
Comparison and Fusion of Co-occurrence, Gabor and MRF Texture Features for Classification of SAR Sea-Ice Imagery

David A. Clausi*

Systems Design Engineering, University of Waterloo, Waterloo, ON N2L 3G1

[Original manuscript received 23 January 2000; in revised form 26 June 2000]

ABSTRACT *Image texture interpretation is an important aspect of the computer-assisted discrimination of Synthetic Aperture Radar (SAR) sea-ice imagery. Co-occurrence probabilities are the most common approach used to solve this problem. However, other texture feature extraction methods exist that have not been fully studied for their ability to interpret SAR sea-ice imagery. Gabor filters and Markov random fields (MRF) are two such methods considered here. Classification and significance level testing shows that co-occurrence probabilities classify the data with the highest accuracy, with Gabor filters a close second. MRF results significantly lag Gabor and co-occurrence results. However, the MRF features are uncorrelated with respect to co-occurrence and Gabor features. The fused co-occurrence/MRF feature set achieves higher performance. In addition, it is demonstrated that uniform quantization is a preferred quantization method compared to histogram equalization.*

RÉSUMÉ [Traduit par la rédaction] *L'interprétation de la texture des images est un aspect important de la discrimination assistée par ordinateur des images de la glace de mer produites par le radar à antenne synthétique (RAS). Les probabilités de cooccurrence sont l'approche la plus courante utilisée pour résoudre ce problème. Il existe cependant d'autres méthodes d'extraction des caractéristiques de texture qui n'ont pas encore été complètement explorées du point de vue de leur aptitude à interpréter les images de la glace de mer du RAS. Les filtres de Gabor et les champs aléatoires de Markov (CAM) sont deux de ces méthodes examinées ici. Les tests de classification et de niveau de signification montrent que ce sont les probabilités de cooccurrence qui permettent d'obtenir le taux de classification le plus élevé, les filtres de Gabor n'étant pas loin derrière. Les résultats obtenus à l'aide des CAM sont nettement inférieurs à ceux produits au moyen des filtres de Gabor et de la cooccurrence. Toutefois, les caractéristiques obtenues par CAM sont sans corrélation avec les caractéristiques de la cooccurrence et des filtres de Gabor. L'ensemble fusionné des caractéristiques de la cooccurrence et des CAM permet d'obtenir de meilleurs résultats. En outre, on montre que la quantification uniforme est la méthode de quantification à privilégier par comparaison avec celle de l'égalisation d'histogrammes.*

1 Introduction

Synthetic Aperture Radar (SAR) is a powerful tool for resource management and environmental monitoring applications. Not only is SAR a sensor invariant to cloud cover and darkness, it is also especially useful for monitoring huge spatial regions that are not easily accessible. This is especially true for monitoring sea-ice concentrations and extents. Sea-ice information is important for assisting ship navigation in ice-infested waters and climate change monitoring in polar regions (Carsey, 1989; Barber et al., 1992).

Sea-ice image products must be produced in a timely fashion. Given the abundance of SAR digital imagery that must be analyzed, it makes sense to provide computer-aided techniques for interpretation. Semi-automated or automated dedicated methods should generate faster and cheaper SAR sea-ice image end products. Also, it is hoped that the computer-assisted techniques will be better able to produce unbiased evaluations of the sea-ice coverage.

The need for automated analysis of SAR sea-ice imagery has been clearly identified in the recent Canadian Global

Climate Observing System (GCOS) Plan for the Cryosphere (Agnew et al., 1999). For example, one of the "Two to Ten Year Action Items" is the continuation of the development of automated procedures at the Canadian Ice Service (CIS) to estimate geophysical parameters from RADARSAT (Brown, 1999). Also, a recent National Ice Center (NIC) Science Plan indicates that one of the primary activities which needs to be addressed is the development of SAR-based algorithms that can partially automate the generation of tactical ice products (Partington and Bertoia, 1997).

The development of reliable, robust methods for the consistent classification of SAR sea-ice data has been inconclusive, even though considerable effort has been made (Barber and LeDrew, 1991; Sholer, 1991; Barber et al., 1993; Soh and Tsatsoulis, 1999; Shanmugan et al., 1981; Holmes et al., 1984; Nystuen and Garcia, 1992). Since SAR sea-ice imagery contains spatially dependent class characteristics, texture extraction methods have been commonly used to generate feature information for sea-ice classes. The most common texture feature extraction method for remotely sensed

*Corresponding author's e-mail: dclausi@engmail.uwaterloo.ca

data is co-occurrence probabilities. However, a host of promising (Randen and Husoy, 1999) techniques in the research literature have not been extensively compared nor thoroughly assessed for their ability to discriminate SAR sea-ice imagery.

This paper compares the ability of three different texture feature methods to classify SAR sea-ice image samples. The three methods are: co-occurrence probabilities, Gabor filters and Markov random fields. Section 2 more fully defines the objectives of this paper. Section 3 details the texture methods, discriminant and data used. Testing and results are found in Section 4 with a discussion in Section 5.

2 Objectives

There are many studies that investigate specific texture measures, however, there are relatively few studies (Augusteijn et al., 1995; Clausi, 1996; Connors and Harlow, 1980; Pichler et al., 1996; Weszka et al., 1976) comparing these methods, especially with regard to SAR sea-ice image classification (Barber et al., 1993). No other work that directly compares the usefulness of each of these particular texture feature extraction techniques (co-occurrence probabilities, Gabor filters, and Markov random fields) for classifying SAR sea-ice imagery is known to the author.

Not only are these techniques considered independently, they are also considered for their ability to work synergistically. Features that are uncorrelated to each other are assumed to provide additional feature information to the classification system and combined to promote the accuracy of the classification. The author knows of no other research in the SAR sea-ice texture analysis field that considers the fusion of different texture feature types. This paper will investigate feature fusion with respect to the classification of SAR sea-ice imagery.

Grey level quantization is recognized to be a necessary step when creating features based on co-occurrence probabilities. Other studies have considered the impact of varying the quantization level on the classification ability of the co-occurrence features (Shokr, 1991; Soh and Tsatsoulis, 1999; Clausi and Jernigan, 1999). The greater the number of grey levels, the greater the computational costs yet the information content of the texture should be better preserved. With reduced grey levels, the computational costs are significantly reduced and the effect of noise in the data is reduced.

Soh and Tsatsoulis (1999) indicate that grey level quantization can be performed using one of three schemes: 1) uniform quantization 2) Gaussian quantization and 3) equal probability quantization (also commonly referred to as histogram equalization). There are other pixel-to-pixel transforms (e.g., log, square root, etc.) that could be considered for quantization, but these are not considered here. Uniform quantization linearly scales the grey levels into the desired range without considering the grey level distribution. Gaussian quantization and histogram equalization quantization methods map the grey level distribution of the image into Gaussian and uniform distributions, respectively. Soh and Tsatsoulis (1999) state that the Gaussian quantization method is inappropriate for SAR sea-ice imagery since the Gaussian

distribution is not an appropriate model for this data. The necessity of having the grey level mapping scheme match the expected grey level distribution of the data in order to perform optimal textural differentiation is an open question. However, there is still uncertainty about whether uniform quantization or histogram equalization is the preferred method for quantizing image data prior to creating the co-occurrence features. Soh and Tsatsoulis (1999) assume that uniform quantization would be superior to histogram equalization when applied to SAR sea-ice imagery, but no quantitative evidence is provided. Research has not been undertaken to demonstrate a preference for one quantization technique over the other when performing texture analysis of SAR sea-ice imagery. Here, a direct comparison of these two methods (uniform quantization and histogram equalization) will be conducted.

