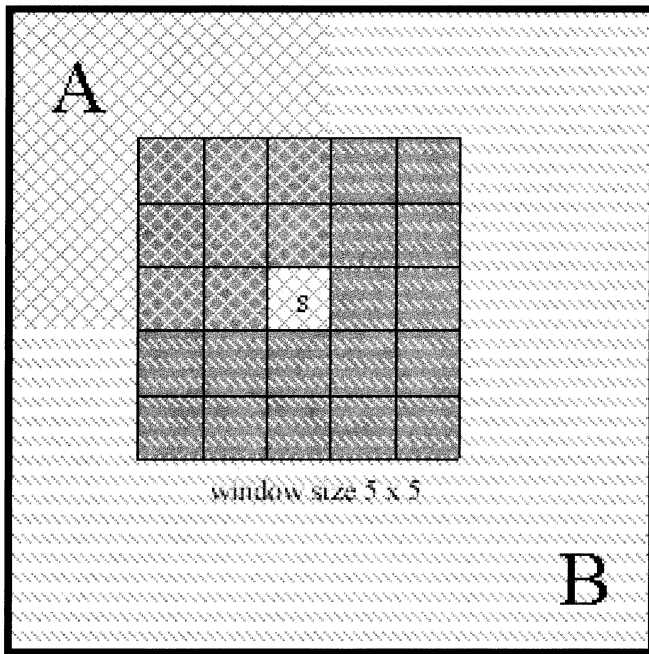


Fig. 7. Estimated texture parameters of pixel  $s$  using a  $5 \times 5$  window should be more similar to the parameters of texture B than A.

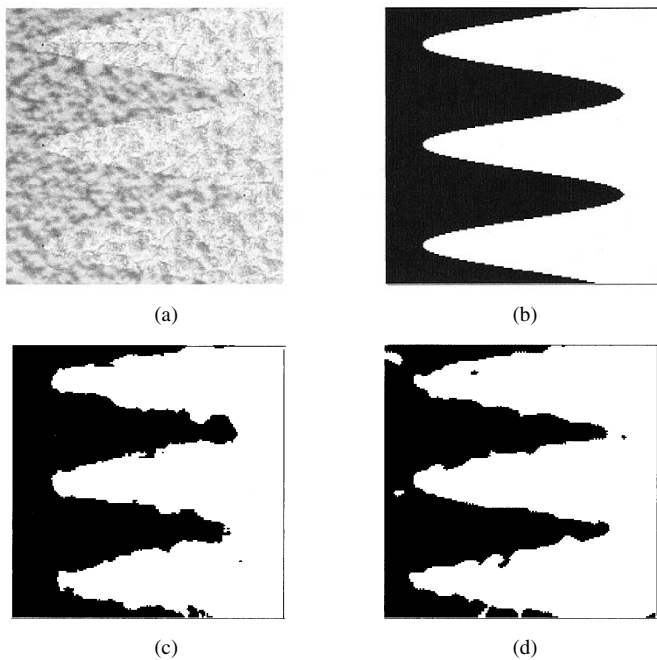


Fig. 8. Segmentation (using  $n = 16$ ) of Brodatz texture image with a three-cycle per image texture boundary. (a) Original image. (b) True segmentation. (c) GLCP result. (d) GMRF result.

matically, and then to exploit this information in a Bayesian framework to indicate which label has the highest probability. During the segmentation process, the prior distribution models are needed to capture the tendencies and constraints that characterize the scenes of interest. This method will inherently encourage nearest neighbor pixels to have a higher probability of belonging to the same class.

For the segmentation problem, image  $X$  is composed of two components, namely a pixel intensity array and a corresponding

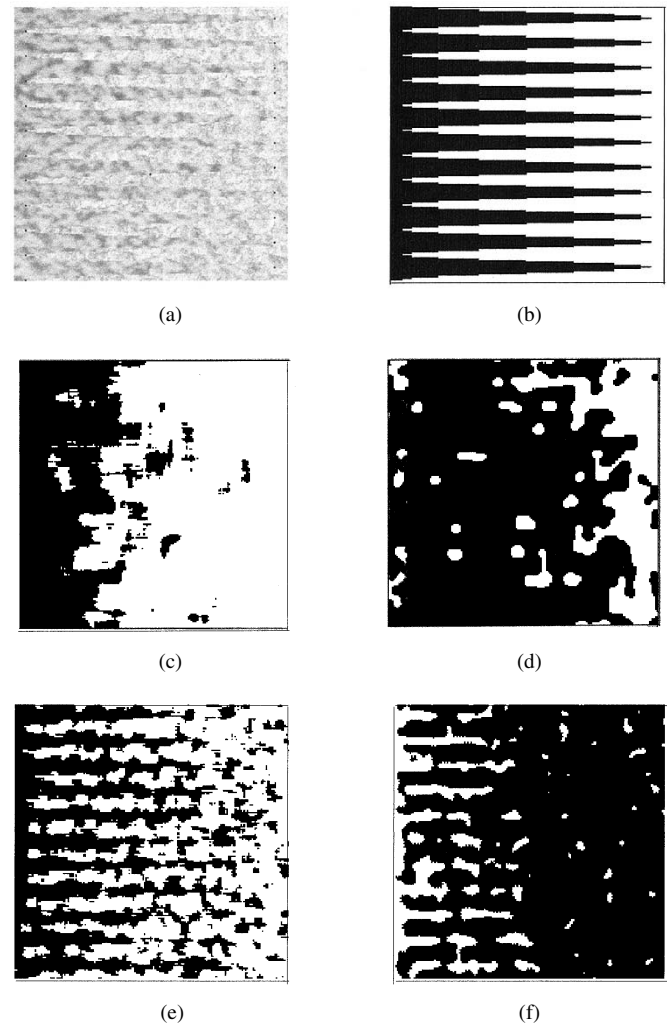


Fig. 9. Segmentation of Brodatz texture image with an 11-cycle per image texture boundary. (a) Original image. (b) True segmentation. (c) GLCP result ( $n = 16$ ). (d) GMRF result ( $n = 16$ ). (e) GLCP result ( $n = 8$ ). (f) GMRF result ( $n = 8$ ).

texture label array, i.e.,  $X = \{X^I, X^L\}$ . A Bayesian model is used to describe the probability of a certain label for a pixel given the pixel intensity. The iterated conditional mode (ICM) method is used to perform the GMRF image segmentation [26]. For succinctness, details of this procedure are not detailed here. Further details of the MRF models and Bayesian segmentation procedure are found in [25]. Briefly, the segmentation proceeds in two stages. First, the GMRF parameters must be estimated. This must be done because the GMRF parameters are dependent on the image data assigned to a particular label, and this information is unknown *a priori*. So, to estimate the GMRF parameters, the image is divided into nonoverlapping windows. For each window, the GMRF parameters are determined using (4). Then,  $k$ -means is used to group like estimates and produce an estimate for a given label. Second, image pixels are assigned to a particular label based on a Gibbs estimate. The estimate is based on a Gaussian component (requiring the local image data) and a label term (based on the neighborhood label assignments). This is performed in an iterative fashion (over each pixel in the image) until a stopping criteria or some minimum energy criterion is met.

