# Texture Analysis of SAR Sea Ice Imagery Using Gray Level Co-Occurrence Matrices

Leen-Kiat Soh, *Member, IEEE*, and Costas Tsatsoulis, *Senior Member, IEEE*

*Abstract*—This paper presents a preliminary study for mapping sea ice patterns (texture) with 100-m ERS-1 synthetic aperture radar (SAR) imagery. We used gray-level co-occurrence matrices (GLCM) to quantitatively evaluate textural parameters and representations and to determine which parameter values and representations are best for mapping sea ice texture. We conducted experiments on the quantization levels of the image and the displacement and orientation values of the GLCM by examining the effects textural descriptors such as entropy have in the representation of different sea ice textures. We showed that a complete gray-level representation of the image is not necessary for texture mapping, an eight-level quantization representation is undesirable for textural representation, and the displacement factor in texture measurements is more important than orientation. In addition, we developed three GLCM implementations and evaluated them by a supervised Bayesian classifier on sea ice textural contexts. This experiment concludes that the best GLCM implementation in representing sea ice texture is one that utilizes a range of displacement values such that both microtextures and macrotextures of sea ice can be adequately captured. These findings define the quantization, displacement, and orientation values that are the best for SAR sea ice texture analysis using GLCM.

*Index Terms*—Co-occurrence matrix, SAR, sea ice, texture.

## I. INTRODUCTION

**I**N THIS paper, we present a set of experiments on textural parameters and representations and a quantitative evaluation of these experiments, which shows which textural parameter values and texture representations are best for describing sea ice in synthetic aperture radar (SAR) imagery. We selected seven different sea ice textural contexts, i.e., sea ice texture types, and used them as our test set. These textural contexts have no definitive, intrinsic geophysical significance and were so selected because they were *visually* separable by a human, without being too different as to make the separation trivial. We computed the texture matrix representations of these sample contexts and used a supervised Bayesian classifier to evaluate how well the texture matrices could describe and recognize the textural contexts.

The texture matrix used was the gray-level co-occurrence matrix (GLCM). In designing the GLCM for texture representation, there are three fundamental parameters that must

be defined: the quantization levels of the image and the displacement and orientation values of the measurements. We performed a set of experiments in which we systematically varied these parameters and studied how the variations affected GLCM standard texture descriptors for SAR sea ice images. From these experiments, we concluded which quantization levels and displacement and orientation values are best for representing sea ice texture in SAR. We developed and evaluated three different implementations of GLCM, the mean displacement and mean orientation (MDMO) matrix, the $\chi^2$-optimal displacement and mean orientation (ODMO) matrix, and the $\chi^2$-optimal displacement and $\chi^2$-optimal orientation (ODOO) matrix. The implementations were evaluated as to their ability to separate between sea ice texture contexts. Based on these experiments, we concluded which texture matrix representation is best at separating sea ice texture types in SAR imagery.

The ability to represent sea ice texture contexts well, i.e., in a way that allows classification and separation of the contexts, is extremely significant in sea ice analysis, classification, and description. Statistical texture analysis is important in SAR sea ice imagery research since it allows better representation and segmentation of sea ice regions, compared to analysis based on intrinsic gray levels only. It has been shown that the inclusion of texture as a descriptor can improve the classification of sea ice and the description of sea ice deformations [38], [56], [63]. Some work has attempted to identify which textural measurements provide better descriptors for sea ice [65], but no work has provided a comprehensive experiment of textural parameters and representations with an evaluation of the quality of the representation based on quantifiable metrics. Our work is the first one to evaluate all possible textural representation parameters and to make specific recommendations about the representation of sea ice texture in SAR imagery.

## II. BACKGROUND ON GLCM

The definition of GLCM's is as follows [35]. Suppose an image to be analyzed is rectangular and has $N_x$ columns and $N_y$ rows. Suppose that the gray level appearing at each pixel is quantized to $N_g$ levels. Let $L_x = \{1, 2, \cdots, N_x\}$ be the columns, $L_y = \{1, 2, \cdots N_y\}$ be the rows, and $G_x = \{0, 1, \cdots, N_g - 1\}$ be the set of $N_g$ quantized gray levels. The set $L_y \times L_x$ is the set of pixels of the image ordered by their row–column designations. The image $I$ can be represented as a function that assigns some gray level in $G$ to each pixel or pair of coordinates in $L_y \times L_x$; $I:L_y \times L_x \rightarrow G$. The texture-context information is

specified by the matrix of relative frequencies $P_{ij}$ with two neighboring pixels separated by distance $d$ occur on the image, one with gray level $i$ and the other with gray level $j$. Such matrices of gray-level co-occurrence frequencies are a function of the angular relationship and distance between the neighboring pixels. Formally, for angles quantized to $45°$ intervals, the unnormalized frequencies are defined as shown in the equations at the bottom of hte page where $\#$ denotes the number of elements in the set.

We used ten textural features in our study. The following equations define these features. Let $p(i, j)$ be the $(i, j)$th entry in a normalized GLCM. The mean and standard deviations for the rows and columns of the matrix are

$$\mu_x = \sum_i \sum_j i \cdot p(i, j), \quad \mu_y = \sum_i \sum_j j \cdot p(i, j)$$

$$\sigma_x = \sum_i \sum_j (i - \mu_x)^2 \cdot p(i, j),$$

$$\sigma_y = \sum_i \sum_j (j - \mu_y)^2 \cdot p(i, j).$$

The features are as follows.

1) Energy:

$$f_1 = \sum_i \sum_j p(i, j)^2.$$

2) Contrast:

$$f_2 = \sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) \Big| |i - j| = n \right\}.$$

3) Correlation:

$$f_3 \frac{\sum_i \sum_j (ij)p(i, j) - \mu_x\mu_y}{\sigma_x\sigma_y}.$$

4) Homogeneity:

$$f_4 = \sum_i \sum_j \frac{1}{1 + (i - j)^2} p(i, j).$$

5) Entropy:

$$f_5 = -\sum_i \sum_j p(i, j) \log(p(i, j)).$$

6) Autocorrelation:

$$f_6 = \sum_i \sum_j (ij)p(i, j).$$

7) Dissimilarity:

$$f_7 = \sum_i \sum_j |i - j| \cdot p(i, j).$$

8) Cluster Shade:

$$f_8 = \sum_i \sum_j (i + j - \mu_x - \mu_y)^3 p(i, j).$$

9) Cluster Prominence:

$$f_9 = \sum_i \sum_j (i + j - \mu_x - \mu_y)^4 p(i, j).$$

10) Maximum Probability:

$$f_{10} = \underset{i, j}{\text{MAX}}\, p(i, j).$$

Note that energy is also popularly known as angular second moment [30]. The two cluster parameters were introduced in [19] to emulate human perceptual behavior. Maximum probability was discussed in [34]. Energy, contrast, correlation, homogeneity, entropy, autocorrelation, and dissimilarity were formulated in [35].

Zucker and Terzopoulos [70] proposed an algorithm for selecting GLCM for texture classification using a $\chi^2$ test. The notion of the structure capturable by GLCM is related to the confidence regarding the variable $x$ given the variable $y$, and vice versa. The null hypothesis of the test that these two variables are independent is in the form of

$$H_0: p(i, j) = p(i, \circ)p(\circ, j), \qquad i = 1, \cdots, N_g$$

where $p(i, \circ)$ is the probability corresponding to the $i$th row and $p(\circ, j)$ is the probability corresponding to the $j$th column. From the derivation presented in [70], the authors arrived at a computationally efficient expression for the test

$$\chi^2 = N \left( \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{p(i, j)^2}{r_i c_j} - 1 \right)$$

where $r_i = \sum_{j=1}^{N_g} p(i, j)$, $c_j = \sum_{i=1}^{N_g} p(i, j)$, and $N = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j)$. The optimal matrix is the matrix that

---

$$P(i, j, 0°) = \#\{((k, l), (m, n)) \in (L_y \times L_x) \times (L_y \times L_x) \Big| k - m = 0, |l - n| = d, I(k, l) = i, I(m, n) = j\}$$

$$P(i, j, 45°) = \#\{((k, l), (m, n)) \in (L_y \times L_x) \times (L_y \times L_x) \Big| (k - m = d, l - n = -d) \text{ or}$$
$$(k - m = -d, l - n = d), I(k, l) = i, I(m, n) = j\}$$

$$P(i, j, d, 90°) = \#\{((k, l), (m, n)) \in (L_y \times L_x) \times (L_y \times L_x) \Big| |k - m| = d, l - n = 0, I(k, l) = i, I(m, n) = j\}$$

$$P(i, j, 135°) = \#\{((k, l), (m, n)) \in (L_y \times L_x) \times (L_y \times L_x) \Big| (k - m = d, l - n = d) \text{ or}$$
$$(k - m = -d, l - n = -d), I(k, l) = i, I(m, n) = j\}$$

yields the highest value of $\chi^2$. As a result, we can determine the displacement and orientation parameters for a certain texture class by simply examining the optimal GLCM. The authors applied the $\chi^2$ test on several Brodatz [9] patterns. It was observed that for values of displacement and orientation that captured texture structure well, the corresponding $\chi^2$ values were high, supporting the validity of the test. As we have mentioned, we used this test in our experiments.

### A. Applications

Haralick *et al.* [35] illustrated the applications of textural features based on GLCM on three different kinds of image data: photomicrographs of different kinds of sandstones [60], panchromatic aerial photographs of land-use categories, and earth resources technology satellite (ERTS) multispectral imagery containing land-used categories. Kruger *et al.* [45] employed GLCM to capture visual texture-context information in an interrib space of X-ray imagery. Chien and Fu [13] computed the average and variance of five augmented GLCM textural features to identify texture changes of X-ray pictures to detect venus hypertension in lung field. Davis *et al.* [21] used generalized GLCM to impose spatial constraints. Shanmugan *et al.* [61] used segments of digitally correlated SEASAT-A SAR imagery in their attempt to classify radar images based on GLCM textural features. Conners *et al.* [19] used texture to segment a high-resolution black and white image of an urban area. Gotlieb and Kreyszig [31] derived a general model for analysis and interpretation of experimental results in texture analysis when raw and composite textural features were used. Barber and LeDrew [5] reported univariate and multivariate analyses in describing the separability of SAR sea ice feature space based on GLCM, tested on a STAR-1 SAR image. Sali and Wolfson [59] used a clustering algorithm based on a generalized Lloyd algorithm and an iterative region merging process based on the phagocytes heuristic to classify SPOT satellite images using GLCM-based textural features. Kushwaha *et al.* [47] used GLCM to classify IRS LISS-II sensor data on forest analysis in northeastern India. Franklin and Peddle [26] improved classification of SPOT HRV imagery for a moderate-relief environment in eastern Canada from 51.1 (spectral alone) to 86.7% (spectral data plus GLCM features). The authors also conducted texture analysis of land systems from Landsat MSS data [25]. Baraldi and Parmiggiani [3] used GLCM to classify SPOT urban areas. Chen and Pavlidis [11] combined a GLCM and a split-and-merge algorithm to segment images in a multiresolution approach. Trivedi *et al.* [66] presented a module that was able to detect fixed orientation objects from a wide variety of backgrounds. They used a supervised parametric method based on a $\chi^2$ distribution to guide a forward sequential search algorithm in the object detection phase. Kovalev and Petrou [44] used multidimensional GLCM to perform classification of various images of CT brain scans, several types of microscope images, and photographs of signatures. Haddon and Boyce [32], [33] proposed one interesting application of co-occurrence matrices in which the matrices were used to detect edges and estimate optic flow field. Beauchemin *et al.* [6]
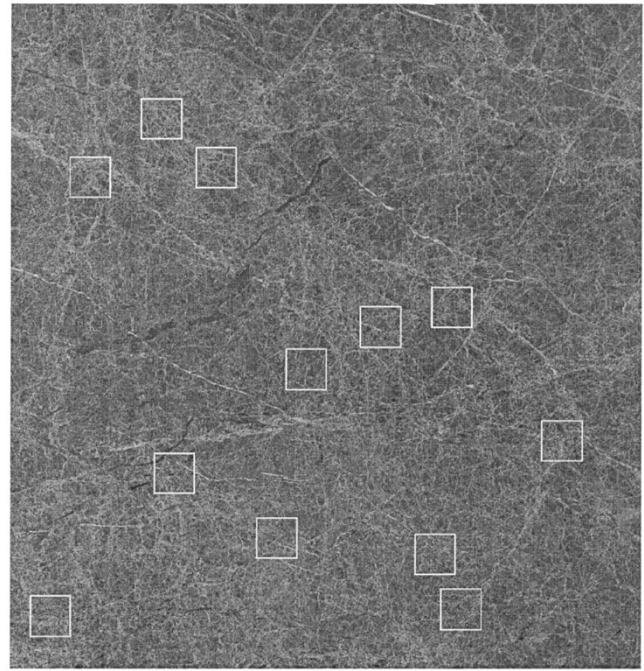