Based on the theories identified above, the following research questions will be addressed with respect to the given SAR sea-ice dataset:

- Q1. Which of the three methods (co-occurrence probabilities, Gabor filters, Markov random fields) offers preferred classification ability?
- Q2. Is uniform quantization or histogram equalization a better approach for producing discriminating feature sets? Do the number of quantization levels influence the ability of the techniques to discriminate between ice classes?
- Q3. Does combining feature sets generate an improved classification? Which combination of feature sets generates an improved performance?

3 Methods

a Texture Methods

1 GREY LEVEL CO-OCCURRENCE PROBABILITIES

Grey level co-occurrence texture features have been used for supervised classification of SAR sea-ice imagery (Shokr, 1991; Barber et al., 1993; Soh and Tsatsoulis, 1999). The co-occurrence probabilities are the conditional joint probabilities of all pairwise combinations of grey levels (i, j) in the spatial window of interest given two parameters: interpixel distance (δ) and orientation (θ) (Haralick et al., 1973). To generate texture features based on the co-occurrence probabilities, statistics are applied to the probabilities. Generally, these statistics identify some structural aspect of the arrangement of probabilities within a matrix indexed on i and j , which in turn reflects some characteristic of the texture. There are many statistics that can be used (Haralick et al., 1973). However, due to the redundancy in these statistics, only three statistics are advocated for SAR sea-ice classification since this should generate preferred discrimination with the least redundancy (Barber and LeDrew, 1991). The selected statistics are dissimilarity, entropy and correlation. These have been used in this study since each measure tends to be independent compared to other co-occurrence statistics and each represents a different characteristic of the co-occurrence matrix (Clausi, 1996). In addition, each statistic is insensitive to grey level shifts.

There is no known rigorous optimal method for selecting θ and δ . Given no other information concerning the window of interest, the preferred parameters should utilize adjacent neighbours, motivating the use of $\theta = 0, 45, 90, 135$ degrees and $\delta = 1$. SAR sea-ice 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. There is a potential for anisotropic behaviour in SAR sea-ice imagery, so orientations of 0, 45, 90 and 135 degrees are advocated. This combination of offset and orientation has characterized SAR texture well and has been identified as being preferred for such applications (Barber and LeDrew, 1991; Shokr, 1991). The use of four orientations, one pixel spacing and three statistics generates a 12-dimensional feature space. Where not explicitly stated, the grey level quantization level is uniform and set to 32. Rapid feature extraction is performed by using the linked list method advocated by Clausi and Jernigan (1998).

2 GABOR FILTERS

Success in the texture analysis research literature motivates the investigation of Gabor filters (Jain and Farrokhnia, 1991; Bovik et al., 1990; Tuener et al., 1995). Gabor functions, implemented as pseudo-wavelet operators, generate appropriate texture features in a more computationally efficient manner than the co-occurrence method (Clausi, 1996). Gabor texture features have been demonstrated as successful for texture segmentation problems, but few efforts have investigated their ability to specifically classify SAR sea-ice imagery. Barber et al. (1993) included Gabor filtering in their comparative study, however, they reduced the outputs of the filter bank to two features, even though the Gabor filter bank outputs, when set up in an appropriate wavelet manner, are effectively independent. Improved discriminating behaviour should occur with the use of preferred Gabor filter banks (Clausi and Jernigan, 1998).

A Gabor filter bank (Jain and Farrokhnia, 1991; Bovik et al., 1990) is a pseudo-wavelet filter bank where each filter generates a near-independent estimate of the local frequency content. Gabor filters are attractive for texture interpretation because of their mathematical tractability, their ease of implementation for multi-channel filtering, optimal joint spatial/spatial-frequency resolution, and their ability to model the filter characteristics of simple cells found in the visual cortex (Jain and Farrokhnia, 1991; Bovik et al., 1990; Daugman, 1985). Design of a Gabor filter bank for image texture segmentation was proposed by Jain and Farrokhnia (1991). The technique for extracting texture features used here is based on a preferred technique developed by Clausi and Jernigan (1999).

The Gabor filters used here are implemented with the following parameters. Filter bandwidths in the 2-d spatial-frequency plane are one octave and 30 degrees. To achieve proper spatial-frequency coverage, centre frequencies are also spaced by one octave and 30 degrees. For example, given 16×16 image samples, the top two centre frequencies are 0.35 and 0.18 cycles per pixel (cpp) (using octave spacing). This imple-

mentation allows generation of a wavelet approximation using Gabor filters i.e., doubling the filter centre frequency cuts the spatial extent of the filter in half. Thus, the spatial localization as a function of spatial-frequency is maintained for all filters in the filter bank. To determine the standard deviations (σ_x, σ_y), the filter cutoff is set to -6 db. To ensure average grey level insensitivity, the DC component of the filter is set to zero.

Given that the filters are well localized in space, local noise can generate misleading filter outputs. To alleviate this, smoothing of the magnitude images as a function of the same Gaussian used in the Gabor function is performed. Using $\gamma = 2/3$ for the smoothing Gaussian $g(\gamma x, \gamma y)$ proved effective, as Bovik et al. (1990) also experienced. Given the top two octave bands and six orientations, the Gabor filter bank generates a total of twelve texture features. Since each pixel has a response to each filter, then a 12-dimensional texture feature vector represents each centre pixel of each window sample.

3 MARKOV RANDOM FIELDS

Markov random fields (MRFs) have been demonstrated to be quite effective for texture characterization (Chellappa and Chatterjee, 1985; Cohen et al., 1991; Yamayaki and Gingras, 1995). Studies of their ability to classify SAR sea-ice imagery are unknown to the author. Application of MRFs to the identification of operational forestry parameters compares well with the co-occurrence technique (Wiebe, 1998). The MRF method generates parameters for a particular distribution function based on Markov assumptions. As a result, these texture measures are assumed to be quite different in nature relative to the other two methods summarized above.

MRF texture measurements can be made based on various orientations and pixel spacings. Feature measurements are determined by applying a specific distribution model to the underlying data to create texture features. A unique aspect of these texture features is that they not only provide a measurement of the texture, but they can also be used to create a facsimile of the analyzed texture (Chellappa, 1985). Most other texture models, including the co-occurrence technique, have not demonstrated this ability.

A Markov random field is a 2-d lattice of points, where each point is assigned a value based upon a probabilistic model. An MRF is a random field with Markovian properties, namely, a point's value on the lattice is only influenced by particular neighbouring values. The specific definition of *neighbours* and their influence on other points give MRFs the freedom to model many types of textures. MRF texture models for SAR images consider every image cell's backscatter intensity as a function of other image cells' intensities in its neighbourhood (Chellappa, 1985; Besag, 1974).