### III. RESEARCH QUESTIONS

The GLCP method is the preferred method for texture feature extraction for SAR sea ice imagery [1], [3], [27]. However, to obtain an optimal solution, exhaustive combinations of quantizations, displacements, window sizes, orientations, and statistics must be performed. Brodatz imagery is often used as a test base for texture analysis algorithms in the research literature. Encouraging results have been obtained for unsupervised texture segmentation using Brodatz textures [28], [29]. Unfortunately, the texture appearance of a consistent ice type is not as regular as a Brodatz texture, and most of the time, different ice types as well as open water are interwoven with each other. Can GMRF methods be used in SAR sea ice image segmentation? Which of GLCP and GMRF produce a better segmentation result? There exists limited published research investigating the potential of GMRF methods in SAR sea ice image segmentation. Although many efforts consider the supervised classification potential of GLCPs, limited research investigating its unsupervised segmentation potential are known.

An important aspect of texture segmentation (using GLCPs or GMRFs) is the window coverage. Large windows produce better estimates of the texture; however, in the case of segmentation, large windows can also lead to the undesirable situation of multiple texture classes within a common window. Small windows are less likely to contain multiple classes; however, the limited coverage may produce misleading features. These issues are investigated within the context of the following research questions.

#### A. Research Questions

1) *Influence of Window Size:* How does window size influence the estimated individual GLCP texture features and GMRF model parameters?

The same window size ( $n$ ) can be used to determine the GLCP texture features as well as the GMRF model parameters. For both methods, the size of  $n$  is a very important parameter, and a different number of pixels in  $n$  can generate different estimation results. It is necessary to evaluate the effect of window size  $n$  on the stability of each estimated GLCP texture feature and GMRF model parameter. Conclusions with regard to this research question can assist the selection of a suitable window size for each method.

2) *Cluster Separability:* How does the window size  $n$  influence the separability of the clusters of the estimated GLCP texture features and GMRF model parameters?

This research question compares GLCP texture features versus GMRF model parameters for feature space separability. For a given pair of textures, the feature space separability for each of GLCP and GMRF features provides a means to evaluate their texture identification ability. This evaluation will act as a basis to help understand the cause of an ineffective segmentation, presented later in the paper.

From this second research question, another issue can also be addressed, i.e., given a sufficiently large window size, which method is better at characterizing texture? The question is inspired by the fundamental difference of statistical-based and

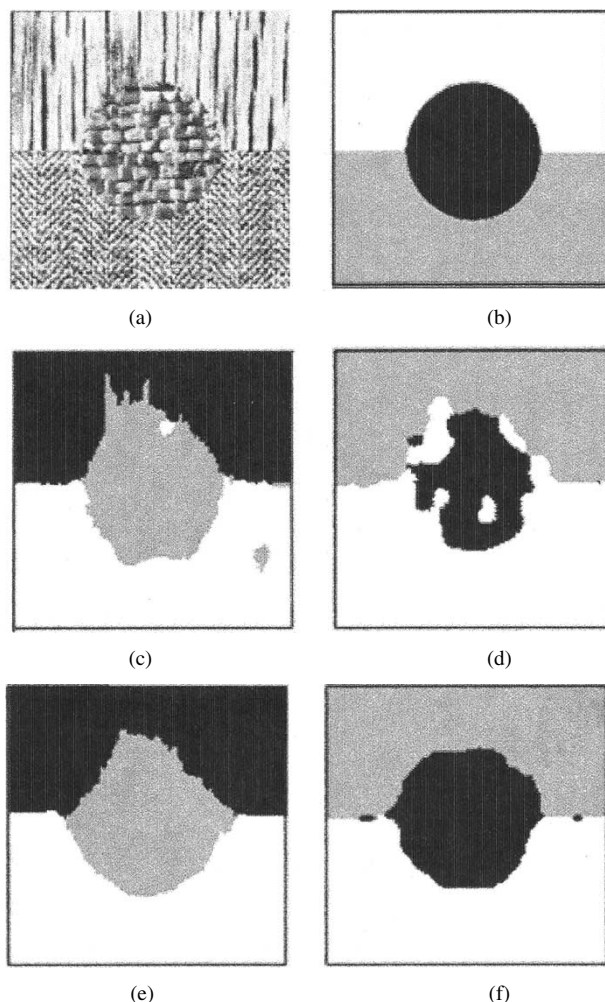


Fig. 10. Segmentation of Brodatz texture image. (a) Original image. (b) True segmentation. (c) GLCP result ( $n = 16$ ). (d) GMRF result ( $n = 16$ ). (e) GLCP result ( $n = 32$ ). (f) GMRF result ( $n = 32$ ).

model-based texture analysis techniques. MRF texture models are associated with the natural formation process of the analyzed textures. The model-based texture analysis technique is able to describe and reconstruct image textures. But, there is no known method to synthesize textures based on the GLCP texture features. From this point, it is a logical inference that the MRF method is expected to contain more complete texture structure information for distinguishing various textures than the GLCP method.

3) *Multiple Textures and Irregular Boundaries:* Given a window, what is the effect on the estimated GLCP texture features and GMRF model parameters if the window contains multiple textures with the possibility of irregular boundaries?

The first and second research questions deal with homogeneous texture samples. For the segmentation problem, some of the selected texture samples will contain multiple textures. The larger a window, the greater the probability that various classes will be included in it. In this case, the estimation of GLCP texture features or GMRF model parameters may be damaged. What is the relationship of the estimated texture features from the multitexture window with the texture features of each texture in the window?

#### IV. METHODS AND TESTING

##### A. Research Question 1: Influence of Window Size

1) *Method for Research Question 1: Influence of Window Size:* How does the window size influence the estimated individual GLCP texture features and GMRF model parameters?

For a given texture, the relative change of a feature's standard deviation as a function of window size  $n$  is calculated. The experiment is designed to check the effect of different window sizes on the stability of the estimation in each dimension for each method.