Fig. 1. Image of mostly multiyear ice with heavy ridging and deformation, categorized as Web. Each sample site is $64 \times 64$, or 40.96 km$^2$. © ESA.

used GLCM to design an adaptive speckle-removal filter and an edge detector.

### III. SAR SEA ICE IMAGERY AND TEXTURAL CONTEXTS

We analyzed over 2000 ERS-1 SAR low-resolution images of the Bering, the Beaufort, and the Chukchi seas for every month of the year and compiled seven classes of SAR sea ice textural contexts based on human visual inspection. Note that these seven types of textural contexts do not fully describe all SAR sea ice imagery, and they certainly do not correspond specifically to all ice types. These classes were selected so that we could compare different implementations of GLCM in terms of their classification power and demonstrate the feasibility of GLCM-based textural contexts in differentiating sea ice imagery.

*Class 1—Web:* Fig. 1 shows an image taken on March 27, 1992, at 73.466° N and 156.19° E. The image consists of mostly multiyear ice with heavy ridging and deformation. The web-like structure that these ridges or deformations build with each other characterizes this type of images.

*Class 2—3-D:* Fig. 2 shows an image taken on February 7, 1993, at 58.54° N and 163.63° W. The image consists of crushing floes creating extreme deformations in the marginal ice zone (MIZ). Boundaries of floes have been roughened such that they give high backscatter return. These thick enclosing structures portray a three-dimensional (3-D) perceptual effect that characterizes this type of image.

*Class 3—Fractal:* Fig. 3 shows an image taken on September 8, 1993, at 77.35° N and 145.79° W. The image consists of mostly new ice and melt ponds. This phenomenon occurs at the end of the summer melt season when ice floes have been broken up, brushed, and rubbled. Boundaries of floes
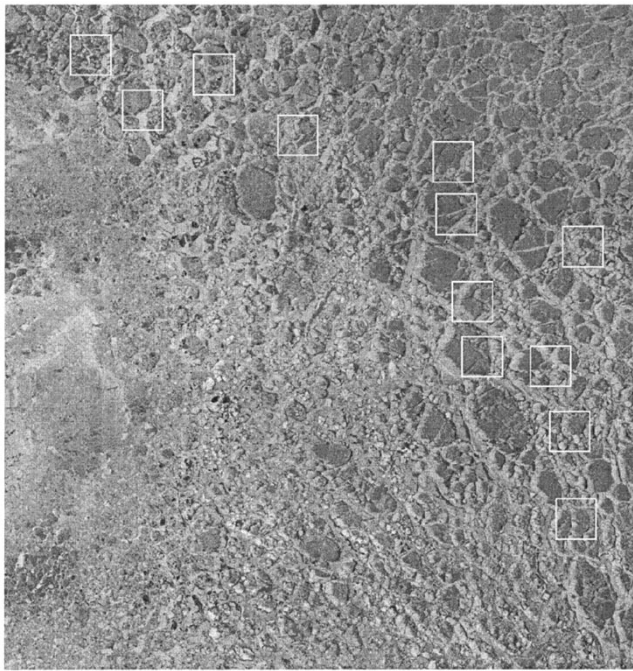
Fig. 2. Image of crushing floes with extreme deformations at MIZ, categorized as 3-D. Each sample site is $64 \times 64$, or 40.96 km$^2$. © ESA.
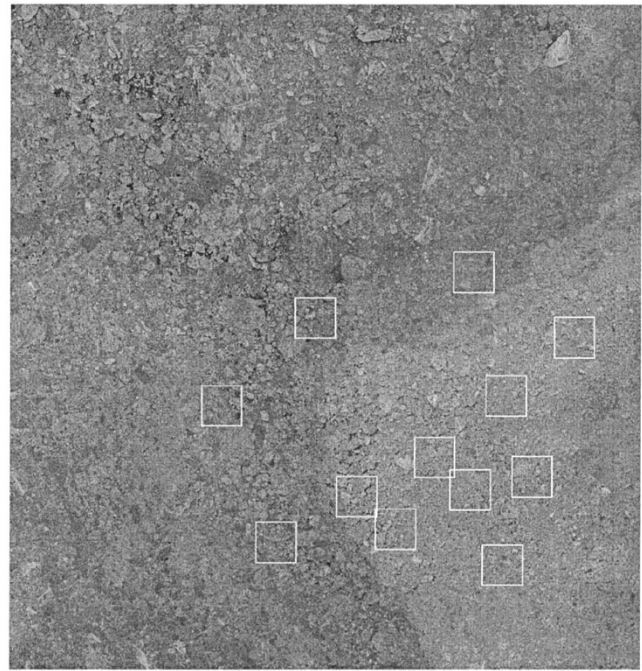


Fig. 4. Image of thawing multiyear (small circular structures) and first-year ice, categorized as pebble-like. Each sample site is $64 \times 64$, or 40.96 km$^2$. © ESA.
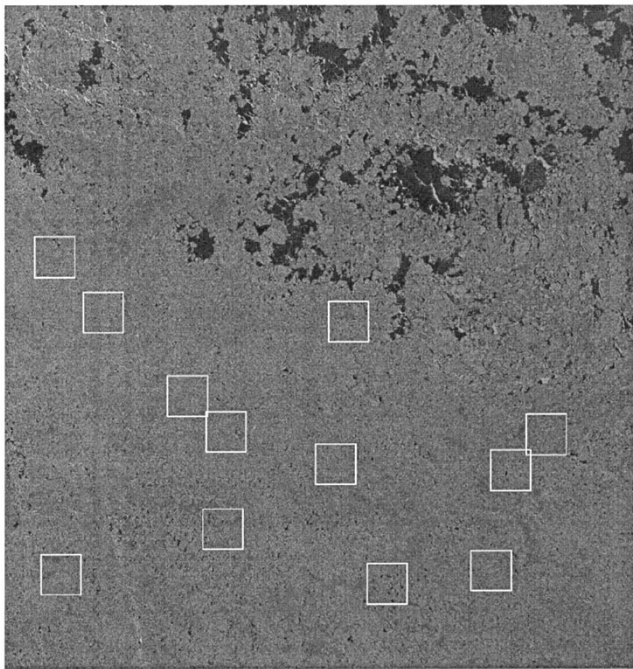


Fig. 3. Image of mostly new ice and melt ponds, categorized as fractal. Each sample site is $64 \times 64$, or 40.96 km$^2$. © ESA.



Fig. 5. Image of highly undeformed ice features, categorized as smooth. Each sample site is $64 \times 64$, or 40.96 km$^2$. © ESA.

look foamy and wiggly fractal, and it is this property that characterizes this type of images.

*Class 4—Pebble-Like:* Fig. 4 shows an image taken on July 12, 1993, at 72.26° N and 160.47° W, showing the start of the summer melt season. Multiyear ice can be seen as little round features; younger ice as quite homogeneous but grainy patches. Pebbles of ice floes are embedded in about-to-melt younger ice formations. The pebbles characterize this type of images.

*Class 5—Smooth:* Fig. 5 shows an image taken on February 1, 1994 at 72.06° N and 176.39° W. This image shows very smooth multiyear and first-year ice with bright refrozen leads cutting across the region. As we can see, dark patches (probably multiyear) and bright sheets (probably young ice) are extremely homogeneous and uniform. These are considered as highly undeformed ice features, characterizing this type of images.

Fig. 6.   Image of regions with substantial size and of two highly contrasting intensities, categorized as high-contrast. Each sample site is 64 × 64, or 40.96 km$^2$. © ESA.
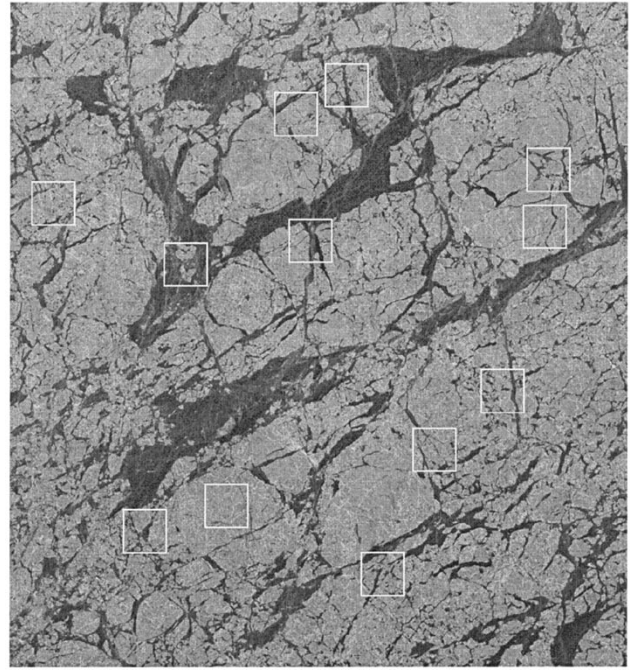


Fig. 7.   Image of slightly ridged but mostly undeformed multiyear ice, and with good three-class separation (MY, FY, and OW/NI), categorized as packed. Each sample site is 64 × 64, or 40.96 km$^2$. © ESA.

*Class 6—High Contrast:* Fig. 6 shows an image taken on November 17, 1993, at 72.27° N and 154.75° W. This image shows on the one hand large dark multiyear ice floes and on the other hand, large refrozen young first-year ice and thin pancake ice. This combination of phenomena tells us that the floes had been mobile to create water lodgings and stationary enough for pancake or young ice to form. These regions with substantial size and of two highly contrasting intensities characterize this type of images.

*Class 7—Packed:* Fig. 7 shows an image taken on March 17, 1992, at 72.85° N and 143.83° W. This image shows a piece of packed multiyear ice broken by large leads. There are refrozen leads developing in some areas, and open water regions are rare. These multiyear ice conglomerates are slightly ridged but mostly undeformed. This context is frequently observed around high latitude regions or in the middle period of winter since floes are frozen and relatively immobile.