For the purposes of texture analysis, let $X(i,j)$ be a random variable which represents the value at (i,j) on an $N \times M$ lattice L . For simplicity, we shall index X with only one variable, e.g., $X(c)$ where $c = 1, 2, 3, \dots, N \times M$. For MRFs, if point m is a neighbour of point c then $p(X(c))$ depends on the value $X(m)$. A Markov random field is a joint probability density on

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

Fig. 1. Symmetric neighbours used to define the Markov random field.

the set of all possible digital numbers (representing the backscatter) of L such that $p(X(c)) > 0$ and

$$P(X(c) | X(m), m = 1, 2, \dots, N \times M, c \neq m) \\ = p(X(c) | \text{neighbours of } c).$$

For example, assuming that the conditional probability of a specific configuration of neighbours about point c is Gaussian we have:

$$P(X(c) | X(m), m \text{ is a neighbour of } c, c \neq m) \\ = (2\pi\sigma^2)^{-1/2} \exp[-(X(c) - \sum \beta_{c,m} \{X(m) + X(m')\})^2 / 2\sigma^2]$$

where σ represents the standard deviation and $\beta_{c,m}$ the parameters (i.e., the texture features) of the MRF. The summation (in the exponential) is taken over all symmetric neighbours of $X(n)$. Symmetric neighbours consist of a pair of image cells the same distance from the centre cell, c , but at opposing angles. The variable m' is defined as being symmetric to cell m . Looking at Fig. 1, the 1.1 entries are symmetric neighbours. The number of parameters (order of the model) depends on the number of symmetric neighbours used in the model. In Fig. 1, up to 5th order symmetric neighbours are depicted. For example, a 3rd order symmetric MRF model centred on image cell c , would include those cells marked 1.x, 2.x and 3.x and be characterized by six parameters (the 1.1 entries determine one parameter, the 1.2 entries the next parameter, 2.1 entries the third parameter, etc.). A number of different techniques for the determination of the MRF texture features exist: coding (Besag, 1974, 1986; Cross and Jain, 1983), least squares (Chellappa and Chatterjee, 1985; Manjurath and Chellappa, 1991) and maximum likelihood estimates (Besag, 1974, 1986).

When $p(X(c) | \text{neighbours of } c)$ is Gaussian, a difference equation can be used to represent the Markov process (Woods, 1972). The symmetric difference equation is:

$$X(c) = \sum \beta_{c,m} [X(c+m) + X(c-m)] + e_c$$

where e_c is zero mean Gaussian distributed noise, m is an offset from the centre cell c and $\beta_{c,m}$ is a parameter which weights a pair of symmetric neighbours to the centre cell. The summation is over all valid values for m as determined by the order of the model. For a 2nd order model, the summation would be over two values of m , namely $m = \{(1,0), (0,1)\}$. The $X(c-m)$ term is the symmetric neighbour of $X(c+m)$. In matrix notation, this equation is represented by:

$$X(c) = \beta^T \mathbf{Q}_c + e_c$$

where β is a vector composed of $\beta_{c,m}$ and \mathbf{Q}_c is a vector defined by:

$$\mathbf{Q}_c = \begin{bmatrix} X(c+m_1) + X(c-m_1) \\ X(c+m_2) + X(c-m_2) \\ X(c+m_3) + X(c-m_3) \\ \dots \end{bmatrix}.$$

The parameters (which represent the texture features) are estimated using a least squares approach:

$$\beta = \left[\sum_{c \in L} \mathbf{Q}_c \mathbf{Q}_c^T \right]^{-1} \left[\sum_{c \in L} \mathbf{Q}_c X(c) \right].$$

For every pixel in the window under consideration, a \mathbf{Q}_c is determined. The $N \times M$ window is defined by neighbourhood L . Then for every window, β is estimated which provides the texture features. The MRF features are, in practice, insensitive to grey level shifts.

b Information Content

As in Kurvonen et al. (1999), the information content of the texture measures are evaluated with (1) a separability index and (2) test classification. A separability index provides a relative estimate for information content of the texture measure under scrutiny that provides a basis for comparing different scenarios. Test classification considers the absolute accuracy of the texture measures with respect to the true classification.

1 TEST CLASSIFICATION

A common technique for the classification of feature vectors is the use of the maximum likelihood (ML) classifier. This technique is not used here because the restricted number of samples per class creates a class feature space representation that is sparse, given the dimension of the feature space. The Fisher linear discriminant (FLD) (Duda and Hart, 1973) is less sensitive to the number of features compared to ML because FLD uses a pooled covariance matrix compared to the individual class covariance matrix employed by ML (Tom and Miller, 1984). Also, the FLD is an appealing index since it is non-parametric.

The FLD can be implemented to generate a projection of a pair of n -dimensional class feature sets onto a one-dimensional vector. The projection provides an optimal separation of the two classes. Then, a maximum likelihood classifier can be used where, as in this case, there are sufficient samples to describe each one-dimensional Gaussian distribution. Only two classes can be compared at a time using the one-dimensional projection, so, to determine whether a sample belongs to one of C classes, $\binom{C}{2}$ comparisons are made. Using all comparisons, a sample assigned to a particular class the greatest number of times is the class to which the sample is assigned. Sometimes, two classes are assigned an equal number of times. In this case, the sample is designated "unclassified".

Using separate training and test datasets, error matrices are created. In this paper, kappa (κ) coefficients and associated

TABLE 1. Classification results for each feature extraction method. κ – kappa coefficient, σ – standard deviation for κ , Min. and Max. $J(\omega)$ – minimum and maximum Fisher criterion.

	Validated Data Used for Training				Inspected Data Used for Training			
	κ	σ	Min. $J(\omega)$	Max. $J(\omega)$	κ	σ	Min. $J(\omega)$	Max. $J(\omega)$
Co-occurrence	0.475	0.0329	1.4	165.4	0.512	0.0324	0.9	204.1
Gabor Filters	0.392	0.0325	1.0	113.4	0.448	0.0330	1.1	96.3
MRF order 3	0.141	0.0270	0.2	6.5	0.192	0.0282	0.3	6.9
MRF order 4	0.184	0.0280	0.4	9.5	0.173	0.0276	0.3	10.1
MRF order 5	0.158	0.0262	0.4	10.7	0.158	0.0271	0.3	11.0

confidence intervals (σ) are used to evaluate each error matrix (Bishop et al., 1975). When two error matrices are compared, the following test statistic can be used to determine a significance level (Congalton et al., 1983)

$$Z \sim \frac{\kappa_1 - \kappa_2}{\sqrt{\sigma_1^2 + \sigma_2^2}}$$

Note that there are typographical errors for the equations determining κ and σ in Congalton et al. (1983) and one is best to use the equations provided by Bishop et al. (1975).

2 SEPARABILITY INDEX

The Fisher criterion (Duda and Hart, 1973) can be used as a measure of the separability of two classes in the feature space. In this work, the scatter matrix (used in the formulation of the Fisher criterion) is actually defined as a true covariance matrix i.e., the scatter matrices are weighted by the number of samples. This is important for the case of classes which have a different number of samples.

c Synthetic Aperture Radar Images

Detailed information concerning the RADARSAT-1 imagery is contained in Yackel et al. (this issue) and Mundy and Barber (this issue). This uncalibrated imagery was collected as part of the international North Water (NOW) Polynya project. In summary, ScanSAR data was acquired from the CIS in a raw, uncalibrated mode with 150 m nominal resolution and 100 m pixel spacing. Although Mundy and Barber (this issue) observed 14 separate surface classes, only 9 classes are considered for this texture study. Classes were omitted due to an insufficient number of observed samples to perform a classification study. The classes studied here include: nilas/new ice, grey ice, young ice floes, medium first-year ice floes, rough open water, calm open water, smooth first-year ice, rubble landfast ice, and multiyear ice.