One synthetic texture ( $1024 \times 1024$ ) [generated using the model parameters "A" in Table I and illustrated in Fig. 2(a)] one Brodatz texture ( $1024 \times 1024$ ) (pigskin) (illustrated on the right-and side of Fig. 3(b), and one SAR sea ice image ( $768 \times 768$ ) [the multiyear component of Fig. 2(c)] are used for the experiments. The SAR image, as obtained from the Canadian Ice Services (CIS), was captured from the Beaufort Sea, October 13, 1997, using Radarsat ScanSAR SW1 mode ( $2 \times 2$  block averaged to 100-m pixels). Note that the SAR images used in this paper have accompanying operational ice charts produced at CIS. This information is used to determine ice types in the imagery. From each texture image, 60 window samples are randomly selected for each window size ( $n = 8, 16, 32, 64, 96$ ). The standard deviation  $\sigma$  per feature per window size is determined for each set of 60 samples. To measure the relative change, the value of  $\sigma$  is normalized by the  $\sigma$  of the 96 window size for a given feature.

2) *Results for Research Question 1: Influence of Window Size:* Representative plots are provided in Fig. 4 using the pigskin texture of Brodatz image (each dataset produced similar results). These plots show relative changes of  $\sigma$  for each estimated GLCP texture feature and each GMRF model parameter from window size 96 to window size 8. Fig. 4(a) plots six representative GLCP texture features, and Fig. 4(b) plots the six third-order GMRF model parameters. A pair of observations are noted. First, larger windows lead to more stable texture features estimates than smaller windows. Similarly, with decreasing window size, the  $\sigma$  of each estimated GLCP texture feature and the GMRF model parameter increase exponentially. Second, with decreasing window size, the  $\sigma$  of each GMRF model parameter increases faster than the GLCP texture features, especially from window size 32 to window size 8.

These observations are supported by the values in Table II, which indicates the average standard deviation ratios of each estimation from all three textures used in this experiment. For example, using the GLCP method, the largest average standard deviation ratio is 9.33 using a synthetic texture based on comparing window sizes 8–96, whereas, the largest ratio using the GMRF method is 28.32, which also occurred using the synthetic texture and the window size 8 compared to 96. The GMRF's average standard deviation ratio of the smaller window sizes (64, 32, 16, 8) to the larger window size (96) always exceeds that of the GLCPs. Positive and negative values of the GMRF model parameter behave quite differently in the texture formation process. In this case, large standard deviations could significantly damage the texture model estimation.

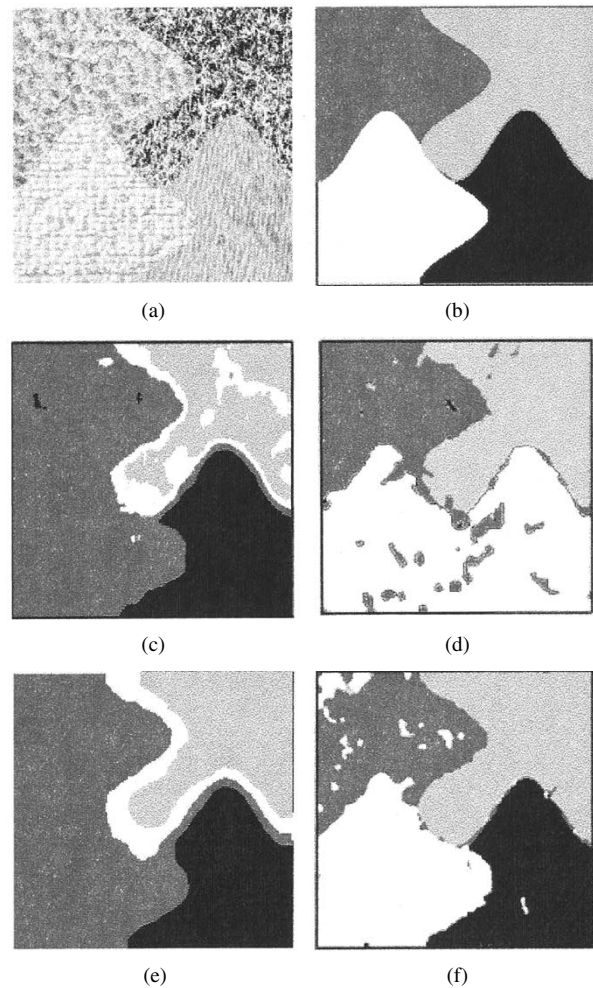


Fig. 11. Segmentation of Brodatz texture image. (a) Original image. (b) True segmentation. (c) GLCP result ( $n = 16$ ). (d) GMRF result ( $n = 16$ ). (e) GLCP result ( $n = 32$ ). (f) GMRF result ( $n = 32$ ).

Based on this experiment, to obtain a stable estimation, the GMRF method requires a relatively larger window size than the GLCP to obtain the same degree of stability in the feature estimates. It is noted here that textures are generally multiscale and may require the use of more than one window size to perform segmentation properly. For brevity, only the relative ability of the GLCP and GMRF methods will be investigated in this paper.

##### B. Research Question 2: Cluster Separability

How does the window size  $n$  influence the separability of the clusters of estimated GLCP texture features and GMRF model parameters?

1) *Method for Research Question 2: Cluster Separability:* This research question explores the effect of different window sizes on the multidimensional feature space separability. The Fisher criterion [20] is used as a nonparametric measure of the separability of two class clusters in the feature space. This criterion calculates a ratio of the between-class separability and the within-class variation. Larger Fisher criterion values demonstrate improved separation of two classes. Further insight can be obtained by calculating the upper bound of classification error between feature cluster pairs using the Bhattacharyya error bound [20]. In this case, smaller errors



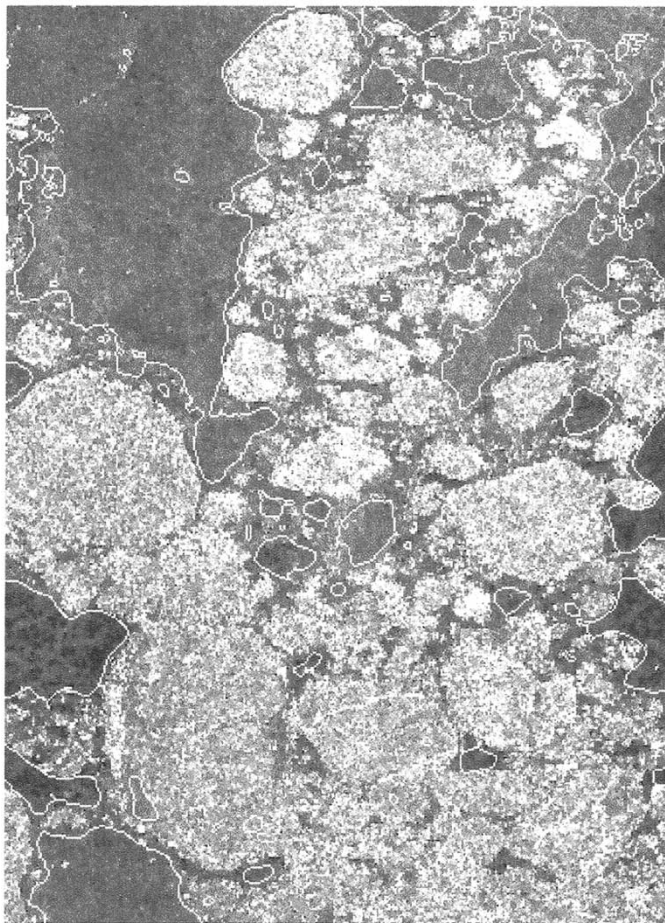


Fig. 12. GLCP segmentation result for STAR-1 image.