## IV. QUANTIZATION, DISPLACEMENT, AND ORIENTATION FACTORS ON GLCM

There are several important parameters to consider when designing a GLCM, as follows: 1) the region size, 2) the quantization levels, $N_g$, 3) the displacement value $d$, and 4) the orientation value $\theta$. The region size gives the dimensions of the region of which GLCM is computed. In Haverkamp *et al.* [36], the region sizes of 32 × 32 and 64 × 64 were used to perform dynamic local thresholding on SAR sea ice imagery successfully. In order to capture sea ice textural contexts, we prefer the larger region size and set it to 64 × 64 during our experiments with GLCM. In the following, we will examine the quantization and displacement factors directly and the orientation factor indirectly in Section V.

The approach of our experiment is as follows. For quantization, we compute a set of quantized images of each sample site. Then, a second-order Euclidean distance measure between each successive pair of the quantized images is computed based on the second-order statistics of GLCM of the quantized images. The visual observation of the measurements serves as the basis for our conclusions regarding the quantization factors. For displacement, once again, we compute a set of quantized images of each sample site. Then we compute various second-order GLCM statistics of the quantized images using a set of displacement values. We visually inspect the curves generated by plotting the statistics against the displacement values for each sample site and draw conclusions from the observation. Finally, for orientation, we design three implementations of GLCM and by comparing the classification powers of the design, we are able to draw some conclusions regarding the orientation factors for sea ice textural contexts.

### A. Quantization

The number of gray levels is an important factor in the computation of GLCM. The decision that we have to make is how many levels are needed to represent a set of textures successfully. The more levels included in the computation, the more accurate the extracted textural information, with, of course, a subsequent increase in computation costs since the quantization scheme smoothes an image and thus reduces noise-induced effects to some degree. We assume that the information gain in noise-effect reduction does not compensate the loss of information as a result of quantization. There are three major quantization schemes: 1) uniform quantization, 2) Gaussian quantization, and 3) equal probability quantization. The uniform quantization scheme is the simplest, in which

gray levels are quantized into separate bins with uniform tolerance limits or spaces with no regard to the gray-level distribution of the image. This technique is not always preferable. For example, if a range of gray levels occurs more frequently than others, we might desire to finely quantize that particular range. The Gaussian quantization technique is one such scheme. According to this scheme, the gray-level distribution of the original image is assumed to behave normally. Each quantization bin has the same area under the curve and thus different space smaller spaces in the middle of the distribution and larger spaces at the tails of the distribution. The equal probability quantization [35] scheme has also been used such that each bin has a similar probability, and it has been shown to retain an accurate representation of the original picture in terms of textural features based on GLCM [17]. The Gaussian quantization scheme assumes a Gaussian gray-level distribution, which is not always true for SAR sea ice imagery. Equal probability quantization normalizes different image samples so that a bright feature and a dark feature, given the same texture, would have the same co-occurrence matrix, which is undesirable since the backscatter of sea ice types is a key parameter in sea ice analysis. Thus, in our experiment, we have focused on the uniform quantization scheme.

We extracted sample sites, each $64 \times 64$, from 18 images with different textural contexts in this experiment. We devised a test using six textural features (energy, contrast, correlation, entropy, autocorrelation, and homogeneity) and six different uniform quantization schemes: 8, 16, 32, 64, 128, and 256 levels. We set displacement to one and orientations to 0, 45, 90, and 135°. Taking the average of the orientations for each image sample yields a $6 \times 6$ matrix. Each entry of this textural vector is defined as $V_{i,j,m}$, where $i = 1 \cdots 6$, $j = 1 \cdots 6$, and $m = 1 \cdots M$. $i$ is the index of the textural features, $j$ is the index of the quantization schemes (such that scheme $j$ has $2^{2+j} = 4 \cdot 2^j$ quantization levels), and $m$ denotes the site label ($M$ = total number of sample sites). The different values of the quantization are evaluated by five measures based on Euclidean distance along each textural feature between each pair of sample sites. Note that the objective of the following Euclidean distance measures is to provide a visual presentation of the trend between the differences between each successive pair of quantization schemes. These measures were designed to capture such differences without assuming the distribution or any prior knowledge of the samples. Therefore, statistical distance measures (that are also computationally more complicated), such as Mahalanobis [51], Kolmogorov [1], Bhattacharyya [7], Bayesian distance [8], Chernoff [12], Matsusita [52], and divergence [40], [46], and those utilized in pattern recognition [50], [57], have not been used in our experiment.

The first measure is the normalized first-order differentiation

$$D_{i,j,m,n}^k = \frac{V_{i,j,m} - V_{i,j,n}}{V_{i,j,m}}; \quad k \text{ is the image label.}$$

This value measures the normalized difference between two sample sites from the same image in terms of a particular textural feature. The collective results for each sample pair is a $6 \times 6$ matrix. Observing this matrix gives us an idea how the degree of dissimilarity between two samples varies with the number of quantization levels.

The second measure is the neighbored second-order autodifferentiation

$$N_{i,j}^{kk} = \frac{2}{M^2} \sum_{m=1}^{M} \sum_{n=1}^{M} |D_{i,j,m,N}^k - D_{i,j+1,m,n}^k|,$$
$$i = 1 \cdots 6 \quad \text{and} \quad j = 1 \cdots 5.$$

This value measures the average absolute difference between the normalized first-order differentiation between two neighboring quantized schemes, taken over all sample pairs of the same image. This matrix tells us whether there exists a systematic relationship between two neighboring quantized schemes.

The third measure is the biased second-order autodifferentiation

$$B_{i,j}^{kk} = \frac{2}{M^2} \sum_{m=1}^{M} \sum_{n=1}^{M} |D_{i,j,m,n}^k - D_{i,6,m,n}^k|,$$
$$i = 1 \cdots 6 \quad \text{and} \quad j = 1 \cdots 5.$$

This value measures the average absolute difference between the normalized first-order differentiation between a $2^{2+j}$-quantization and the 256-quantization (or the original image, in our case) scheme, where $j = 1 \cdots 5$, taken over all sample pairs of the same image. This matrix tells us which quantization scheme best imitates the result of 256-level quantization scheme.

The fourth measure is the neighbored second-order cross differentiation

$$N_{i,j}^{kl} = \frac{2}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} |D_{i,j,m,n}^k - D_{i,j+1,m,n}^l|,$$
$$i = 1 \cdots 6, j = 1 \cdots 5, \quad \text{and} \quad k \neq l.$$

Note that $N$ is the total number of sample sites in the image with the label $l$. This value measures the average absolute difference between the normalized first-order differentiation between two neighboring quantization schemes, taken over all sample pairs of two different images.

The fifth measure is the biased second-order cross differentiation

$$B_{i,j}^{kl} = \frac{2}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} |D_{i,j,m,n}^k - D_{i,6,m,n}^l|,$$
$$i = 1 \cdots 6, j = 1 \cdots 5, \quad \text{and} \quad k \neq l.$$

This value measures the average absolute difference between the normalized first-order differentiation between a $2^{2+j}$-quantization and the 256-quantization (or the original image, in our case) scheme, where $j = 1 \cdots 5$, taken over all sample pairs of two different images.

Fig. 8 shows one graphical example of the normalized first-order differentiation matrix. Note that this graph shows the differences between two sample sites in terms of six textural features when the sites are quantized with eight levels, 16 levels, and so on, up to 256 levels (the original). We can see that for the eight-level quantization scheme, the differentiation
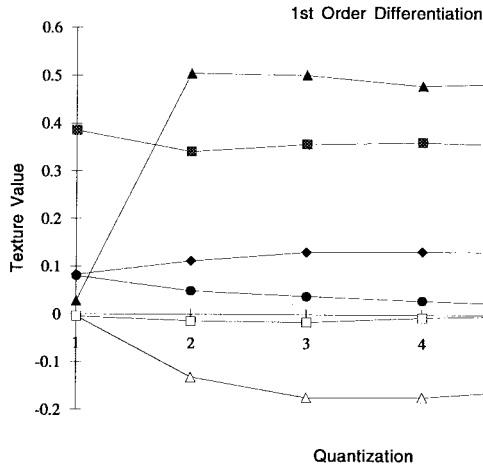
Fig. 8. Example of the first-order differentiation of textural feature measurements versus quantization schemes. There is no noticeable trend of how degree of similarity between two samples varies with the number of quantization levels. Note: quantization scheme 1 denotes eight-level quantization; 2 denotes 16-level, and so on. Filled triangle: entropy, filled square: autocorrelation, filled diamond: contrast, filled circle: correlation, square: homogeneity, and triangle: energy.
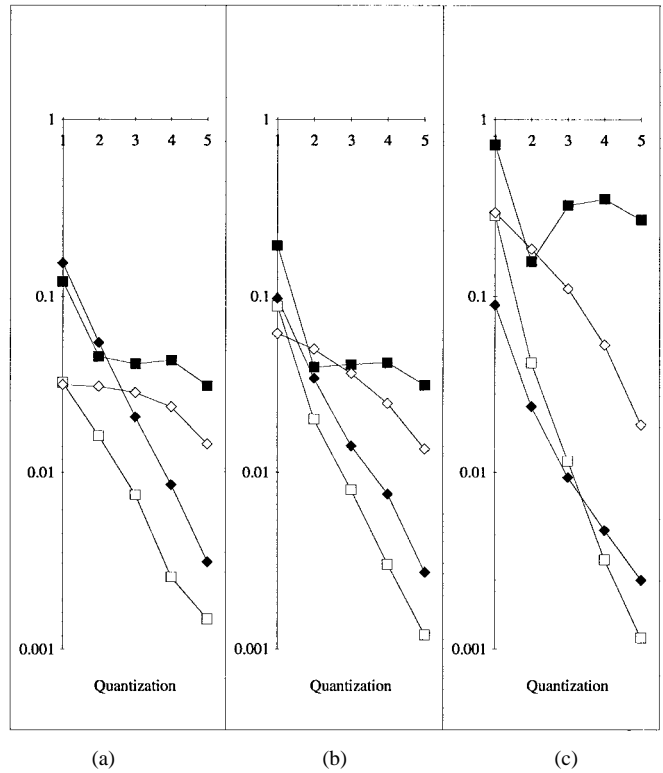


Fig. 9. Example of (a) neighbored second-order autodifferentiation, (b) neighbored second-order cross differentiation (within same context), and (c) neighbored second-order cross differentiation (with different contexts). All plots are in log scale. Difference between two neighboring feature measurements is generally smaller as the number of quantization levels increases. Note: quantization scheme 1 denotes eight-level quantization; 2 denotes 16-level, and so on. Graphs show energy (filled square), contrast (square), autocorrelation (filled diamond), and homogeneity (diamond).

values are noticeably different from others. Also, we can see that, for other quantization levels, the differences between the two sample sites are very similar. This hints that if our objective is to distinguish between two sea ice contexts, the number of quantization levels (excluding eight) do not matter. To analyze whether the textures can be used to represent sea ice contexts, we turn to second-order differentiation values. Fig. 9 shows some examples of our neighbored second-order differentiation: (a) $N_{i,j}^{kk}$, (b) $N_{i,j}^{kl}$ where images $k$ and $l$ are of the same context, and (c) $N_{i,j}^{km}$ where images $k$ and $m$ are of different contexts. Note that only four textural features are shown so the scale of the texture value can be displayed adequately for view. Generally, the distance between features of two neighboring quantization schemes decreases as more quantization levels are involved $N_{i,j}^{kk} < N_{i,j-1}^{kk}$ for all $j > 1$. This indicates that, if we assume the original image to be accurate, information loss to quantization is gradual: more when the number of quantization levels is smaller and vice versa. From another perspective, if we choose a high enough number of quantization levels, we will be able to preserve enough information found in the image. In addition, we observe that such distance is the smallest for the samples from the same image and the largest for two samples from different textural contexts $N_{i,j}^{kk} < N_{i,j}^{kl} < N_{i,j}^{km}$, $|N_{i,j}^{kk} - N_{i,j}^{kl}| < |N_{i,j}^{kl} - N_{i,j}^{km}|$. This indicates that not only GLCM can recognize different contexts, whether they are from same or different images, but it can do it for a different number of quantization levels. Fig. 10 shows some examples of biased second-order differentiation measures: (a) $B_{i,j}^{kk}$, (b) $B_{i,j}^{kl}$ where images $k$ and $l$ are of the same context, and (c) $B_{i,j}^{km}$ where images $k$ and $m$ are of different contexts, respectively. Generally, we can derive similar observations, as discussed in the neighbored case $B_{i,j}^{kk} < B_{i,j-1}^{kk}$, for all $j > 1$, $B_{i,j}^{kk} < B_{i,j}^{kl} < B_{i,j}^{km}$, $|B_{i,j}^{kk} - B_{i,j}^{kl}| < |B_{i,j}^{kl} - B_{i,j}^{km}|$. So, while the representative power of the GLCM deteriorates with increased quantization, the discriminative power of GLCM

does not, if used in characterizing and recognizing sea ice contexts.