This dataset is difficult to classify for two reasons: (1) nine different classes are discriminated simultaneously and (2) relatively smaller window sizes (16×16) are used. Most SAR classification testing considers either fewer ice classes or larger windows or both. For example, Barber and LeDrew (1991) used three ice classes (multiyear, first-year rough, first-year smooth) with 25×25 windows. Shokr (1991) compares 5×5 , 7×7 and 9×9 windows to classify either three or four classes. Classification accuracy increases with increased window size. Soh and Tsatsoulis (1999) consider seven classes,

however, they also used considerably larger 64×64 samples which are recognized as being helpful to the classification process. Here, the 16×16 window size provides sufficient sampling to populate the 32×32 co-occurrence matrix. Some ice types did not have the spatial coverage to provide sufficient non-overlapping samples at larger window sizes.

Two sets of image samples were created. The “validated” dataset contains those regions with field observations during the NOW program that were co-registered in the images. The “inspected” samples are based on regions that were in the NOW vicinity but not directly observed during the field program. These class samples were selected by inspection of the images having similar visual characteristics as the validated classes. Thirty-two non-overlapping samples of each class within each set were selected.

4 Testing and results

Each of the hypotheses proposed in Section 2 will now be addressed.

- Q1. Which of the three methods (co-occurrence probabilities, Gabor filters, or Markov random fields) offers preferred classification ability?

To study this question, each method was used to perform two tests. First, the ground-validated data was used to train the system and the visually inspected data was classified. Then, the visually inspected data was used for training and the validated data was used for test classification. In addition to recording the kappa coefficients (κ) and confidence intervals (σ), the minimum and maximum Fisher criterion (from exhaustive class-pairwise analysis) are recorded based on the training data. Three MRF orders (3, 4 and 5) are tested to determine which order produces the most favourable results. Results are displayed in Table 1.

Co-occurrence results generated the best classification accuracies and the highest Fisher distances (see results in Q3 for a discussion of the significance levels between the feature methods). Gabor filters performed favourably but not to the same extent as the co-occurrence method. MRF results were poorer than Gabor and co-occurrence results. Based on using independent feature sets, co-occurrence was demonstrated to be performing the best.

One may assume that 0.512 kappa value (the best co-occurrence classification) is not a strong result. However, note that other studies (Shokr, 1991; Barber et al., 1993) usually only

TABLE 2. Significance levels between validated versus inspected error matrices for each feature extraction method. There is no statistically significant difference among the various methods (5% significance level).

Method	Significance Level
Co-occurrence	0.787
Gabor Filters	0.886
MRF order 3	0.903
MRF order 4	0.388
MRF order 5	0.751

TABLE 3. Joint results obtained by determining classification results of an error matrix created by combining the "Validation" and "Inspected" error matrices of Table 1.

	Joint Validation and Inspected Results			
	κ	σ	Min. $J(\omega)$	Max. $J(\omega)$
Co-occurrence	0.494	0.0231	0.9	204.1
Gabor Filters	0.420	0.0232	1.0	113.4
MRF order 3	0.178	0.0197	0.2	6.9
MRF order 4	0.167	0.0196	0.3	10.1
MRF order 5	0.145	0.0189	0.3	11.0

consider three classes (first-year rough, first-year smooth and multiyear) while this study is considering a total of nine different classes. For example, Barber and LeDrew (1991) had test kappas in the approximate range of 0.60 to 0.70, depending on the experiment. Shokr (1991) performs a host of comparative testing, but his results use classification accuracies instead of kappa coefficients, which does not allow for proper comparison. Tripling the number of classes dramatically increases the difficulty of the problem. Here, a random assignment of pixels would only generate about 11% classification accuracy, thus, a kappa value of 0.512 is quite strong considering this circumstance. The 4th order MRF kappa value of 0.184 is above random assignment and thus, distinguishing information is provided in this feature set. Note that each feature set only captures textural information and not the average grey level i.e., the methods employed are insensitive to grey level shifts.

The Fisher distances support the classification evidence. Co-occurrence and Gabor filtered features have high, similar Fisher distances. The distances for the MRF methods are about 10% of the values noted for co-occurrence and Gabor methods.

The statistical validity comparing the error matrix produced by classifying the inspected data using a validated data discriminant versus the error matrix produced by classifying the trained data using an inspected data discriminant has been determined (Table 2). The null hypothesis states that, for each feature extraction method, the pair of error matrices are the same. Table 2 reveals that there is no statistical significance between the error matrices created by classifying the test data using either validation or inspected data for training (in this paper, a 5% threshold is used i.e., 0.025 and 0.975 are the confidence levels). Since the null hypothesis is not false, the

TABLE 4. (a) Co-occurrence (b) Gabor filter and (c) MRF (4th order) feature classification and Fisher distances across three histogram equalization grey levels (64/32/16) and three linear quantizations (64/32/16).

(a)	Co-occurrence Probabilities			
	κ	σ	Min. $J(\omega)$	Max. $J(\omega)$
Hist. Eq. 64	0.261	0.0217	0.8	114.4
Hist. Eq. 32	0.263	0.0217	1.0	90.8
Hist. Eq. 16	0.285	0.0221	1.2	73.5
Unif. Quant. 64	0.459	0.0232	0.8	197.4
Unif. Quant. 32	0.494	0.0231	0.9	204.1
Unif. Quant. 16	0.465	0.0231	0.8	219.8
(b)	Gabor Filters			
	κ	σ	Min. $J(\omega)$	Max. $J(\omega)$
Hist. Eq. 64	0.268	0.0217	1.1	69.7
Hist. Eq. 32	0.262	0.0217	1.3	60.7
Hist. Eq. 16	0.283	0.0218	1.2	50.7
Unif. Quant. 64	0.459	0.0232	0.9	140.8
Unif. Quant. 32	0.448	0.0232	1.0	132.5
Unif. Quant. 16	0.434	0.0232	0.9	118.5
(c)	MRF order 4			
	κ	σ	Min. $J(\omega)$	Max. $J(\omega)$
Hist. Eq. 64	0.210	0.0203	0.4	9.3
Hist. Eq. 32	0.219	0.0206	0.4	9.6
Hist. Eq. 16	0.228	0.0206	0.4	9.4
Unif. Quant. 64	0.210	0.0203	0.4	9.3
Unif. Quant. 32	0.197	0.0203	0.3	7.7
Unif. Quant. 16	0.154	0.0192	0.3	6.6

error matrices are not demonstrated to be statistically different. As a result, for the rest of this paper, the test error matrices produced by the validation and inspected data are combined into a joint error matrix. The classification accuracy provided by this joint error matrix is provided in Table 3. Results are in agreement with those found in Table 1.

Q2. Is uniform quantization or histogram equalization a better method for producing discriminating feature sets? Do the number of quantization levels influence the ability of the techniques to discriminate between ice classes?

To test this set of questions, each texture method was used to determine classification accuracy based on quantization levels of 64, 32 and 16 for both histogram equalization and uniform quantization. Quantization was applied to each dataset. A 4th order MRF was used for MRF results since this produced the best MRF result in the first question. The null hypothesis states that, within each feature extraction method, the results should be consistent regardless of quantization method or quantization level. Classification results are presented in Table 4 and significance levels for these results are displayed in Table 5. Tables in the format of Table 5 are to be read in the following manner: for each item in the table, what is the statistical difference of the column entry with respect to the row entry? Statistical levels below 0.025 are significantly lower and statistical levels above 0.975 are significantly higher.

Comparing Texture Features for Classifying SAR Sea-Ice Imagery / 189

TABLE 5. Comparisons of significance levels across all quantization methods for each feature extraction technique: (a) co-occurrence (b) Gabor filters (c) MRF 4th order.