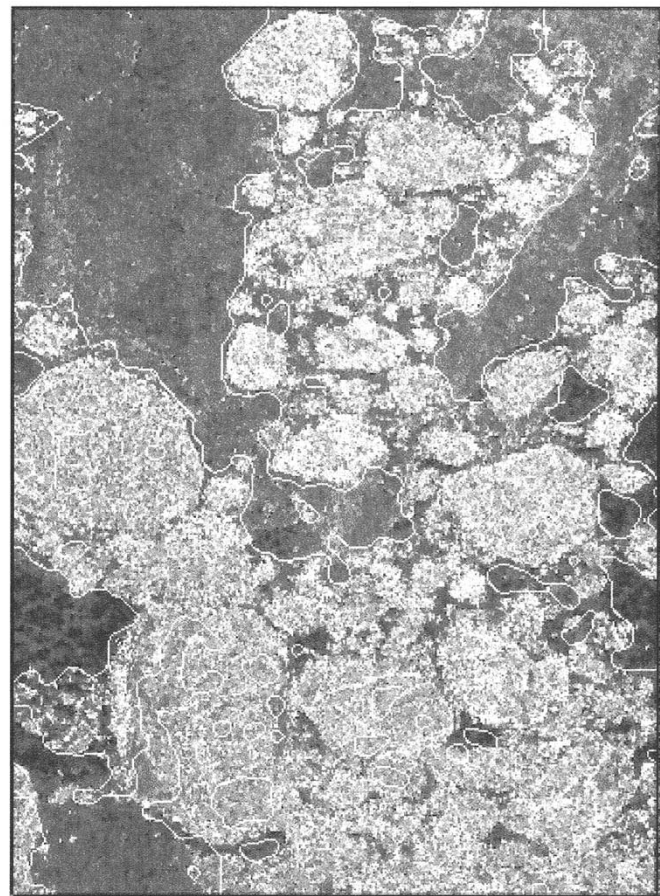


Fig. 13. GMRf segmentation result for the STAR-1 image.

indicate better separability of two class clusters in the feature space.

Three texture image pairs are used for testing, as illustrated in Fig. 2. These include synthetic ( $1024 \times 1024$ ) (models A and B in Table I), Brodatz ( $1024 \times 1024$ ) (wood grain and raffia), and SAR sea ice texture images ( $768 \times 768$ ). The multiyear ice is obtained from the same image used in Research Question 1. The first-year sample uses the same modality and processing but was captured from the Gulf of St. Lawrence on February 19 1997. Sixty sample windows with sizes 8, 16, 32, and 64 are randomly selected from each image. The two images in each pair represent internally consistent but mutually distinct texture patterns.

2) *Results for Research Question 2: Cluster Separability:* Table III reports the Bhattacharyya error bounds (BEBs) of each texture pair and the Fisher criterion ( $J$ ) of each projected texture pair. A number of observations can be made. When  $n = 8$ , all of the GLCP pairs have a lower Bhattacharyya bound as well as a higher Fisher criterion compared to the GMRf pairs. In contrast, for  $n = 16$  and  $n = 32$ , the GMRf has lower Bhattacharyya bound and higher Fisher criterion. Separability is relatively strong given smaller windows for GLCP texture features compared to GMRf texture features. However, for large windows, GMRf features are more separable than GLCP features. Corresponding line projections (for succinctness, these are not displayed here) visibly validate

these numerical results. Testing for  $n = 64$  produces similar results as for  $n = 32$ .

Based on this experiment, another conclusion in the second research question can be obtained. With a large window size, the GMRf model and GLCP texture features have capable separability given homogeneous samples. The BEBs are all at most on the order of  $10^{-3}$ , indicating that the homogeneous  $32 \times 32$  windows contain sufficient data to distinguish the given class pairs.

### C. Research Question 3: Multiple Textures and Irregular Boundaries

Given a window, what is the effect on the estimated GLCP texture features and GMRf model parameters if the window contains multiple textures with the possibility of an irregular boundary?

1) *Method for Research Question 3: Multiple Textures and Irregular Boundaries:* Features for windows containing multiple textures can produce misleading segmentation results. Based on the calculation of GLCP texture features and GMRf model parameters, a reasonable assumption is that the extracted GLCP or GMRf texture features from a multitexture window should have a linear weighting proportional to the area coverage of individual texture features. For example, given two texture types A and B in window  $n$ , then  $F = a \times F^A + b \times F^B$ , where  $F$  is any texture feature of window  $n$ ,  $a$  is the area proportion

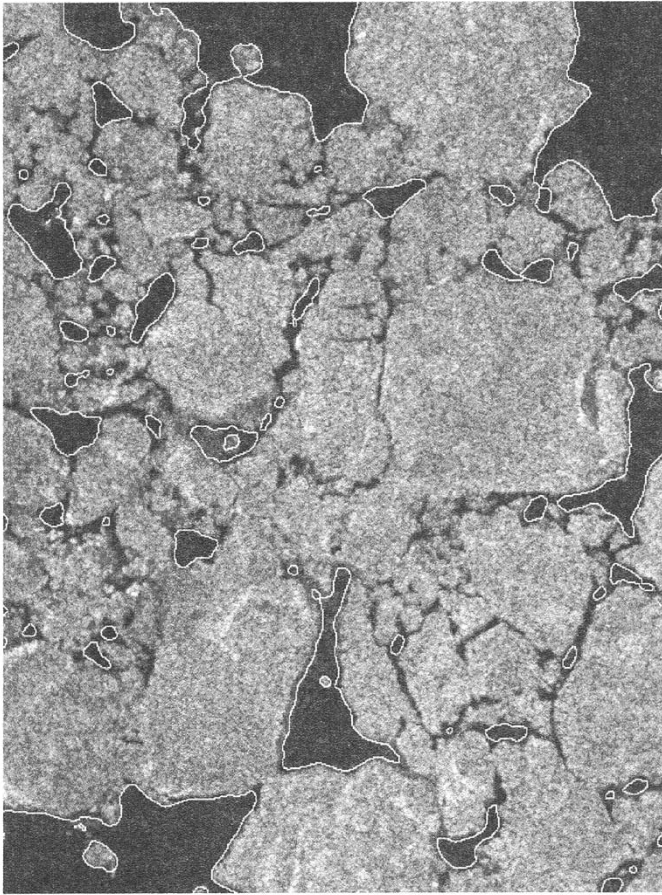


Fig. 14. GLCP segmentation result for Radarsat image captured over the Gulf of St. Lawrence (February 19, 1997).

of texture A in window  $n$ ,  $b$  is the area proportion of texture B in window  $n$ , and  $a + b = 1$ .