Several conclusions have been drawn from the experiments, as follows.

1) There is no noticeable trend to indicate that the degree of dissimilarity between two samples varies with the number of quantization levels. But there are noticeable differences, more random than systematic.

2) There is a noticeable trend that shows diminishing difference when comparing results of a pair of larger numbers of quantization levels. This indicates that the GLCM-based results are more consistent when using a higher number of quantization levels. This disagrees with the findings of [53]. The probable cause of this disagreement lies at the number of quantization levels analyzed: in [53], the authors used *only* the 16- and 32-level quantization schemes and found their textural features to be similar; and we compared quantization levels to the original image.

3) Degree of dissimilarity is the smallest when the samples are of the same image, small when the samples are of two images of the same context, and the largest when the samples are of two images of different contexts. This indicates that the GLCM can be utilized to differentiate two different contexts and recognize two similar contexts.
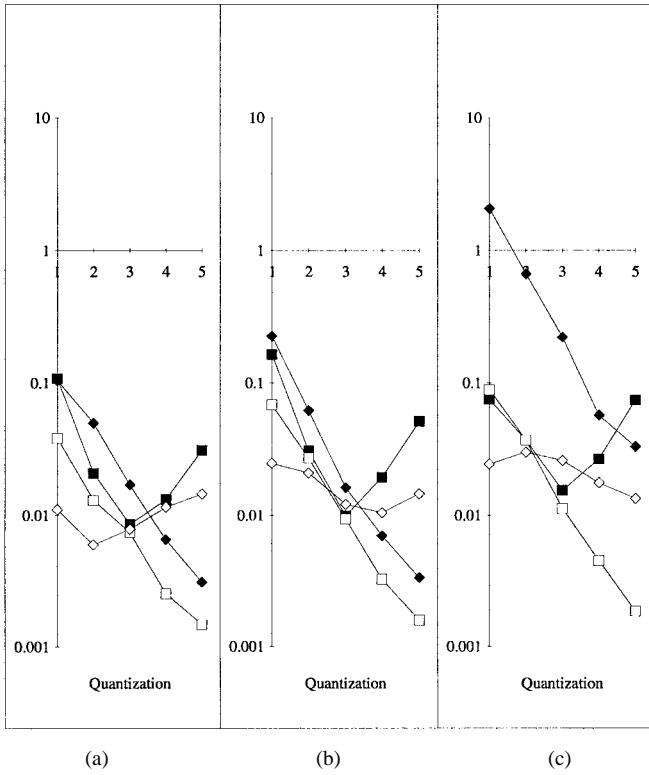
Fig. 10. Example of (a) biased second-order autodifferentiation, (b) biased second-order cross differentiation (within same context), and (c) biased second-order cross differentiation (with different contexts). All plots are in log scale. Second-order differentiation values between other feature measurements and those of 256-level quantization is generally smaller as the number of quantization levels increases. Note: quantization scheme 1 denotes eight-level quantization; 2 denotes 16-level, and so on. Graphs show energy (filled square), contrast (square), autocorrelation (filled diamond), and homogeneity (diamond).

4) As expected, as the number of quantization levels increases, features imitate those of the original quantization more closely. This trend is clearly noticeable for all sample pairs.

5) The eight-level quantization scheme is consistently and noticeably different from other schemes. Thus, it should not be used.

## B. Displacement

The displacement parameter $d$ is important in the computation of GLCM. Applying a large displacement value to a fine texture would yield a co-occurrence matrix that does not capture the textural information, and vice versa. Chen *et al.* [10] used $d = 1, 2, 4, 8, 16, 32, 64$ and found that overall classification accuracies with $d = 1, 2, 4, 8$ were essentially equivalent in differentiating cloud types. However, for higher displacement values, the authors found that the classification accuracies decreased. They also concluded that the classification result was best when using features from matrices of $d = 1, 2$. This indicates that single-displacement features might not be sufficient to represent textures. Another study [22] showed that a displacement value equal to the size of the texture element would tend to improve the classification result with texture features. For our experiment, we used $d = 1, 2, \cdots, 32$. We used 32 as our upper limit of range as
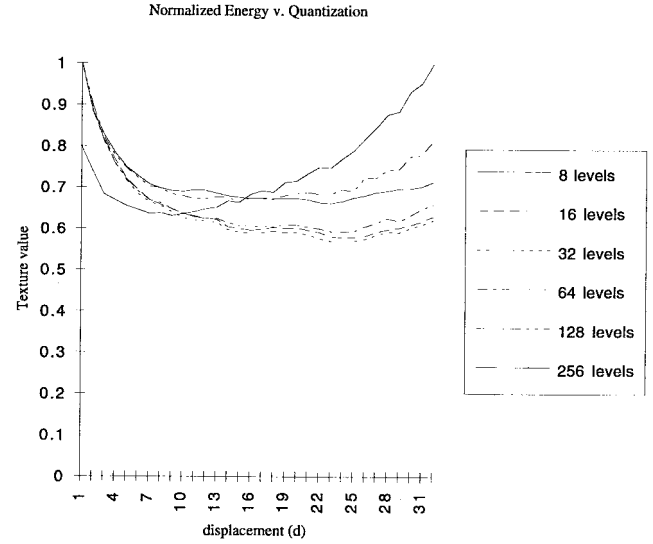


Fig. 11. Graph of an example of normalized energy for all displacement curves across all quantization schemes.



Fig. 12. Graph of an example of normalized contrast for all displacement curves across all quantization schemes.

it was half the region size of each sample. To better compare the texture values across different quantization schemes, we normalized each series by dividing it by its maximum value. Figs. 11–18 show the graphs of normalized texture values versus quantization schemes. In general, we can see that across different quantization schemes, displacement curves preserve nicely. In particular, for contrast and dissimilarity, all displacement curves, after normalization, are almost identical. This indicates that quantization does not affect the two textural features. For energy, correlation, homogeneity, entropy, and autocorrelation, the behaviors of the curves are similar in terms of locations of upward and downward slopes along the curves. The differences are in the degree of these slopes, which are probably caused by reduced resolutions in the reduction of quantization levels. This indicates that, although the loss of information is visible, the inherent structures of relationships among pixels for these textural features are still intact and

Fig. 13.   Graph of an example of normalized correlation for all displacement curves across all quantization schemes.

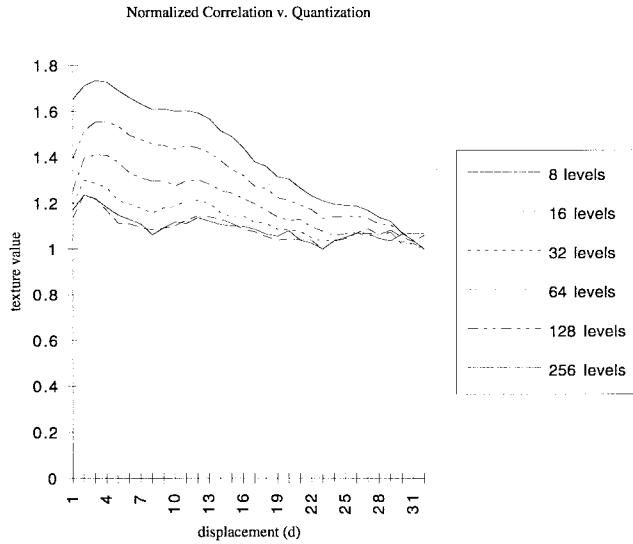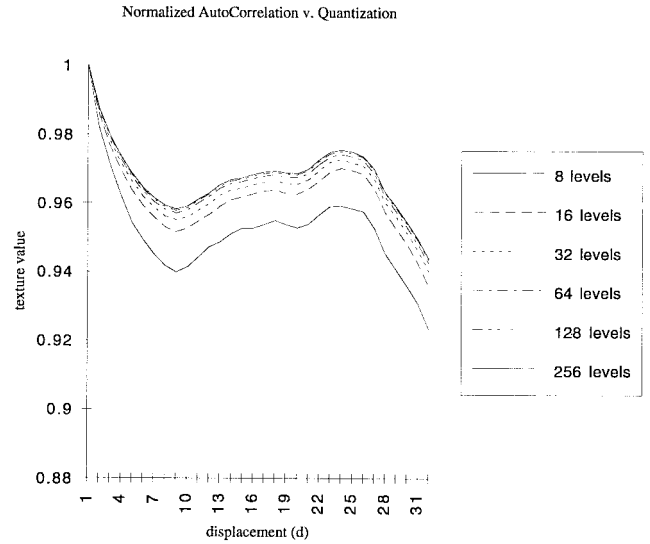Fig. 15.   Graph of an example of normalized entropy for all displacement curves across all quantization schemes.
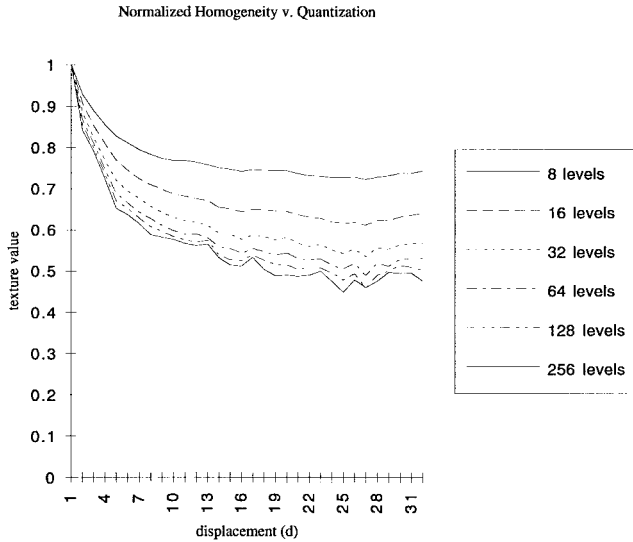
Fig. 14.   Graph of an example of normalized homogeneity for all displacement curves across all quantization schemes.
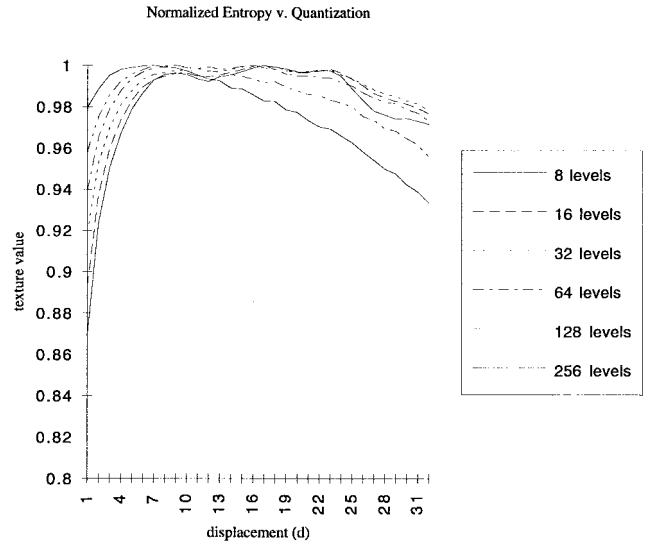
Fig. 16.   Graph of an example of normalized autocorrelation for all displacement curves across all quantization schemes.