(a)	Hist. Eq. 64	Hist. Eq. 32	Hist. Eq. 16	Unif. Quant. 64	Unif. Quant. 32	Unif. Quant. 16
Hist. Eq. 64	x	0.515	0.781	1.000	1.000	1.000
Hist. Eq. 32	x	x	0.769	1.000	1.000	1.000
Hist. Eq. 16	x	x	x	1.000	1.000	1.000
Unif. Quant. 64	x	x	x	x	0.854	0.573
Unif. Quant. 32	x	x	x	x	x	0.193
Unif. Quant. 16	x	x	x	x	x	x
(b)	Hist. Eq. 64	Hist. Eq. 32	Hist. Eq. 16	Unif. Quant. 64	Unif. Quant. 32	Unif. Quant. 16
Hist. Eq. 64	x	0.414	0.680	1.000	1.000	1.000
Hist. Eq. 32	x	x	0.754	1.000	1.000	1.000
Hist. Eq. 16	x	x	x	1.000	1.000	1.000
Unif. Quant. 64	x	x	x	x	0.374	0.225
Unif. Quant. 32	x	x	x	x	x	0.332
Unif. Quant. 16	x	x	x	x	x	x
(c)	Hist. Eq. 64	Hist. Eq. 32	Hist. Eq. 16	Unif. Quant. 64	Unif. Quant. 32	Unif. Quant. 16
Hist. Eq. 64	x	0.627	0.732	0.500	0.328	0.022
Hist. Eq. 32	x	x	0.615	0.373	0.221	0.010
Hist. Eq. 16	x	x	x	0.268	0.144	0.004
Unif. Quant. 64	x	x	x	x	0.328	0.022
Unif. Quant. 32	x	x	x	x	x	0.060
Unif. Quant. 16	x	x	x	x	x	x

TABLE 6. Feature correlations. Correlations are determined by averaging individual class correlations and then averaging correlations over all directions for each feature type illustrated.

	Gabor 0.18cpp	Gabor 0.35cpp	Co-occ. DIS	Co-occ. ENT	Co-occ. COR	MRF 1 pixel	MRF $\sqrt{2}$ pixel	MRF 2 pixel	MRF $\sqrt{5}$ pixel
Gabor – 0.18cpp	0.263	0.255	0.419	0.463	0.192	0.022	-0.016	-0.002	-0.002
Gabor – 0.35cpp	–	0.317	0.520	0.531	0.065	0.009	0.002	-0.030	0.015
Co-occ. – DIS	–	–	0.885	0.899	0.125	0.005	-0.013	0.001	0.006
Co-occ. – ENT	–	–	–	0.987	0.308	0.015	-0.014	-0.001	0.002
Co-occ. – COR	–	–	–	–	0.517	0.040	-0.025	-0.010	-0.005
MRF – 1 pixel	–	–	–	–	–	-0.300	-0.175	-0.065	-0.065
MRF – $\sqrt{2}$ pixel	–	–	–	–	–	–	-0.570	0.025	-0.048
MRF – 2 pixel	–	–	–	–	–	–	–	-0.070	-0.271
MRF – $\sqrt{5}$ pixel	–	–	–	–	–	–	–	–	0.005

For co-occurrence probabilities and Gabor filters, there was a strong trend, namely, uniform quantization leads to a statistically significant improvement compared to histogram equalization (all significance levels are 1.000). For the co-occurrence probabilities, histogram equalization leads to kappa values around 0.26 while the uniform quantization was consistently around 0.46. A similar trend was observed for the Gabor filters. MRF texture features do not follow this trend. Both the uniform quantization and histogram equalization produce similar MRF results across most of the quantization levels. There was a statistically significant decrease in classification accuracy at a uniform quantization level of 16 where lower classification values are noted (0.154 compared to greater than 0.197 for the rest of the dataset). At such a high quantization, it may be that the MRF is less successful at finding characteristic differences between the textures due to excessive data smoothing. Gabor and co-occurrence error matrices were consistent across different quantizations given each quantization method.

The Fisher distances support these results. Fisher distances for co-occurrence and Gabor were higher for the linearly quantized data compared to the histogram equalization data.

Q3. Does combining feature sets generate an improved classification?

Correlations of the different features provide insight into the potential for fusing the features and producing an improved classification. Correlations are displayed in Table 6. These correlations were determined by first determining class correlations across all features (12 Gabor features + 12 co-occurrence features + 10 MRF 4th order features = 34 features in total). Then, a correlation matrix based on the class averages was produced. Since it is unwieldy to display this 34x34 correlation matrix, for each feature type, the average across all orientations for a given statistic (co-occurrence), frequency (Gabor), or distance (MRF) was determined. This averaging process generates the 9x9 matrix found in Table 6.

TABLE 7. Classification results and Fisher distances of fused feature sets.

	Fused Feature Sets			
	κ	σ	Min. $J(\omega)$	Max. $J(\omega)$
Co-occurrence	0.494	0.0231	0.9	204.1
Gabor Filters	0.420	0.0232	1.0	113.4
MRF order 4	0.178	0.0197	0.3	10.1
Co-occurrence & Gabor Filters	0.478	0.0229	3.0	378.6
Co-occurrence & MRF order 4	0.543	0.0226	5.2	251.5
Gabor Filters & MRF order 4	0.495	0.0228	3.3	143.7
Co-occurrence & Gabor Filters & MRF order 4	0.499	0.0227	10.5	502.0

Table 6 shows that the co-occurrence features tend to have high intra-feature correlations, especially with regards to Dissimilarity (DIS) and Entropy (ENT). This was an expected result (Barber and LeDrew, 1991; Clausi and Jernigan, 1998). Correlation (COR) was not as strongly correlated with DIS or ENT. Gabor filter features generally have lower intra-feature correlations. MRF features have a strong tendency to produce low intra-feature correlations. Gabor and co-occurrence features have relatively strong inter-feature correlations. What is most interesting is that MRF features are *not* well correlated with either the Gabor filtered features or the co-occurrence features. This indicates that the MRF features were providing unique information to the classification process that the Gabor and co-occurrence methods do not measure. By combining MRF features with either or both the Gabor and co-occurrence features, a more successful classification was expected.

Tables 7 (classification results) and 8 (statistical significance levels) reveal the outcome of fusing the feature sets based on using each method's preferred feature set. Gabor filters have a statistically lower classification success than co-occurrence features (0.013). However, fusing co-occurrence and Gabor feature sets combines redundant information since these two feature sets were strongly correlated. When Gabor filters were combined with the co-occurrence probabilities, there was no statistically significant change (0.315) and the classification accuracy actually dropped (from 0.494 to 0.478). Although the co-occurrence probabilities combined with the Gabor filters produced a significant improvement in the classification (0.962), the classification rate does not exceed that of co-occurrence alone, which was the objective of fusing feature sets. As a result, using co-occurrence alone was a better alternative than combining co-occurrence with Gabor features. MRF features alone have a consistently poorer classification than any of the other feature sets, alone or fused.

In contrast, the addition of MRF features to either Gabor or co-occurrence features produced improved results. The clas-

sification accuracy increased from 0.494 to 0.543 by adding MRF features to the co-occurrence features (statistical significance level of 0.936). Although this does not exceed the hard threshold of 0.975, it was a strong result. Similarly, the classification accuracy increased from 0.420 to 0.495 by adding MRF features to the co-occurrence features (statistically significant improvement of 0.989). These results advocate the use of a combined co-occurrence and MRF feature set for classifying SAR sea-ice imagery. Also, adding the MRF features to the fused co-occurrence/Gabor set improved the classification accuracy (from 0.478 to 0.499), however, the level was not statistically significant (0.740). This was in contrast to the fusing of a feature set with a correlated feature set. For example, fusing Gabor filters with a co-occurrence/MRF feature set decreased the classification accuracy from 0.543 to 0.499 (a strong significance level of 0.084). Fusing co-occurrence features with a fused Gabor/MRF feature set improved classification accuracy slightly from 0.495 to 0.499, but the significance level indicates that this change was definitely not significant (0.549).