2) *Results for Research Question 3: Multiple Textures and Irregular Boundaries:* Fig. 3(a)–(c) shows three bipartite texture images, each  $256 \times 256$ : synthetic, Brodatz, and SAR sea ice, respectively. Texture estimates for each image are estimated based on two window sizes ( $n = 16$ ,  $n = 32$ ) for samples centered at each pixel across 50 randomly selected rows. Selected results are only presented for the sea ice image [Fig. 3(c)] because results for other images are similar in nature. Figs. 5 and 6 depict the average feature values for the 50 samples for each of GLCP and GMRF texture features, respectively. The vertical line pairs represent the window extent when centered on the boundary of the two textures. The resulting parameters change in approximately a linear manner from one texture to another.

This experiment demonstrates that texture measures using either method can be damaged by the multitexture windows. Given a window with multiple textures, each estimated GLCP texture feature seems to be a weighted linear combination of the corresponding GLCP texture features of each texture. The same conclusion applies to each estimated GMRF model parameter.

Consequently, significant segmentation error can occur, as illustrated by Fig. 7. The boundary between textures A and B is a right angle. Nine pixels in a  $5 \times 5$  window are located on pixel  $s$ , and 16 pixels belong to texture B. Pixel  $s$  belongs to texture A.

Based on the above experiment, for both methods, the derived features will be closer to texture B than to texture A, encouraging pixel  $s$  to be misclassified as texture B. In this case, the corner pixel  $s$  would be eroded, since it would be erroneously classified as texture B. This example illustrates that textured images with irregular boundaries will strongly affect the performance of the texture segmentation methods. Commonly, texture feature images in the research literature utilize straight boundaries, which can provide misleading results if extrapolated to natural imagery with irregular texture boundaries.

How do these damaged estimations affect the segmentation results for each method? Figs. 8 and 9 show the segmentation results of two bipartite images with different texture boundaries. To concentrate on the effect of the textural boundary on the segmentation, each image contains the same two Brodatz textures (paper and pigskin). In the case of a straight vertical boundary given both  $n = 8$  and  $n = 16$ , GLCP and GMRF texture features and their associated segmentation schemes successfully segment the image (not shown for brevity).

Fig. 8(a) has a sinusoidal boundary with three periods, which makes it difficult to distinguish compared to a straight boundary. For each GLCP and GMRF, a  $16 \times 16$  window size is used, and the segmentation result is effective. By increasing the frequency of the sinusoidal boundary, a more difficult segmentation problem is produced (see Fig. 9). The segmentation results for both GLCP and GMRF using a  $16 \times 16$  window are unsuccessful [see Fig. 9(c) and (d)]. Fig. 9(e) and (f) shows the segmentation results using an  $8 \times 8$  window size. In this case, the GLCP method produces a much better segmentation than the GMRF method. Also,  $n = 8$  generates a much better result than  $n = 16$ , given the GLCP texture features.

For multiple texture images with simple boundaries, sample windows will contain single textures more often than such images with complex boundaries. For simple boundaries, both the GLCP and GMRF methods can produce correct feature estimates, which leads to an accurate segmentation. For complex boundaries, both methods may have damaged features estimates that can erode the quality of the segmentation. To minimize the effect of multiple texture windows, small windows should be used. In such cases, the GLCP method should be employed, as supported by the first two research questions.

#### D. Image Segmentation Results

In this section, the GLCP and GMRF methods will be applied to the segmentation of two Brodatz and three SAR sea ice images. Note that the Brodatz images contain boundaries that are simple relative to the boundaries found in SAR sea ice imagery. The first example is a three-class Brodatz image [Fig. 10(a)]. The true segmentation is illustrated in Fig. 10(b). Given a window size of  $n = 16$ , GLCPs [Fig. 10(c)] are better able to segment the image than GMRFs [Fig. 10(d)]. When the window size is increased to  $n = 32$ , GMRFs generate a marginally better segmentation [Fig. 10(f)] than the GLCPs [Fig. 10(e)]. Both methods produce improved segmentations when the window size is increased; however, the relative improvement for the GMRF method is significant. The GMRF method requires a larger spatial extent to estimate the model

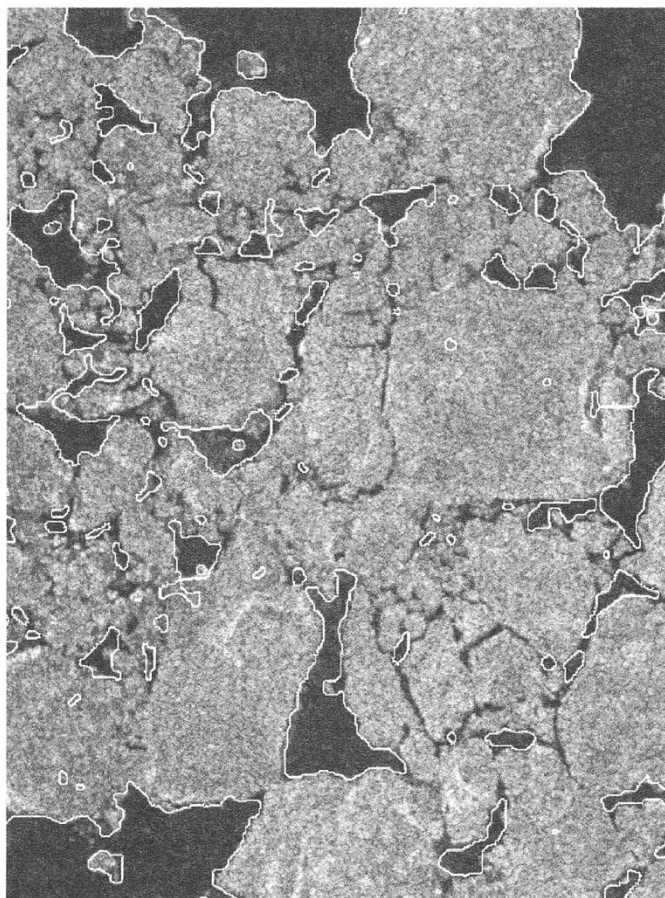


Fig. 15. GMRF segmentation result for a Radarsat image captured over the Gulf of St. Lawrence (February 19, 1997).