can be captured and represented with GLCM. For maximum probability, peaks and valleys are more prominent as the resolution of the sample site increases. Note that the definition of max probability is $f_{10} = \mathrm{MAX}_{i,j}\, p(i,j)$. Unlike others, it is not a function of an aggregate of $p(i,j)$. Thus, the value of this parameter is more volatile and susceptible to loss of information due to quantization. As we can see from Fig. 18, peaks and valleys have been lost, and when the number of quantization levels reduces to eight (the thin solid line), the displacement curve becomes the smoothest. In general, all displacement curves still follow a similar trend.

Several conclusions have been drawn from this experiment, as follows.

1) Across quantization schemes, each textural curve preserves nicely. This hints that the number of quantization levels can be arbitrarily chosen as long as a range of displacement values are used in computing the GLCM, and thus, we eliminate the need for 128-level and 256-level schemes, which are computationally very costly. In [10], the authors showed that classification of cloud fields was still successful even with the highly quantized version of GLCM features. Compared to that finding, our observation is a more restricted case.

2) Instead of a multiresolution approach that coordinates different quantization schemes, it is sufficient to use one quantization scheme with a range of displacement values since the dynamics of curves of different number of quantization levels are similar.

3) There is no hint on determining *a priori* a general displacement value for all samples in which each sample approaches an asymptotic plateau. This indicates that it is necessary to compute matrices of different displacement values to obtain accurate textural features (even
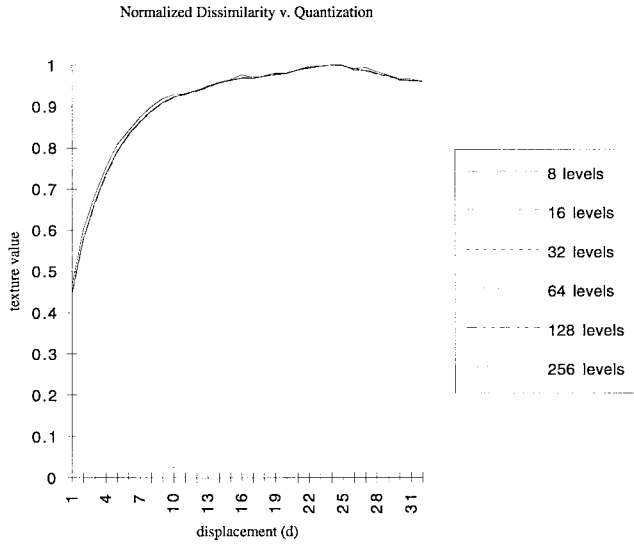
Normalized Dissimilarity v. Quantization



Fig. 17. Graph of an example of normalized dissimilarity for all displacement curves across all quantization schemes.

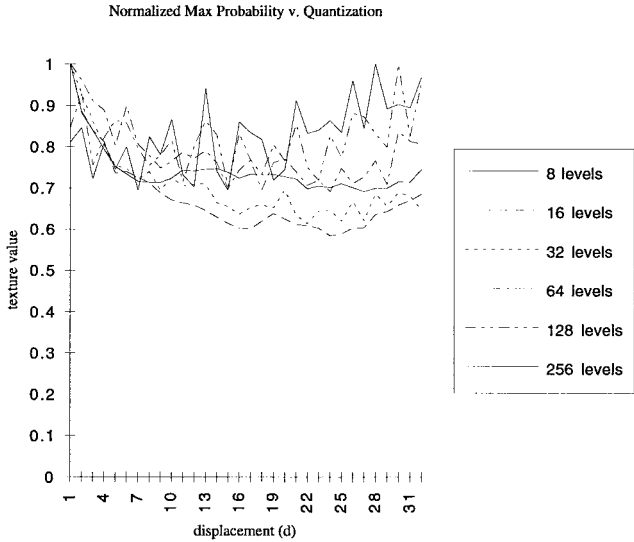Normalized Max Probability v. Quantization



Fig. 18. Graph of an example of normalized maximum probability for all displacement curves across all quantization schemes.

though it is not necessary to use all such matrices). *A single displacement value for GLCM to represent sea ice textural contexts is not advisable.*

4) Averaging features over a range of displacement values might be used to obtain a reliable and economical representation, especially for asymptotic curves. For example, each sample has four features, eight displacement values, and four quantization schemes, yielding a total of 128 values. The averaged version has only 16 values.

### C. Orientation

The orientation parameter $\theta$ is relatively less important compared to other factors in co-occurrence matrices. Some authors used the average and range, some used certain series of orientations; for example, $\theta = 19, 75, 90, 109,$ and $165°$ to accommodate man-made urban structures [19], range and

average of $\theta = 0, 45, 90,$ and $135°$ [35], average of two by $126°$ apart, average of three by $66°$ apart, and average of four by $37°$ apart [56] average of $\theta = 0, 45, 90,$ and $135°$ [38], average, variance and range of $\theta = 0, 45, 90,$ and $135°$ [45], and prespecified orientation for each image [66].

For SAR sea ice imagery, there are no systematic patterns based on orientation. Ice features rotate and position themselves in all possible orientations. Therefore, we argue that the orientation factor is not important in SAR sea ice research. We used $\theta = 0, 45, 90,$ and $135°$ for they cover efficiently all directions of SAR sea ice imagery, as we shall show in the next section.

## V. THREE CO-OCCURRENCE MATRICES FOR SEA ICE TEXTURAL CONTEXT

In this experiment, we designed and compared three different GLCM implementations by subjecting them to training and testing of a database of sea ice sample sites. We evaluated the classification powers of the designs based on a Bayes classifier. Based on the classification accuracies, we were able to derive the importance of the orientation factor and an effective way of utilizing displacement values.

The first design is the MDMO matrix. Feature measures of the matrices of the four orientations of 0, 45, 90, and 135° are averaged, and further averaged over the displacement range. Note that the average over the displacement range is our choice of aggregating feature measures of different displacement values. Other approaches, such as curve approximation or parametric polynomial modeling, can be used to describe the displacement curves for a more complete representation. Note that other higher level treatment of features of different displacement values is also possible, such as principal component analysis of the eigenvector of different features.

Let the set of all displacement values be $\Delta$ and the set of all orientations be $\Theta$; let $M(d, \theta)$ be the co-occurrence matrix computed with displacement $d$ and orientation $\theta$; and let $f_i(d, \theta)$ be the feature $i$ computed with displacement $d$ and orientation $\theta$. Formally

$$F_i^{\text{MDMO}} = \frac{1}{|\Delta|} \sum_{d \in \Delta} \frac{1}{|\Theta|} \sum_{\theta \in \Theta} f_i(d, \theta)$$

where $|\bullet|$ is the number of members in the set $\bullet$. The MDMO implementation assumes that every matrix of specific displacement and orientation is partially and cumulatively representative for the sample.

The second GLCM is the $\chi^2$-ODMO matrix. $\chi^2$ values of all four matrices of different orientations are calculated and averaged for each displacement $d$, and the matrix accumulating the most $\chi^2$ value is the optimal matrix. Let $\chi^2(M(d, \theta))$ be the $\chi^2$ of $M(d, \theta)$. Formally

$$f_i^{\text{ODMO}} = \frac{1}{|\Theta|} \sum_{\theta \in \Theta} f_i(d, \theta) \Bigg|$$

$$\sum_{\theta \in \Theta} \chi^2(M(d, \theta)) = \max_{k \in \Delta} \left( \sum_{\theta \in \Theta} \chi^2(M(k, \theta)) \right).$$

The ODMO implementation assumes that only the matrix whose $\chi^2$ value is the highest with a specific displacement value is truly and sufficiently representative for the sample, with no regard to selective orientation.

The third GLCM is the $\chi^2$-optimal displacement and ODOO matrix. Formally

$$f_i^{\mathrm{ODOO}} = f_i(d_o, \theta_o)\Big|$$
$$\chi^2(M(d_o, \theta_o)) = \max_{d \in \Delta; \, \theta \in \Theta}(\chi^2(M(d, \theta))).$$

The ODOO implementation assumes that the matrix whose $\chi^2$ value is the highest with specific displacement and orientation is truly and sufficiently representative for the sample.

We elected to use the ability of a GLCM to classify between different textures as a metric of its efficacy. Classification was performed using a Bayesian supervised classifier. Using all ten features outlined in Section II, we have for each sample three vectors of ten entries $\hat{f}^{\mathrm{MDMO}}$, $\hat{f}^{\mathrm{ODMO}}$, and $\hat{f}^{\mathrm{ODOO}}$ grouped into three groups for all samples, $\hat{G}^{\mathrm{MDMO}}$, $\hat{G}^{\mathrm{ODMO}}$, and $\hat{G}^{\mathrm{ODOO}}$, respectively. For each group of data $\hat{G}$, samples of seven textural contexts, $\hat{G}_k$ for $k = 1 \cdots 7$, are used to train the Bayes classifier. For each context $k$, we calculated its discriminant function

$$h_k(\hat{x}) = \tfrac{1}{2}(\hat{x} - \hat{\mu}_k)^T \sum_k{}^{-1}(\hat{x} - \hat{\mu}_k) + \log\left|\sum_k\right|$$

where $\hat{x}$ is the unknown sample vector of $x$, $\hat{\mu}_k$ the calculated mean vector for $\hat{G}_k$ and $\sum_k$ is the calculated covariance for $\hat{G}_k$. The classifier assigns a sample $x$ to $k$ if

$$h_k(\hat{x}) > h_j(\hat{x}) \text{ for all } k \neq j.$$

Note that, in all of the experiments that follow, we used the uniform 64-level quantization scheme. This number of quantization levels was chosen since, used over a range of displacement values, it preserves textural information well and it does not incur as much computational load as other higher level quantization schemes.

### A. Experimental Evaluation of GLCM's

In our first experiment, we used ten features, 228 samples, 18 images, and six contexts. The resubstitution classification or training set classification accuracies of the MDMO, ODMO, and ODOO implementations are tabulated in Table I. The numbers represent the accuracy by which a specific GLCM can classify a sample into one of six contexts. Consequently, it is a metric of representational quality. From the table, we observe that the MDMO implementation has the best classification results on the training set with overall accuracy above 90%.