Note that the ability of the training set to discriminate between the classes increased with additional features (both minimum and maximum Fisher criteria increase – refer to Table 1 as well), however, this increase did not have an impact on the classification of the test data. The highest Fisher distances were calculated when all three feature sets were combined, however, this did not lead to strong performance of the classifier when applied to test data.

To see how the MRF features improved the co-occurrence classification rate on a class-by-class basis, consider the two truth tables found in Table 9. Both tables display results based on combining truth tables using validated and inspected data (thus, there were 64 samples per class and some samples were designated unclassified). Columns represent reference data and rows represent classified data. Table 9(a) represents results using co-occurrence texture features alone and Table 9(b) represents results using co-occurrence features combined with MRF features. Strong improvements to the classification of the young ice (23 correctly classified for the co-occurrence versus 41 correctly classified for the co-occurrence/MRF combined), rough water (15 versus 25), and rubble landfast ice (6 versus 20) were noted. The only class with a noticeable reduction in classification was nilas ice (25 for co-occurrence versus 18 for co-occurrence and MRF).

Other improvements were noted using errors of commission/omission. For example, using co-occurrence alone, 20 samples were erroneously assigned to young instead of rubble ice and 16 samples were erroneously assigned to rubble instead of young ice. Combining co-occurrence and MRF feature sets reduced each of these totals to 1. Thus, the MRF features significantly improved the distinction between rubble and young ice types. The same can be said for rough water and young ice types: 20/7 samples were assigned to young ice/rough water that should have been assigned to rough water/young ice. Adding the MRF features reduced these values to 5 and 6, respectively.

Comparing Texture Features for Classifying SAR Sea-Ice Imagery / 191

TABLE 8. Statistical significance levels between combined validation/inspected error matrices for the identified feature sets.

	Co-occurrence	Gabor Filters	MRF 4 th	Co-occurrence & Gabor Filters	Co-occurrence & MRF 4 th	Gabor Filters & MRF 4 th	Co-occurrence & Gabor Filters & MRF 4 th
Co-occurrence	x	0.013	0.000	0.315	0.936	0.514	0.563
Gabor Filters	x	x	0.000	0.962	1.000	0.989	0.992
MRF 4 th	x	x	x	1.000	1.000	1.000	1.000
Co-occurrence & Gabor Filters	x	x	x	x	0.978	0.698	0.740
Co-occurrence & MRF 4 th	x	x	x	x	x	0.067	0.084
Gabor Filters & MRF 4 th	x	x	x	x	x	x	0.549
Co-occurrence & Gabor Filters & MRF 4 th	x	x	x	x	x	x	x

TABLE 9. Combined truth tables for validation and visually inspected classifications given (a) co-occurrence alone (b) co-occurrence with MRF features added. Classes: nilas/new ice, grey ice, young ice floes, medium first-year ice floes, rough open water, calm open water, smooth first-year ice, rubble landfast ice, and multi-year (my) ice.

(a)	nilas	grey	young	medium	rough	calm	smooth	rubble	my	
nilas	41	11	0	2	0	0	4	0	0	58
grey	5	25	2	4	0	1	1	11	0	49
young	0	0	23	1	20	0	0	20	0	64
medium	5	18	12	48	18	0	0	4	2	107
rough	0	0	7	4	15	0	0	15	1	42
calm	0	1	0	0	0	49	11	0	0	61
smooth	7	4	0	0	0	12	47	0	0	70
rubble	2	0	16	2	7	0	0	6	1	34
my	0	0	2	0	0	0	0	3	60	65
	60	59	62	61	60	62	63	59	64	

(b)	nilas	grey	young	medium	rough	calm	smooth	rubble	my	
nilas	39	5	0	0	0	1	4	1	0	50
grey	5	18	2	7	2	3	3	6	0	46
young	0	0	41	1	5	0	0	1	1	49
medium	4	16	9	44	13	0	0	5	2	93
rough	0	0	6	7	25	0	0	15	1	54
calm	0	2	0	0	0	46	8	0	0	56
smooth	7	4	0	0	0	12	46	0	0	69
rubble	0	0	1	2	10	0	0	20	2	35
my	2	0	1	1	0	0	0	7	58	69
	57	45	60	62	55	62	61	55	64	

To simulate results based on fewer classes to verify that the test results were consistent, classes were merged based on their relative Fisher distance. A preferred classification was generated when the co-occurrence and 4th order MRF feature sets were combined. Consider the relative Fisher distances between each pair of classes based on the training data using MRF and co-occurrence features (Table 10). The objective was to merge classes that are closest together in the feature space. The nearest classes were combined until an arbitrary number of four classes remained i.e., combine the closest two classes, namely, calm water and smooth ice at 5.2; then combine nilas/new ice and grey ice since they were the next closest at a distance of 8.4; etc. This merging process, as

expected, grouped classes that were visually similar in the feature set. The four merged classes were as follows:

- Class I – nilas/new ice, grey ice, medium first-year ice floes
- Class II – young ice floes, rough water
- Class III – calm open water, smooth first-year ice
- Class IV – rubble landfast ice, multi-year ice

The classification results based on each of the three texture methods individually, pairs of texture methods, and all three methods together using the four class representation are displayed in Table 11. Table 12 contains the parallel results for significance level testing. Results were consistent with previ-

TABLE 10. Fisher distances between each class using combined cooccurrence and 4th order MRF features. Classes: nilas/new ice, grey ice, young ice floes, medium first-year ice floes, rough open water, calm open water, smooth first-year ice, rubble landfast ice, and multi-year (my) ice.

	nilas	grey	young	medium	rough	calm	smooth	rubble	my
nilas	–	8.4	67.3	24.6	49.3	41.4	31.3	61.7	69.1
grey	–	–	121.6	11.1	68.6	21.4	16.7	93.3	102.1
young	–	–	–	27.0	12.9	176.3	211.6	14.4	43.2
medium	–	–	–	–	13.6	29.2	35.1	31.4	45.7
rough	–	–	–	–	–	133.4	148.5	6.4	30.9
calm	–	–	–	–	–	–	5.2	136.4	155.9
smooth	–	–	–	–	–	–	–	160.4	178.3
rubble	–	–	–	–	–	–	–	–	12.5
my	–	–	–	–	–	–	–	–	–

TABLE 11. Classification test results based on four combined classes.

	Combined Classes			
	κ	σ	Min. J(ω)	Max. J(ω)
Co-occurrence	0.667	0.0243	2.1	81.4
Gabor Filters	0.655	0.0246	2.5	53.6
4 th order MRF	0.251	0.0275	0.4	1.6
Co-occurrence & Gabor Filters	0.684	0.0238	4.0	94.4
Co-occurrence & 4 th order MRF	0.733	0.0224	5.1	89.5
Gabor Filters & 4 th order MRF	0.702	0.0233	5.0	56.5
Co-occurrence & Gabor Filters & 4 th order MRF	0.712	0.0229	6.1	107.4

ous testing. As individual feature sets, co-occurrence and Gabor produce strong results. Here, co-occurrence and Gabor do not have a significant difference (0.364 significance level). As with nine classes, MRF results significantly lagged co-occurrence and Gabor results. Fusing Gabor features with co-occurrence features increased the success compared to using co-occurrence features alone, and this increase was not statistically significant (0.691). Fusing co-occurrence with Gabor filters improved the classification results compared to Gabor alone, however, this result was not statistically significant (0.801).