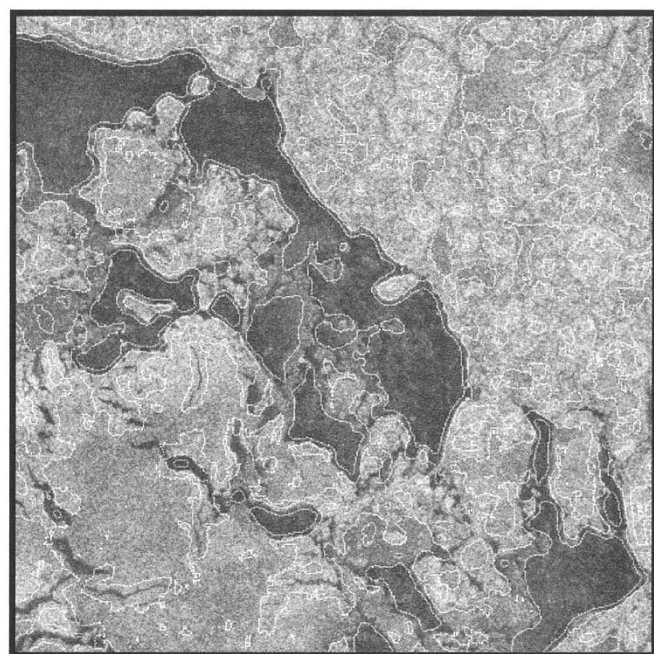


Fig. 16. GLCP segmentation result for a Radarsat image captured over the Beaufort Sea (October 13, 1997).

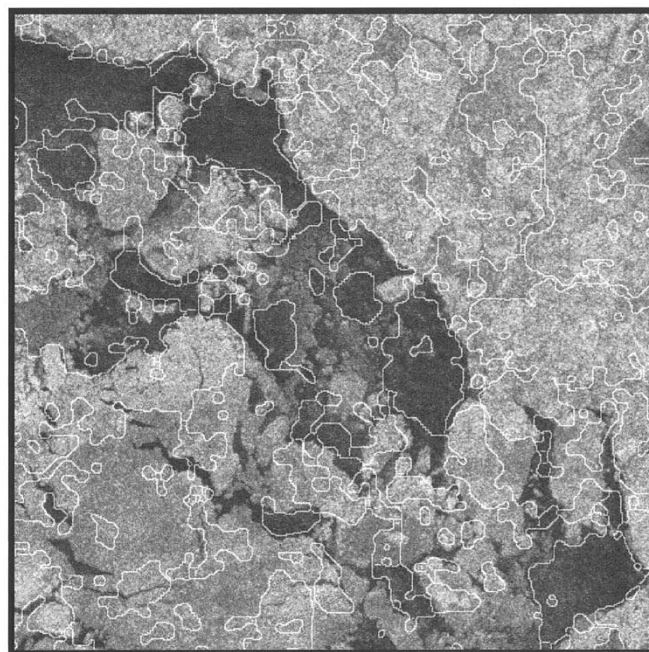


Fig. 17. GMRF segmentation result for a Radarsat image captured over the Beaufort Sea (October 13, 1997).

$n = 32$ , the GMRF method has a better approximation for the curved boundary.

The second Brodatz image [Fig. 11(a)] contains four textures separated by horizontal and vertical sinusoidal boundaries. Once again, the larger window size ( $n = 32$ ) allows the GMRF method to produce a reasonable image segmentation [Fig. 11(f)] compared to a window size of  $n = 16$  [Fig. 11(d)]. The GLCPs are not able to properly identify the boundary region between the textures [Fig. 11(c) and (e)]. The GMRF segmentation procedure produces a superior result for this case. The GMRF segmentation process is advantageous because it uses a label image model that encourages neighbor pixels to be assigned to the same class.

From the earlier testing using different window sizes (see Table II), one can see that both methods prefer a larger window size to obtain a robust estimation. But, the larger window size may cause a segmentation problem using the GLCP method in the boundary area, i.e., the true boundary between textures may be blurred, and sometimes, the pixels along the boundary areas could be misclassified as another texture class. To overcome this boundary problem, the window size should be as small as possible for the GLCP method. Using the GMRF method, such a small window size may compromise the quality of the texture features. As a result, the GMRF window size should be as large as possible. As discussed previously, given complex texture boundaries, a large window could also damage the estimated parameters.

The last three segmentation results are based on SAR sea ice images. Because the sea ice texture has a finer appearance than Brodatz textures, for these images, the window size used in the GLCP method is  $13 \times 13$ , and the window size used in the GMRF method is  $16 \times 16$ . The full sea ice image presented in Figs. 12 and 13 was first published in [31]. The image is

parameters and perform an accurate segmentation. Also, for

a STAR-1 X-HH image (6-m resolution). Here, a  $336 \times 466$  subimage of the full image is used. Two ice types are identified: first-year ice and multiyear ice. Both methods can distinguish the two ice types. The boundaries generated from the GMRF method are more accurate than the GLCP method. Visually observing the boundary locations in these two figures, one can see that the edges determined by the GMRF method present more appropriately the boundaries of different ice types. However, the GMRF method has problems to properly discriminate the multiyear ice in the floes close to the bottom of the image.

The second SAR sea ice image is a Radarsat ScanSAR SW1 C-band HH image obtained from CIS. The pixel spacing is 100 m ( $2 \times 2$  block averaged), and the image size is  $495 \times 669$ . The scene was captured over the Gulf of St. Lawrence on February 19, 1997, in a descending orbit. The image segmentations are found in Figs. 14 (GLCP) and 15 (GMRF). Both methods are readily able to perform the segmentation; however, the GMRF method has relatively better boundary approximations. Many of the thin first-year ice regions separating the multiyear floes are not recognized, due to poor features estimates for such regions.

The third sea ice image was also obtained from CIS. The image has the same parameters as the previous SAR image; however, this image was obtained October 13, 1997, in an ascending orbit over the Beaufort Sea and is  $486 \times 486$ . Edge maps are displayed in Figs. 16 (GLCP) and 17 (GMRF), and region maps are displayed in Figs. 18 (GLCP) and 19 (GMRF). Such a SAR image illustrates a more difficult task for an automated segmentation algorithm for operational ice reconnaissance. Multiyear ice, new ice, and gray ice are contained in the image. The new ice (dark tone) can be found in the middle of the image. Gray ice is found in the top and bottom regions of the image. The multiyear ice floes with a brighter tone than gray ice are mingled with the gray ice and can be found primarily in the top region.

The GLCP method and GMRF methods generate different segmentation results, and both methods do not perform as well as the first two SAR images. Generally speaking, the GLCP method generates a better segmentation image than GMRF method. The GMRF method tends to underrepresent the amount of new ice, confusing it with gray ice. Relatively, the GLCP segmentation generates a better segmentation of the new ice region. In distinguishing multiyear ice and gray ice, the two methods have a different tendency. The GMRF method is unable to properly capture locally uniform texture characteristics, and the segmentation result appears more discontinuous than GLCP method. The GLCP method tends to confuse the multiyear ice with gray ice on the top region of the image, delineating the gray ice more than it should. Note that the human observer would tend to identify the gray ice in the bottom left as gray ice, due to the nature of the shape of the leads in that region. Younger ice (such as gray ice) tends to produce leads with the observed characteristic shapes, providing an additional basis for manual ice classification.