After generating features for all samples, we noticed that the cluster shade and cluster prominence measurements show higher range and range/mean (normalized range) values than the other measurements (Table II). This is probably due to the third and fourth order of moments used in the computation,

TABLE I
CLASSIFICATION ACCURACY AFTER SUPERVISED TRAINING, BASED ON THE TRAINING DATA SETS. MDMO's CLASSIFICATION ACCURACY IS BETTER THAN THE OTHER TWO FOR ALL CLASSES, EXCEPT FOR CLASS 1. ALL THREE IMPLEMENTATIONS FAILED TO CLASSIFY CLASS 2. TEN FEATURES WERE USED

| Context | MDMO | ODMO | ODOO |
|---|---|---|---|
| 1 | 88.33% (53/60) | 100.00% (60/60) | 43.33% (26/60) |
| 2 | 16.67% (2/12) | 16.67% (2/12) | 16.67% (2/12) |
| 3 | 95.83% (46/48) | 81.25% (39/48) | 83.33% (40/48) |
| 4 | 94.44% (34/36) | 72.22% (26/36) | 94.44% (34/36) |
| 5 | 100.00% (12/12) | 75.00% (9/12) | 75.00% (9/12) |
| 6 | 100.00% (60/60) | 10.00% (6/60) | 11.67% (7/60) |
| Total | 90.79% (207/228) | 62.28% (142/228) | 51.75% (118/228) |

TABLE II
FEATURES FOR A CLASS 1 SAMPLE. NOTE THAT BOTH CLUSTER SHADE AND PROMINENCE HAVE THE HIGHEST RANGE AND RANGE/MEAN VALUES. THIS CHARACTERISTIC CREATES A PROBLEM FOR OUR BAYESIAN CLASSIFIER

| | Energy | Contrast | Correlation | Homogeneity | Entropy |
|---|---|---|---|---|---|
| high | 0.00490 | 88.098 | -370.357 | 0.192 | 2.831 |
| low | 0.00180 | 46.812 | -867.525 | 0.127 | 2.500 |
| mean | 0.00296 | 59.120 | -652.684 | 0.155 | 2.647 |
| std. dev. | 0.00054 | 8.475 | 118.060 | 0.012 | 0.062 |
| range | 0.00309 | 41.286 | 497.168 | 0.065 | 0.331 |
| range/mean | 1.04586 | 0.698 | 0.762 | 0.418 | 0.125 |

| | Auto-correlation | Dissimilarity | Cluster Shade | Cluster Prominence | Maximum Probability |
|---|---|---|---|---|---|
| high | 1048.556 | 7.472 | 495.085 | 34639.106 | 0.0144 |
| low | 740.179 | 5.221 | -170.840 | 8356.405 | 0.0048 |
| mean | 907.826 | 6.051 | 149.718 | 14398.884 | 0.0075 |
| std. dev. | 65.352 | 0.461 | 105.173 | 4744.914 | 0.0015 |
| range | 308.377 | 2.251 | 665.925 | 26282.701 | 0.0096 |
| range/mean | 0.340 | 0.372 | 4.448 | 1.825 | 1.2914 |

TABLE III
RESUBSTITUTION CLASSIFICATION ACCURACY AFTER SUPERVISED TRAINING, BASED ONLY ON THE DATA SETS. MDMO's CLASSIFICATION ACURACY WAS IMPROVED WHEN CLUSTER SHADE AND PROMINENCE WERE EXCLUDED FROM THE FEATURE SET

| Context | MDMO (10 features) | MDMO (8 features) |
|---|---|---|
| 1 | 88.33% (53/60) | 88.33% (53/60) |
| 2 | 16.67% (2/12) | 100.00% (12/12) |
| 3 | 95.83% (46/48) | 93.75% (45/48) |
| 4 | 94.44% (34/36) | 86.11% (31/36) |
| 5 | 100.00% (12/12) | 100.00% (12/12) |
| 6 | 100.00% (60/60) | 100.00% (60/60) |
| Total | 90.79% (207/228) | 93.42% (213/228) |

respectively. Using higher order moments amplifies the surface of a cluster. If the cluster was roughly homogeneous, the resulting feature value would be consistent and occupy a compact range. As we can see from Table II, both normalized ranges are higher for cluster shade and prominence. This hints that the quantized images do not possess good cluster-type characteristics. We conducted an experiment to compare the classification results between MDMO with all ten features and MDMO with eight features, excluding cluster shade and prominence. The results are shown in Table III. Overall, the exclusion of cluster shade and prominence improved our classification accuracy. Class 2 classification accuracy was improved from 16.67 to 100.00%. As a result, we believe that the cluster-type features (cluster shade and cluster prominence) tend to extenuate the clustering characteristic of samples and dominate classification resulting in large errors.

TABLE IV
Resubsitution Classification Accuracy After Supervised Training, Based Only on the Data Sets. MDMO's Classification Acuracy Is Better than the Other Two for All Classes. Eight Features Were Used. Seven Contexts Were Represented

| Context | MDMO | ODMO | ODOO |
|---|---|---|---|
| 1 | 95.00% (57/60) | 86.67% (52/60) | 93.33% (56/60) |
| 2 | 100.00% (12/12) | 8.33% (1/12) | 50.00% (6/12) |
| 3 | 95.91% (47/48) | 85.42% (41/48) | 81.25% (39/48) |
| 4 | 97.22% (35/36) | 94.44% (34/36) | 97.22% (35/36) |
| 5 | 100.00% (12/12) | 50.00% (6/12) | 25.00% (4/12) |
| 6 | 100.00% (60/60) | 98.33% (59/60) | 91.67% (55/60) |
| 7 | 100.00% (12/12) | 8.33% (1/12) | 0.00% (0/12) |
| Total | 97.92% (235/240) | 80.83% (194/240) | 81.25% (195/240) |

TABLE V
Confusion Matrices. The Rows Denote the Actual Class of the Sample Sites, and the Columns Denote the Class Found by the Classifier. (a) MDMO with Ten Features. (b) ODOO with Eight Features. (c) MDMO with Eight Features. (d) MDMO with Eight Features. A Perfect Classifier Would Have Nonzero Entries Only in the Diagonal of the Matrix

| C | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| 1 | 53 | | | 4 | | 3 | |
| 2 | | 2 | | | | 10 | |
| 3 | | | 46 | 1 | | 1 | |
| 4 | 1 | | | 34 | | 1 | |
| 5 | | | | | 12 | | |
| 6 | | | | | | 60 | |
| 7 | | | | | | 11 | 1 |

(a)

| C | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| 1 | 56 | | | 4 | | | |
| 2 | 4 | 6 | | | | 2 | |
| 3 | 3 | | 39 | 6 | | | |
| 4 | 1 | | | 35 | | | |
| 5 | | | | | 4 | 8 | |
| 6 | 2 | 3 | | | | 55 | |
| 7 | 1 | | | 9 | | 2 | |

(b)

| C | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| 1 | 52 | | | 6 | | 2 | |
| 2 | | 1 | | | | 11 | |
| 3 | 1 | | 41 | 5 | | 1 | |
| 4 | 1 | | | 34 | | 1 | |
| 5 | | | | | 6 | 6 | |
| 6 | | 1 | | | | 59 | |
| 7 | 2 | | | 7 | | 2 | 1 |

(c)

| C | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| 1 | 57 | | | 3 | | | |
| 2 | | 12 | | | | | |
| 3 | 1 | | 47 | | | | |
| 4 | 1 | | | 23 | | | |
| 5 | | | | | 12 | | |
| 6 | | | | | | 60 | |
| 7 | | | | | | | 12 |

(d)

## B. Confusion Tables

We also conducted a test to see how our implementations of the GLCM's respond to different context classes through the use of confusion tables. We used eight features, 240 samples, 19 images, and seven contexts. The classification results of MDMO, ODMO, and ODOO are shown in Tables IV and V. Again, MDMO outperformed the other two with a training set classification accuracy of 97.95%. Upon closer observation, we can see that both Class 5 and 6 possess clustering characteristics with Class 5 data having smoother composition than Class 6, a difference rendered indistinguishable without cluster-type features as in our earlier experiment. Moreover, we can see that both Class 4 and 7 data have similar compositions if we concentrate on a small region of size, or use a small displacement value for the matrix. ODMO and ODOO failed to classify these data because they did not employ full range of displacement values. On the other hand, MDMO was able to coordinate all matrices of different displacement values and project each textural context onto a global scrutiny. This approach deals with microtextures and macrotextures successfully in this experiment. Using the KHAT [16] statistics (an estimate of KAPPA [15]) to assess accuracy, MDMO with ten features achieved 0.831, ODOO with eight features achieved 0.763, ODMO with eight features achieved 0.757, and MDMO with eight features achieved 0.914. Thus, even with the inclusion of cluster features and its consequent degradation to the classification, MDMO still achieves better separability among sea ice contexts than either ODOO or ODOO.

## C. Classification Power

To test the generality of our MDMO Bayes classifier, we divided the data set into two, each with 120 samples. First, we used 120 samples to train the classifier and then applied the classifier to the other 120 samples. The training set classification accuracy was 99.19%, and the test set classification accuracy was 94.17%. The overall classification accuracy was 96.67%. Once again, the KAPPA results are 0.967 and 0.889 for the training set and test set classification accuracies, respectively.

## D. Conclusions

Several conclusions have been drawn from these experiments, as follows.

1) MDMO implementation is better than the ODMO and ODOO implementation.
2) ODMO and ODOO's performances are about the same. This indicates that the orientation factor is not important in SAR sea ice imagery for textural context research.
3) Cluster shade and cluster prominence are able to highlight cluster-type textures, but, due to their significantly large range of values and thus their influence over other features, the inclusion of these two features is bad for classification.
4) Clusters for MDMO are generally more selective than those for ODMO and ODOO as it actually detects the difference between a low-contrast Web and a high-contrast Web (as shown in Table I when the classification accuracy for Class 1 samples dropped).
5) Range of displacement values is more representative than a single displacement value. This indicates that, when applying GLCM on SAR sea ice imagery, we should make use of matrices of different displacement values. In our experiments, we used the average function to combine the features of all displacement values. This indicates that MDMO is able to capture local and global details of a texture. Local details can be captured since microtextures are sometimes periodic and preservable by the choice of different displacement values; global details can be captured since macrotextures are generally correlated and coverable by the range of displacement values. This conclusion concurs with that of [48] in that both suggest that different displacement (or lag) values should be used to characterize textures.

6) Matrices with the highest $\chi^2$ values do not provide useful data, and they are not sufficiently representative for our samples in the classification process.

7) Classification based on the MDMO implementation is accurate and generalizable as it yields >90% training and test sets classification accuracies.

## VI. RELATED WORK

### A. Texture in Sea Ice

Statistical texture analysis is important in SAR sea ice imagery research since it allows better representation and segmentation of sea ice regions, compared to analysis based on intrinsic gray levels only. For example, Holmes *et al.* [38] classified one SAR image (over the Beaufort Sea) to new/first-year ice and multiyear ice with an overall accuracy of more than 65% using derived textural descriptors on X-band (HV polarization). Nystuen and Garcia [56] used standard and higher order texture statistics generated from co-occurrence matrices to classify SAR sea ice data collected during the marginal ice zone experiment (MIZEX) in April 1987. The combination of the two statistics improved ice type classification to an overall accuracy of 89.5%. Shokr [63] introduced several second-order textural parameters and evaluated them together with existing parameters by examining their usage and performance in sea ice classification for radar imagery. Sun *et al.* [65] used normalized local average and standard deviation as first-order texture parameters to classify sea ice images into open water, young ice, level ice, brash ice, and ice ridges. They concluded that the pixel-based approach, using texture, was more effective than the region-based approach in representing and characterizing sea ice types. Smith *et al.* [64] employed both spectral and textural information to classify sea ice types from ERS-1 (Earth Resource Satellite-1) SAR data. Statistical textures have also been utilized to other sea ice imagery such as Landsat Thematic Mapper (TM) Antarctic scenes [14] and NOAA AVHRR images [4].