Combining MRF with Gabor and combining MRF with co-occurrence produced much improved results over either Gabor or co-occurrence alone (significance levels of 0.977 and 0.918, respectively). The best results were obtained by fusing MRF features with co-occurrence features (kappa of 0.733). Merging all three feature sets dramatically increased the separability of the training data (maximum Fisher criterion of 107.4), but the results were not reflected in the test classification. Fusing Gabor filter features to the co-occurrence/MRF feature set reduced the classification accuracy (from 0.733 to 0.712).

5 Discussion

The Gabor filter and co-occurrence probability methods measure texture features in SAR sea-ice imagery basically by measuring local frequency. Gabor filters directly measure

local frequency components by acting as a bandpass filter centred on the frequency of interest. The two co-occurrence measures that generally perform more strongly (dissimilarity and entropy) (Clausi and Jernigan, 1998) correlate well with local frequency measures. For example, smooth textures tend to have fewer entries in the co-occurrence matrix (low entropy) which are located close to the diagonal (low dissimilarity). MRF features measure something completely different about the local texture compared to Gabor filters and co-occurrence probabilities. Here, model parameters are generated to best fit a Gaussian distribution. However, there is no correlation between the MRF features and the co-occurrence or between the MRF and Gabor features. Yet, the MRF classification accuracy indicates that they provide meaningful information. Fusing MRF features with Gabor filter features improves the classification accuracy of using Gabor features alone. Fusing MRF features with co-occurrence features produces the same effect. The combination of MRF and co-occurrence features is advocated for improved texture feature recognition of SAR sea-ice imagery since it consistently generated preferred results.

The MRF features add to the classification accuracy by improving the distinction of a number of class pairs confused by using only the co-occurrence features. In this dataset, MRF features improved the distinction between rough open water, rubble zones and young ice. Distinctions between these ice types are recognized as being difficult, especially in the context of identifying the other six classes in this dataset. This supports the use of MRF features for assisting the discrimination of SAR sea-ice imagery.

Gabor filters, implemented as wavelet operators, are not ideal candidates for solving the texture classification problem. Their inherent multiresolutional ability is not utilized for fixed-sized samples. The effective window size for the Gabor *wavelet* filters is gauged by the centre frequency of interest. Higher frequencies utilize smaller windows and lower frequencies utilize larger windows. This ability is more suited to the segmentation problem where the window size is not pre-defined. A future recommendation for classification testing is to redesign the filter bank to set the effective spatial filter extent to match that of the window size, regardless of the frequency of interest. This moves away from the concept of a wavelet filter bank, however, to ensure better coverage of the entire window sample by each filter, this approach for feature extraction should be attempted.

Comparing Texture Features for Classifying SAR Sea-Ice Imagery / 193

TABLE 12. Significance levels for each feature set for results based on four combined classes.

	Co-occurrence	Gabor Filters	MRF 4 th	Co-occurrence & Gabor Filters	Co-occurrence & MRF 4 th	Gabor Filters & MRF 4 th	Co-occurrence & Gabor Filters & MRF 4 th
Co-occurrence	x	0.364	0.0	0.691	0.977	0.852	0.911
Gabor Filters	x	x	0.0	0.801	0.991	0.918	0.955
MRF 4 th	x	x	x	1.000	1.000	1.000	1.000
Co-occurrence & Gabor Filters	x	x	x	x	0.934	0.707	0.803
Co-occurrence & MRF 4 th	x	x	x	x	x	0.168	0.254
Gabor Filters & MRF 4 th	x	x	x	x	x	x	0.620
Co-occurrence & Gabor Filters & MRF 4 th	x	x	x	x	x	x	x

In the research literature, some form of quantization is advocated when using co-occurrence probabilities for texture analysis. For the datasets utilized in this study, there is strong support for using uniform quantization as opposed to histogram equalization. The co-occurrence measures behaved quite consistently regardless of the quantization level, given either uniform quantization or histogram equalization. This is probably due to the fact that certain co-occurrence statistics perform better using coarser quantization (such as entropy) and other statistics perform better with more grey levels (such as dissimilarity) (Clausi and Jernigan, 1998). Given the parameters used in this study, varying the quantization levels did not have a noticeable effect on the classification accuracy.

A comment should be made concerning the use of presenting results for classifying training data only. To have reliable classification testing, separate test and training data is required. Some other studies have used the same data for testing and training (Weszka et al., 1976; Chang and Kuo, 1993) which can lead to misleading interpretations. Classification accuracies of training data may indicate strong performance, however, the relevance to test datasets is unknown.

Co-occurrence probabilities are identified as the most successful independent method compared to Gabor filters and

Markov random fields for the supervised classification of the given SAR sea-ice dataset. A more formidable problem is that of *segmentation* of SAR sea-ice imagery. Future testing should involve determining the ability of co-occurrence methods for the segmentation of SAR sea-ice imagery. Segmentation involves consideration of the boundaries between different textures. The ability of the co-occurrence method to identify textural boundaries is uncertain.

Sea ice is notoriously difficult to classify because of numerous nuances and variations within class, across a scene, from scene to scene, seasonally, diurnally, etc. A reliable means of capturing essential textural information is important. Co-occurrence texture features are a proven method for capturing essential texture information. In this paper, augmenting co-occurrence feature sets with features based on Markov random fields has been demonstrated to be a viable means of producing improved texture feature sets.

Acknowledgements

Appreciation is extended to C.J. Mundy and D.A. Barber who generously collected and supplied the necessary datasets for this study.

References

- AGNEW T.; R. BROWN, G. FLATO, H. MELLING and B. RAMSAY. 1999. *Canadian Contributions to GCOS - Sea-Ice*, Atmospheric Environment Service (AES) - Environment Canada, Toronto, Ontario, Canada, 23 pp.
- AUGUSTEIJN, M.F.; L.E. CLEMENS and K.A. SHAW. 1995. Performance evaluation of texture measures for ground cover identification in satellite images by means of a neural network classifier. *IEEE Trans. Geosci. Remote Sens.* **33(3)**: 616–626.
- BARBER, D.G. and E.F. LEDREW. 1991. SAR sea ice discrimination using texture statistics: a multivariate approach. *Photogram. Eng. Remote Sens.* **57(4)**: 385–395.
- ; M.J. MANORE, T.A. AGNEW, H. WELCH, E.D. SOULIS and E.F. LEDREW. 1992. Science issues relating to marine aspects of the cryosphere: implications for remote sensing. *Can. J. Remote Sens.* **18(1)**: 46–55.
- ; M.E. SHOKR, R.A. FERNANDES, E.D. SOULIS and D.G. FLETT. 1993. A comparison of second-order classifiers for SAR sea ice discrimination. *Photogram. Eng. Remote Sens.* **59(9)**: 1397–1408.
- BESAG, J. 1974. Spatial interaction and the statistical analysis of lattice systems. *J. R. Statist. Soc. B.* **36(2)**: 192–236.
- . 1986. On the statistical analysis of dirty pictures, *J.R. Statist. Soc. B.* **48(3)**: 259–302.
- BISHOP, T.; S. FIENBERG and P. HOLLAND. 1975. *Discrete Multivariate Analysis - Theory and Practice*, MIT Press, Cambridge, Massachusetts, 557 pp.
- BOVIK, A.C.; M. CLARK and W.S. GEISLER. 1990. Multichannel texture analysis using localized spatial filters. *IEEE Trans. Patt. Anal. Mach. Int.* **12(1)**: 55–73.
- BROWN, R.; T. AGNEW, M. BURGESS, G. COGLEY, M. DEMUTH, C. DUGUAY, G. FLATO, B. GOODISON, R. KOERNER, H. MELLING, T. PROWSE, B. RAMSAY, M.