Comparing the three ice images, one can notice that the pattern of each ice type in the first two SAR images is more accurately defined than the pattern in the third image. Also, the

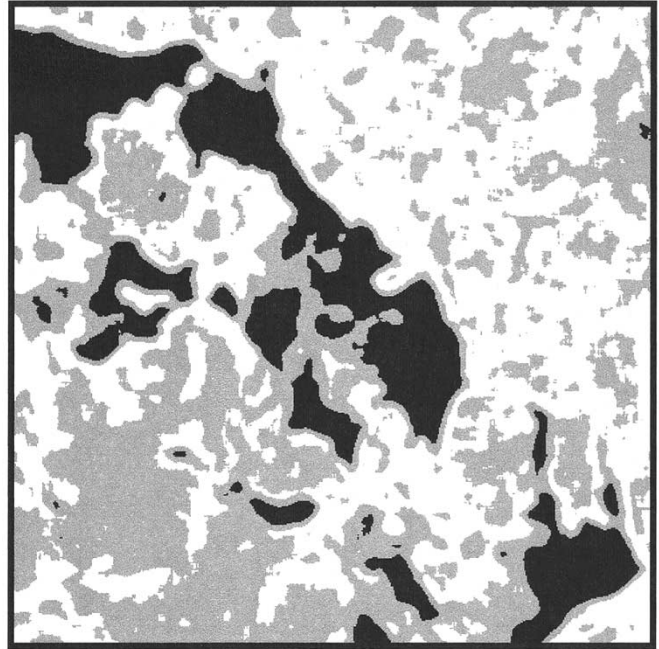


Fig. 18. GLCP segmentation result with regions identified for a Radarsat image captured over the Beaufort Sea (October 13, 1997).

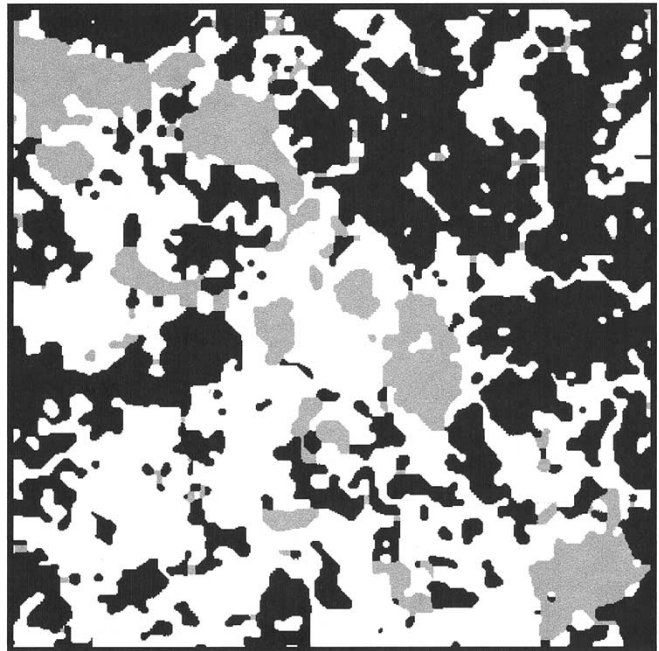


Fig. 19. GMRF segmentation result with regions identified for a Radarsat image captured over the Beaufort Sea (October 13, 1997).

texture patterns of the ice types in the first two images have a more visually distinct difference than those in the third image. Finally, the multiyear ice and gray ice in the third SAR image have considerable mixing, leading to each having smaller local image extent. All of these could result in a poorer model estimation and make the third image more difficult to segment using the GMRF method. It is possible that the pixel resolution of the Radarsat image (100 m) contributes to the poorer segmentation

result relative to the STAR-1 image (6-m resolution). Operationally, it is more cost-effective to utilize the Radarsat imagery, and algorithms should be developed in support of this platform.

## V. SUMMARY

There exists a lack of published research comparing unsupervised texture segmentation methods. Gray-level cooccurrence probabilities and Gaussian Markov random fields are two popular texture methods. GLCPs are commonly applied to remote sensing imagery in the research literature, while GMRFs have not been used extensively for this purpose. The goal of this paper was to develop a better understanding of the ability of each method in the unsupervised segmentation of SAR imagery by considering the role of window size in the segmentation process.

A number of research questions were posed and studied, producing the following results. GMRFs require increased window sizes relative to GLCPs to produce stable texture estimates. GLCPs produce more separable texture features for smaller windows than the GMRF texture features. Across three different datasets (synthetic, Brodatz, and SAR), a window size of 32 was deemed sufficiently large to obtain separable, consistent texture features. However, such a large window size can lead to segmentation error due to the higher risk of multiple classes appearing in the same window. Given a window with multiple textures, a linear weighting of the ratio of the textures in the image generates the overall texture feature. Such a weighting can lead to erroneous boundary delineation and can even misclassify the boundary itself as an incorrect class. The segmentation of classes separated by irregular boundaries found in remote sensing imagery will be affected by this problem. The texture literature often utilizes convenient texture boundaries, yet complex boundaries, coupled with varying local class spatial extents, pose greater challenges to the operational use of image segmentation algorithms.

## ACKNOWLEDGMENT

This work is part of an ongoing collaboration with CIS (<http://www.cis.ec.gc.ca/home.html>). CIS is thanked for encouragement as well as the provision of data. The authors appreciate the assistance of M. Arkett, R. De Abreu, D. Flett, M. Manore, B. Ramsay, and K. Wilson. Thanks to the anonymous reviewers, whose helpful advice improved the quality of the manuscript.

## REFERENCES

- [1] Q. A. Holmes, D. R. Nuesch, and R. A. Shuchman, "Textural analysis and real-time classification of sea-ice types using digital SAR data," *IEEE Trans. Geosci. Remote Sensing*, vol. GE-22, pp. 113–120, Mar. 1984.
- [2] M. E. Shokr, "Evaluation of second-order texture parameters for sea ice classification from radar images," *J. Geophys. Res.*, vol. 96, no. C6, pp. 10 625–10 640, 1991.
- [3] D. G. Barber and E. F. LeDrew, "SAR sea ice discrimination using texture statistics: A multivariate approach," *Photogramm. Eng. Remote Sens.*, vol. 57, no. 4, pp. 385–395, 1991.
- [4] M. Tuceryan and A. K. Jain, "Texture analysis," in *Handbook of Pattern Recognition and Computer Vision*, C. Chen, L. Pau, and P. Wang, Eds. Singapore: World Scientific, 1993, ch. 2.1, pp. 235–276.