In other domains, researchers have also used texture to provide analytical information about an image, showing that the incorporation of texture into segmentation or classification tasks is crucial. For example, Hsu [39] used first-order statistical textural features, such as mean, standard deviation, skewness, kurtosis, etc., to improve land-use mapping to 85–90%. Jensen [41] combined spectral and textural features to improve land-cover classification at the urban fringe. Shih and Schowengerdt [62] indicated that texture might be extremely valuable for classifying geologic/geomorphic surfaces, based on their experiments on bedrock, desert pavement, fluvial deposits, and vegetation from Landsat images. Also, cloud analysts have utilized textural features to improve classification of clouds [23], [28]. Ryherd and Woodcock [58] showed that addition of image texture improved the segmentation process in most areas where there were textural classes in the image.

Previous studies of SAR sea ice imagery have concentrated on the effect of texture on the quality of the analysis to describe sea ice types and deformations. In this paper, we emphasize the statistical textural contexts of SAR sea ice regions for identifying structural composition of ice-water patterns, instead of surface textures that have been used for determining ice types (for example, open water, first year, and multiyear) and degrees of deformation of sea ice (for example, undeformed and deformed first year). Instead of looking at local textures of different ice types, we are looking at a more global composition of sea ice features as a texture. Hence, we are analyzing sea ice texture at a higher granularity. We argue that this approach can *complement* the existing ice-type textural analysis since 1) local textures of different ice types are sometimes not consistently distinguishable, 2) noise effects are more significant if analyzed locally, as in sea ice surface texture analysis, 3) knowledge about sea ice types can be extracted from their relations (i.e., the composition of sea ice types), and 4) global textures are more reliable. We call this composition of sea ice features the *textural context*. Due to different geographical locations and seasonal temperatures, different ice types can coexist in different situations, creating different textural contexts. These contexts are reliable properties for image manipulation since they are more resistant to noise-corruption than surface texture. They are also of a second-order perception level of sea ice features in the image, deemphasizing local details and concentrating on the spatial, relational make up of a region. They can also be linked to sea ice phenomena associated with certain time in the year and region, providing valuable information toward the automation of sea ice image understanding. Indeed, other domains have employed the usage of context to improve classification. For example, Groom *et al.* [29] used contextual correction to improve land cover mapping. Initial classes were reexamined based on terrestrial and maritime contexts, urban and nonurban contexts, and upland and lowland contexts on TM multiband imagery. As a result, the classification accuracy was improved. Flasse and Ceccato [24] used a contextual algorithm to confirm potential fire regions in AVHRR fire detection analysis. This approach allowed the technique to be adaptive to the different environments and, thus, different radiometric imaging of the surrounding regions. In this paper, the seven textural contexts selected, as described in Section III, are by no means representative of all SAR sea ice imagery. To fully utilize textural contexts, further work must be done to study fully a vast amount of SAR sea ice imagery covering different seasons and regions.

### B. Texture Algorithms

There are several statistical texture analysis algorithms designed to represent and recognize textures; for example, GLCM [35], gray-level run length [27], gray-level difference vector [69], Fourier power spectrum [49], max–min texture [55], [62], sum and difference histograms [67], texture spectrum [68], and semivariograms [48], [54] are among the common approaches in the literature. We have chosen the GLCM as our texture analysis algorithm to investigate different SAR sea ice images for two basic reasons. First, perceptual psychology studies [42], [43] have shown this method to match a level of human perception. Second, separate studies have shown this method to outperform the others

in texture discrimination. Weszka *et al.* [69] studied the performances of the four texture analysis algorithms (the first aforementioned in this paragraph) on 54 aerial photographic terrain samples belonging to nine land use classes and extended the experiment to 180 LANDSAT imagery samples belonging to three geological terrain types. Conners and Harlow [18] conducted a comparative study on Markov-generated textures and its translation stationary random fields of order two. Both studies showed that features based on GLCM yielded better classification results than other methods. Gong *et al.* [30] showed that textural features derived from GLCM and the simple statistical transformation method could largely improve the classification accuracies, both generated superior results to the texture spectrum method, and the energy feature of GLCM was the best classification feature. The experiments were conducted on a SPOT high-resolution visible (HRV) band 3 image.

Note that some other works showed GLCM to be inferior. For example, He and Wang [37] demonstrated that, for airborne SAR $C$-band data on lithological units, features derived from texture spectrum fared better than the co-occurrence method [68]. An in-depth comparative study of five texture algorithms [2] reported that no universally best feature set, among the five different texture algorithms chosen, was found when the data from six frequency bands (of a TM image of a mountainous region) were used. However, if using only a single band for classification, features derived from the Fourier spectrum were found to be better. Dikshit [22], investigating a high-resolution (1.25–5.0 m) airborne TM image of a semi-natural scene, concluded that features based on the gray-level difference histogram and the sum and difference histogram approaches showed more promise than those of GLCM. This study is not readily generalizable to our work since SAR sea ice imagery is generated by spaceborne satellites and has lower resolutions (e.g., for ERS-1 satellite, 100 m for low-resolution images, 12.5 m for full-resolution images).

Finally, note that Cossu [20] described a study that concluded experimentally that features such as energy, contrast, variance, correlation, entropy, and inverse difference moment were most relevant among GLCM-based features to represent textures. Baraldi and Parmiggiani [4] extended the investigation further by analyzing the statistical meaning of the six GLCM features, by means of a self-consistent, theoretical assessment in order to remain independent from test images. They supported the theoretical analyses with experiments conducted on an NOAA AVHRR channel 2 image that contained mountains, sea ice, clouds, and continental ice classes. This study arrived at some similar conclusions regarding energy and entropy as textural features. We do not attempt to identify the most relevant GLCM features in our work, but we assume that all of them can be useful in representing certain type of textures.

## VII. CONCLUSIONS

In this paper, we have accomplished three objectives. First, we investigated the quantization, displacement, and orientation factors of GLCM on SAR sea ice imagery. We showed that 1) 256-level representation is not necessary, 2) eight-level representation is undesirable, and 3) 64-level representation is efficient and sufficient. We also showed that the displacement factor is more important than orientation. Second, we compared three different implementations of co-occurrence matrices. We concluded that matrices with a range of displacement values as a collective whole, due to their capabilities of capturing microtextures and macrotextures, are more adequately representative of a texture than a single $\chi^2$-optimal matrix. Also, from the training set classification accuracy of ODMO and ODOO, single-displacement features alone should not be used to describe sea ice textural contexts. Third, we presented a Bayes classifier that yielded 99.19% of training set and 94.17% of test set classification accuracies into seven classes on 240 samples. This indicates that 1) sea ice textural contexts can be successfully represented and used to classify different types of SAR sea ice imagery and 2) sea ice textural contexts can be successfully represented using GLCM-based features. Our future work in this research includes applying sea ice textural contexts to refine classification based on spectral and surface textural information.

The above findings pave the way for SAR sea ice texture analysis using GLCM in a concrete and definitive manner; the choices of quantization, displacement, or orientation values are no longer arbitrary or experimental. We have also conducted a comparative study based on Bayesian classification results on three different GLCM designs. This study also provides evidence for how a GLCM should be designed to adequately capture textures; the evidence is formally supported by statistical measurements of Bayesian classification results.

From the viewpoint of higher granularity, in this paper, we have proposed textural contexts of sea ice, rather than surface texture of sea ice types. We have also described experimental studies and compared different implementations of texture algorithms to prove that such textural contexts can be distinctly characterized. Undoubtedly, this means of successful representation motivates and facilitates further utilization of textural contexts in sea ice classification.

## REFERENCES

[1] B. P. Adhikari and D. D. Joshi, "Distance discrimination et resume exhaustif," *Publ. Inst. Stat.,* vol. 5, pp. 57–74, 1956.

[2] M. F. Augusteijn, L. E. Clemens, and K. A. Shaw, "Performance evaluation of texture measure for ground cover identification in satellite images by means of a neural network classifier," *IEEE Trans. Geosci. Remote Sensing,* vol. 33, pp. 616–626, May 1995.

[3] A. Baraldi and F. Parmiggiani, "Urban area classification by multispectral SPOT images," *IEEE Trans. Geosci. Remote Sensing,* vol. 28, pp. 674–680, May 1990.

[4] _____, "An investigation of the textural characteristics associated with gray level cooccurrence matrix statistical parameters," *IEEE Trans. Geosci. Remote Sensing,* vol. 33, pp. 293–304, Mar. 1995.

[5] D. G. Barber and E. F. LeDrew, "SAR sea ice discrimination using texture statistics: A multivariate approach," *Photogramm. Eng. Remote Sensing,* vol. 57, no. 4, pp. 385–395, 1991.

[6] M. Beauchemin, K. P. B. Thomson, and G. Edwards, "Edge detection and speckle adaptive filtering for SAR images based on a second-order textural measure," *Int. J. Remote Sensing,* vol. 17, no. 9, pp. 1751–1759, 1996.

[7] A, Bhattacharyya, "On a measure of divergence between two statistical populations defined by their probability distributions," *Bull. Calcutta Math. Soc.,* vol. 35, pp. 99–109, 1943.

[8] D. E. Boekee and J. C. A. Van Der Lubbe, "Some aspects of error bounds in feature selection," *Pattern Recognit.,* vol. 11, pp. 353–360, 1979.

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

[10] D. W. Chen, S. K. Sengupta, and R. M. Welch, "Cloud field classification based upon high spatial resolution textural features, 2. Simplified vector approaches," *J, Geophys. Res.,* vol. 94, no. D12, pp. 14 749–14 765, 1989.

[11] P. C. Chen and T. Pavlidis, "Segmentation by texture using a co-occurrence matrix and a split-and-merge algorithm," *Comput. Graph. Image Processing,* vol. 10, no. 2, pp. 172–182, 1979.

[12] H. Chernoff, "A measure of asymptotic efficiency for tests of a hypothesis based on a sum of observations," *Ann. Math. Stat.,* vol. 23, pp. 493–507, 1952.

[13] Y. P. Chien and K.-S. Fu, "Recognition of X-ray picture patterns," *IEEE Trans. Syst., Man, Cybern.,* vol. SMC-4, pp. 145–156, Feb. 1974.

[14] J. Chou, R. C. Weger, J. M. Ligtenberg, K.-S. Kuo, R. M. Welch, and P. Breeden, "Segmentation of polar scenes using multi-spectral texture measures and morphological filtering," *Int. J. Remote Sensing,* vol. 15, no. 5, pp. 1019–1036, 1994.

[15] J. Cohen, "A coefficient of agreement for nominal scales," *Educ. Psychol. Measur.,* vol. 20, no. 1, pp. 37–46, 1960.

[16] R. G. Congalton, "A review of assessing the accuracy of classifications of remotely sensed data," *Remote Sens. Environ.,* vol. 37, no. 1, pp. 35–46, 1991.

[17] R. W. Conners and C. A. Harlow, "Equal probability quantizing and texture analysis of radiographic images," *Comput. Graph. Image Processing,* vol. 8, pp. 447–463, 1978.