- SHARP, S. SMITH and A. WALKER. 1999. Development of a Canadian GCOS workplan for the cryosphere, Atmospheric Environment Service (AES), Environment Canada, Toronto, Ontario, Canada.
- CARSEY, F. 1989. Review and status of remote sensing of sea ice. *IEEE J. Oceanic Eng.* **14(2)**: 127–138.
- CHANG, T. and C.C.J. KUO. 1993. Texture analysis and classification with tree-structured wavelet transform. *IEEE Trans. Image Proc.* **2(4)**: 429–441.
- CHELLAPPA, R. 1985. Two-dimensional discrete Gaussian Markov random field models for image processing. In: *Machine Intelligence and Pattern Recognition: Progress in Pattern Recognition 2*, G.T. Toussaint (Ed.), Elsevier Science Publishers, B.V. (North-Holland). pp. 79–112.
- and S. CHATTERJEE. 1985. Classification of textures using Gaussian Markov random fields. *IEEE Trans. Acous., Speech Signal Proc.* **33(4)**: 959–963.
- CLAUSI, D.A. 1996. *Texture Segmentation of SAR Sea Ice Imagery*. Ph.D. thesis, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1, 175 pp.
- and M.E. JERNIGAN. 1998. A fast method to determine co-occurrence texture features. *IEEE Trans. Geosci. Remote Sens.* **36(1)**: 298–300.
- and ———. 1999. Designing Gabor filters for optimal texture separability. *Pattern Recog.* **33(11)**: 1835–1849.
- COHEN, F.S.; Z. FAN and M.A. PATEL. 1991. Classification of rotated and scaled texture images using Gaussian Markov random field models. *IEEE Trans. Patt. Anal. Mach. Int.* **13(2)**: 192–203.
- CONGALTON, R.G.; R.G. ODERWALD and R.A. MEAD. 1983. Assessing Landsat classification accuracy using discrete multivariate analysis statistical techniques. *Photogram. Eng. Remote Sens.* **49(12)**: 1671–1678.
- CONNERS, R.W. and C.A. HARLOW. 1980. A theoretical comparison of texture algorithms. *IEEE Trans. Patt. Anal. Mach. Int.* **2(3)**: 204–222.
- CROSS, G.R. and A.K. JAIN. 1983. Markov random field texture models. *IEEE Trans. Patt. Anal. Mach. Int.* **PAMI-5(1)**: 25–39.
- DAUGMAN, J.G. 1985. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortex filters. *J. Opt. Soc. Am. A* **2(7)**: 1160–1169.
- DUDA, R.O. and P.E. HART. 1973. *Pattern Classification and Scene Analysis*, John Wiley and Sons. Toronto, 482 pp.
- HARALICK, R.M.; K. SHANMUGAN and I. DINSTEN. 1973. Textural features for image classification. *IEEE Trans. Syst. Man. Cybern.* **SMC-3(6)**: 610–621.
- HOLMES, Q.A.; D.R. NUESCH and R.A. SHUCHMAN. 1984. Textural analysis and real-time classification of sea-ice types using digital SAR data. *IEEE Trans. Geosci. Remote Sens.* **GE-22(2)**: 113–120.
- JAIN, A.K. and F. FARROKHNI. 1991. Unsupervised texture segmentation using Gabor filters. *Pattern Recog.* **24(12)**: 1167–1186.
- KURVONEN, L.; J. PULLIAINEN and M. HALLIKAINEN. 1999. Retrieval of biomass in boreal forests from multitemporal ERS-1 and JERS-1 SAR images. *IEEE Trans. Geosci. Remote Sens.* **37(1)**: 198–205.
- MANJUNATH, B.S. and R. CHELLAPPA. 1991. Unsupervised texture segmentation using Markov random field models. *IEEE Trans. Patt. Anal. Mach. Int.* **13(5)**: 478–482.
- MUNDY, C.J. and D.A. BARBER. 2001. On the Relationship between Spatial Patterns of Sea-Ice Type and the Mechanisms which Create and Maintain the North Water (NOW) Polynya. *ATMOSPHERE-OCEAN*, **39**: 327–341.
- NYSTUEN, J.A. and F.W. GARCIA JR. 1992. Sea ice classification using SAR backscatter statistics. *IEEE Trans. Geosci. Remote Sens.* **30(3)**: 502–509.
- PARTINGTON, K.C. and C. BERTOIA. 1997. Science Plan - Version 6.0, National Ice Center, Washington, DC.
- PICHLER, O.; A. TEUNER and B.J. HOSTICKA. 1996. A comparison of texture feature extraction using adaptive Gabor filtering, pyramidal and tree structured wavelet transforms. *Pattern Recog.* **29(5)**: 733–742.
- RANDEN, T. and J.H. HUSOY. 1999. Filtering for texture classification: A comparative study. *IEEE Trans. Patt. Anal. Mach. Int.* **21(4)**: 291–310.
- SHANMUGAN, K.S.; V. NARAYANAN, V.S. FROST, J.A. STILES and J.C. HOLTZMAN. 1981. Textural features for radar image analysis. *IEEE Trans. Geosci. Remote Sens.* **GE-19(3)**: 153–156.
- SHOKR, M.E. 1991. Evaluation of second-order texture parameters for sea ice classification from radar images. *J. Geophys. Res.* **96(C6)**: 10625–10640.
- SOH, L.K. and C. TSATSOUKIS. 1999. Texture analysis of SAR Sea ice imagery using gray level co-occurrence matrices. *IEEE Trans. Geosci. Remote Sens.* **37(2)**: 780–795.
- TEUNER, A.; O. PICHLER and B.J. HOSTICKA. 1995. Unsupervised texture segmentation of images using tuned matched Gabor filters. *IEEE Trans. Image Proc.* **4(6)**: 863–869.
- TOM, C.H. and L.D. MILLER. 1984. An automated land-use mapping comparison of the Bayesian maximum likelihood and linear discriminant analysis algorithms. *Photogram. Eng. Remote Sens.* **50(2)**: 193–207.
- WESZKA, J.S.; C.R. DYER and A. ROSENFELD. 1976. A comparative study of texture measures for terrain classification. *IEEE Trans. Syst. Man. Cybern.* **SMC-6(4)**: 269–285.
- WIEBE, J.A. 1998. *Texture Estimates Of Operational Forestry Parameters*. M.Sc. Thesis, University of Calgary, Calgary, Alberta. 99 pp.
- WOODS, J.W. 1972. Two-dimensional discrete Markovian fields. *IEEE Trans. Info. Theory.* **18(2)**: 232–240.
- YACKEL, J.J.; D.A. BARBER and T.N. PAPPYRIAKU. 2001. On the examination of spring melt in the North Water (NOW) using RADARSAT-1. *ATMOSPHERE-OCEAN*, **39**: 195–208.
- YAMAZAKI, T. and D. GINGRAS. 1995. Image classification using spectral and spatial information based on MRF models. *IEEE Trans. Image Proc.* **4(9)**: 1333–1339.