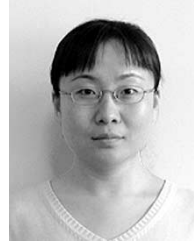
- [5] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for images classification," *IEEE Trans. Syst., Man Cybern.*, vol. 3, no. 6, pp. 610–621, 1973.
- [6] J. E. Besag, "Spatial interaction and the statistical analysis of lattice systems (with discussion)," *J. R. Statist. Soc. B*, vol. 36, no. 2, pp. 192–236, 1974.
- [7] A. K. Jain and F. Farrokhnia, "Unsupervised texture segmentation using Gabor filters," *Pattern Recognit.*, vol. 24, no. 12, pp. 1167–1186, 1991.
- [8] D. A. Clausi and M. E. M. E. Jernigan, "Establishing Gabor filter parameters for optimal texture separability," *Pattern Recognit.*, vol. 33, no. 11, pp. 1835–1849, 2000.
- [9] T. Randen and J. H. Husoy, "Filtering for texture classification: A comparative study," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 21, pp. 291–310, Apr. 1999.
- [10] P. Brodatz, *Textures: A Photographic Album for Artists and Designers*. New York: Dover, 1966.
- [11] O. Pichler, A. Teuner, and B. J. Hosticka, "A comparison of texture feature extraction using adaptive Gabor filtering, pyramidal and tree structured wavelet transforms," *Pattern Recognit.*, vol. 29, no. 5, pp. 733–742, 1996.
- [12] T. Ojala, M. Pietikainen, and D. Harwood, "A comparative study of texture measures with classification based on feature distributions," *Pattern Recognit.*, vol. 29, no. 1, pp. 51–59, 1996.
- [13] P. P. Ohanian and R. C. Dubes, "Performance evaluation for four classes of textural features," *Pattern Recognit.*, vol. 25, no. 8, pp. 819–833, 1992.
- [14] R. W. Connors and C. A. Harlow, "Theoretical comparison of texture algorithms," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-2, pp. 204–222, May 1980.
- [15] J. S. Weszka, C. R. Dyer, and A. Rosenfeld, "A comparative study of texture measures for terrain classification," *IEEE Trans. Syst., Man Cybern.*, vol. 6, no. 4, pp. 269–285, 1976.
- [16] J. R. Carr and F. P. de Miranda, "The semivariogram in comparison to the co-occurrence matrix for classification of image texture," *IEEE Trans. Geosci. Remote Sensing*, vol. 36, pp. 1945–1952, Nov. 1998.
- [17] P. Gong, D. J. Marceau, and P. J. Howarth, "A comparison of spatial feature extraction algorithms for land-use classification with SPOT HRV data," *Remote Sens. Environ.*, vol. 40, pp. 137–151, 1992.
- [18] D. A. Clausi, "Comparison and fusion of co-occurrence, Gabor and MRF texture features for classification of SAR sea-ice imagery," *Atmos.-Ocean*, vol. 39, no. 3, pp. 183–194, 2001.
- [19] —, "An analysis of co-occurrence texture statistics as a function of grey level quantization," *Can. J. Remote Sens.*, vol. 28, no. 1, pp. 45–62, 2002.
- [20] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. New York: Wiley, 2001.
- [21] R. Chellappa and S. Chatterjee, "Classification of texture using Gaussian Markov random fields," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-33, pp. 959–963, Aug. 1985.
- [22] F. S. Cohen, Z. Fan, and M. A. Patel, "Classification of rotated and scaled texture images using Gaussian Markov random fields," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, pp. 192–203, Feb. 1991.
- [23] S. Geman and C. Graffigne, "Markov random field image models and their applications to computer vision," in *Proc. Int. Congr. Mathematicians*, 1986.
- [24] R. L. Kashyap and R. Chellappa, "Estimation and choice of neighbors in spatial-interaction models of images," *IEEE Trans. Inform. Theory*, vol. 29, no. 1, pp. 61–72, 1983.
- [25] B. Yue, "SAR sea ice recognition using texture methods," Master's thesis, Univ. Waterloo, Waterloo, ON, Canada, 2001.
- [26] J. Besag, "On the statistical analysis of dirty pictures," *J. R. Statist. Soc. B*, vol. 48, pp. 259–302, 1986.
- [27] L. K. Soh and C. Tsatsoulis, "Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices," *IEEE Trans. Geosci. Remote Sensing*, vol. 37, pp. 780–795, Mar. 1999.
- [28] B. S. Manjunath and R. Chellappa, "Unsupervised texture segmentation using Markov random field models," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 13, pp. 478–482, May 1991.
- [29] H. Derin and W. S. Cole, "Segmentation of textured images using Gibbs random fields," *Comput. Vis., Graph. Image Process.*, vol. 35, pp. 72–98, 1986.
- [30] S. Krishnamachari and R. Chellappa, "Multi-resolution Gauss-Markov random field models for texture segmentation," *IEEE Trans. Image Processing*, vol. 6, pp. 251–267, Feb. 1997.
- [31] D. G. Barber, M. E. Shokr, R. A. Fernandes, E. D. Soulis, D. G. Flett, and E. F. LeDrew, "A comparison of second-order classifiers for SAR sea ice discrimination," *Photogramm. Eng. Remote Sens.*, vol. 59, no. 9, pp. 1397–1408, 1993.



**David Clausi** (S'93–M'96–SM'03) received the B.A.Sc., M.A.Sc., and Ph.D. degrees from the University of Waterloo, Waterloo, ON, Canada, in 1990, 1992, and 1996, respectively.

From 1996 to 1997, he worked in the medical imaging field with Mitra Imaging, Inc., Waterloo. From 1997 to 1999, he was an Assistant Professor in the Department of Geomatics Engineering, University of Calgary, Calgary, AB, Canada. In 1999, he returned to the University of Waterloo and is currently an Associate Professor in the Systems

Design Engineering program. His research interests include automated image interpretation, digital image processing, and pattern recognition with applications in remote sensing and medical imaging.



**Bing Yue** received the undergraduate degree in electrical engineering from Hebei University of Technology, Tianjin, China, in 1991, and the M.A.Sc. degree in systems design engineering from the University of Waterloo, Waterloo, ON, Canada, in 2001.

From 1991 to 1999, she was an Engineer with the Remote Sensing Geology Department, Research Institute of Petroleum Exploration and Development, Beijing, China. Since 2002, she has been a Research Assistant with Noetix Research Inc., Ottawa, ON, under contract with the Canada Center for Remote Sensing, Ottawa. Her current research interests include remotely sensed digital image processing, target detection, SAR polarimetry, interferometry, and pattern recognition.