[18] ———, "A theoretical comparison of texture algorithms," *IEEE Trans. Pattern Anal. Machine Intell.,* vol. PAMI-2, pp. 204–222, Mar. 1980.

[19] R. W. Conners, M. M. Trivedi, and C. A. Harlow, "Segmentation of a high-resolution urban scene using texture operators," *Comput. Vision, Graph., Image Processing,* vol. 25, pp. 273–310, 1984.

[20] R. Cossu, "Segmentation by means of textural analysis," *Pixel,* vol. 1, no. 2, pp. 21–24, 1988.

[21] L. S. Davis, S. A. Johns, and J. K. Aggarwal, "Texture analysis using generalized co-occurrence matrices," *IEEE Trans. Pattern Anal. Machine Intell.,* vol. PAMI-1, pp. 251–259, Mar. 1979.

[22] O. Dikshit, "Textural classification for ecological research using ATM images," *Int. J. Remote Sensing,* vol. 17, no. 5, pp. 887–915, 1996.

[23] E. Ebert, "A pattern recognition technique for distinguishing surface and cloud types in the polar regions," *J. Clim. Appl. Meteorol.,* vol. 26, pp. 1412–1427, 1987.

[24] S. P. Flasse and P. Ceccato, "A contextual algorithm for AVHRR fire detection," *Int. J. Remote Sensing,* vol. 17, no. 2, pp. 419–424, 1996.

[25] S. E. Franklin and D. R. Peddle, "Spectral texture for improved class discrimination in complex terrain," *Int. J. Remote Sensing,* vol. 10, pp. 1437–1443, 1989.

[26] ———, "Classification of SPOT HRV imagery and texture features," *Int. J. Remote Sensing,* vol. 11, no. 3, pp. 551–556, 1990.

[27] M. M. Galloway, "Texture analysis using gray level run lengths," *Comput. Graph. Image Processing,* vol. 4, pp. 172–179, 1975.

[28] L. Garand, "Automated recognition of oceanic cloud patterns, I. Methodology and application to cloud climatology," *J. Climate,* vol. 1, pp. 20–39, 1988.

[29] G. B. Groom, R. M. Fuller, and A. R. Jones, "Contextual correction: Techniques for improving land cover mapping from remotely sensed images," *Int. J. Remote Sensing,* vol. 17, no. 1, pp. 69–89, 1996.

[30] P. Gong, J. D. 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, no. 2, pp. 137–151, 1992.

[31] C. C. Gotlieb and H. E. Kreyszig, "Texture descriptors based on co-occurrence matrices," *Comput. Vision, Graph. Image Processing,* vol. 51, pp. 70–86, 1990.

[32] J. F. Haddon and J. F. Boyce, "Image segmentation by unifying region and boundary information," *IEEE Trans. Pattern Anal. Machine Intell.,* vol. 12, pp. 929–948, Oct. 1990.

[33] ———, "Co-occurrence matrices for image analysis," *Electron. Commun. Eng. J.,* vol. 4, pp. 71–83, 1993.

[34] R. M. Haralick, "Statistical and structural approaches to texture," *Proc. IEEE,* vol. 67, pp. 786–804, May 1979.

[35] R. M. Haralick, K. Shanmugan, and I. H. Dinstein, "Textural features for image classification," *IEEE Trans. Syst., Man, Cybern.,* vol. SMC-3, pp. 610–621, May 1973.

[36] D. Haverkamp, L.-K. Soh, and C. Tsatsoulis, "A comprehensive, automated approach to determining sea ice thickness from SAR data," *IEEE Trans. Geosci. Remote Sensing,* vol. 33, pp. 46–57, Jan. 1995.

[37] D.-C. He and L. Wang, "Recognition of lithological units in airborne SAR images using new texture features," *Int. J. Remote Sensing,* vol. 11, no. 2, pp. 2337–2344, 1990.

[38] 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.

[39] S.-Y. Hsu, "Texture-tone analysis for automated land-use mapping," *Photogramm. Eng. Remote Sensing,* vol. 44, no. 11, pp. 1393–1404, 1978.

[40] H. Jeffreys, "An invariant form for the prior probability in estimation problems," *Proc. R. Stat. Soc. A,* vol. 186, pp. 453–461, 1946.

[41] J. R. Jensen, "Spectral and textural features to classify elusive land cover at the urban fringe," *Prof. Geographer,* vol. 31, no. 4, pp. 400–409, 1979.

[42] B. Julesz, "Visual pattern discrimination," *IRE Trans. Inform. Theory,* vol. 8, pp. 84–92, 1962.

[43] B. Julesz, E. N. Gilbert, L. A. Shepp, and H. L. Frisch, "Inability of humans to discriminate between visual textures that agree in second-order statistics—Revisited," *Perception,* vol. 2, pp. 391–405, 1973.

[44] V. Kovalev and M. Petrou, "Multidimensional co-occurrence matrices for object recognition and matching," *Graph. Models Image Processing,* vol. 58, no. 3, pp. 187–197, 1996.

[45] R. P. Kruger, W. B. Thompson, and A. F. Turner, "Computer diagnosis of pneumoconiosis," *IEEE Trans. Syst., Man, Cybern.,* vol. SMC-4, pp. 40–49, Jan. 1974.

[46] S. Kullback, *Information Theory and Statistics.* New York: Wiley, 1959.

[47] S. P. S. Kushwaha, S. Kuntz, and G. Oesten, "Applications of image texture in forest classification," *Int. J. Remote Sensing,* vol. 15, no. 11, pp. 2273–2284, 1994.

[48] R. M. Lark, "Geostatistical description of texture on an aerial photograph for discriminating classes of land cover," *Int. J. Remote Sensing,* vol. 17, no. 11, pp. 2115–2133, 1996.

[49] G. O. Lendaris and G. L. Stanley, "Diffraction pattern sampling for automatic pattern recognition," *Proc. IEEE,* vol. 58, pp. 198–216, Feb. 1970.

[50] T. Lissack and K. S. Fu, "Error estimation in pattern recognition via L-distance between posterior density functions," *IEEE Trans. Inform. Theory,* vol. 22, pp. 34–45, 1976.

[51] P. C. Mahalanobis, "On the generalized distance in statistics," *Proc. Nat. Inst. Sci.,* India,vol. 12, pp. 49–55, 1936.

[52] K. Matusita, "Decision rules based on the distance for problems of fit, two samples and estimation," *Ann. Math. Stat.,* vol. 26, pp. 631–640, 1955.

[53] D. J. Marceau, P. J. Howarth, J.-M. M. Dubois, and D. J. Gratton, "Evaluation of the gray-level co-occurrence matrix method for land-cover classification using SPOT imagery," *IEEE Trans. Geosci. Remote Sensing,* vol. 28, pp. 513–519, July 1990.

[54] F. P. Miranda, J. A. MacDonald, and J. R. Carr, "Application of the semivariogram textural classifier (STC) for vegetation discrimination using SIR-B data of Borneo," *Int. J. Remote Sensing,* vol. 13, pp. 2349–2354, 1992.

[55] O. R. Mitchell, C. R. Myers, and W. Boyne, "A max–min measure for image texture analysis," *IEEE Trans. Comput.,* vol. C-26, pp. 408–414, Apr. 1977.

[56] J. A. Nystuen and F. W. Garcia, Jr., "Sea ice classification using SAR backscatter statistics," *IEEE Trans. Geosci. Remote Sensing,* vol. 30, pp. 502–509, May 1992.

[57] E. A. Patrick and F. P. Fisher, "Nonparametric feature selection," *IEEE Trans. Inform. Theory,* vol. IT-15, pp. 577–584, 1969.

[58] S. Ryherd and C. Woodcock, "Combining spectral and texture data in the segmentation of remotely sensed images," *Photogramm. Eng. Remote Sensing,* vol. 62, no. 2, pp. 181–194, 1996.

[59] E. Sali and H. Wolfson, "Texture classification in aerial photographs and satellite data," *Int. J. Remote Sensing,* vol. 13, no. 18, pp. 3395–3408, 1992.

[60] K. Shanmugan and R. M. Haralick, "Computer classification of reservoir sandstones," *IEEE Trans. Geosci. Electron.,* vol. GE-11, pp. 171–177, Jan. 1973.

[61] K. Shanmugan, V. Narayanan, V. S. Frost, J. A. Stiles, and J. C. Holtzman, "Textural features for radar image analysis," *IEEE Trans. Geosci. Remote Sensing,* vol. GE-19, pp. 153–156, May 1981.

[62] E. H. H. Shih and R. A. Schowengerdt, "Classification of arid geomorphic surfaces using LANDSAT spectral and textural features," *Photogramm. Eng. Remote Sensing,* vol. 49, no. 3, pp. 337–347, 1983.

[63] M. E. Shokr, "Evaluation of second-order textural parameters for sea ice classification in radar images," *J. Geophys. Res.,* vol. 96, no. C6, pp. 10 625–10 640, 1991.

[64] D. M. Smith, E. C. Barett, and J. C. Scott, "Sea ice type classification from ERS-1 SAR data based on gray level and texture information," *Polar Record,* vol. 31, pp. 135–146, 1995.

[65] Y. Sun, A. Carlstrom, and J. Askne, "SAR image classification of ice in the Gulf of Bothnia," *Int. J. Remote Sensing,* vol. 13, no. 13, pp. 2489–2514, 1992.

[66] M. M. Trivedi, C. A. Harlow, R. W. Conners, and S. Goh, "Object detection based on gray level co-occurrence," *Comput. Vision, Graph., Image Processing,* vol. 28, pp. 199–219, 1984.

[67] M. Unser, "Sum and difference histograms for texture classification," *IEEE Trans. Pattern Anal. Machine Intell.,* vol. PAMI-8, pp. 118–125, Jan. 1986.

[68] L. Wang and D.-C. He, "A new statistical approach for texture analysis," *Photogramm. Eng. Remote Sensing,* vol. 56, pp. 61–66, 1990.

[69] J. S. Weszka, C. R. Dyer, and A. Rosenfeld, "A comparative study of texture measures for terrain classification," *IEEE Trans. Syst., Man, Cybern.,* vol. SMC-6, pp. 269–285, Feb. 1976.

[70] S. W. Zucker and D. Terzopoulus, "Finding structure in co-occurrence matrices for texture analysis," *Comput. Graph. Image Processing,* vol. 12, pp. 286–308, 1980.

**Costas Tsatsoulis** (M'88–SM'98) received the Ph.D. degree in electrical engineering in 1987 from Purdue University, West Lafayette, IN.

He is an Associate Professor in the Department of Electrical Engineering and Computer Science, University of Kansas, Lawrence. His research interests include artificial intelligence, image processing, computer vision, data mining, and multiagent systems. He recently co-edited a book entitled *Analysis of SAR Data of the Polar Oceans* (Berlin, Germany: Springer-Verlag).

Dr. Tsatsoulis is a member of the ACM, AAAI, Eta Kappa Nu, and Sigma Xi.

**Leen-Kiat Soh** (S'91–M'98) was born in Johor, Malaysia. He received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of Kansas, Lawrence, in 1991, 1993, and 1998, respectively.

He is a Research Scientist in the Intelligent Systems and Information Management Laboratory, University of Kansas. His research interests include image processing, computer vision, machine learning, and knowledge discovery in databases.

Dr. Soh is a member of the ACM, AAAI, Phi Kappa Phi, Tau Beta Pi, Eta Kappa Nu, and Phi Beta Delta.