Texture Segmentation of SAR Sea Ice Imagery

by

David Anthony Clausi

A thesis
presented to the University of Waterloo
in fulfilment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Systems Design Engineering

Waterloo, Ontario, Canada, 1996

I hereby declare that I am the sole author of this thesis.

I authorize the University of Waterloo to lend this thesis to other institutions or individuals for the purpose of scholarly research.

I further authorize the University of Waterloo to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

The University of Waterloo requires the signatures of all persons using or photocopying this thesis. Please sign below, and give address and date.

Abstract

Texture Discrimination of SAR Sea Ice Imagery

The differentiation of textures is a critical aspect of SAR sea ice image segmentation. Provision of images that identify pertinent ice types is important for the operational (ice breakers, ships, oil platforms) and scientific (ie. global warming monitoring) communities. Although a human is readily able to visually segment any textured image, no unsupervised machine method has been designed that consistently and robustly performs the same task.

Two steps are followed to perform image texture segmentation. First, feature vectors representing local characteristics are determined for each pixel. The more distinct the features, the better they are able to distinguish different classes. Second, the feature vectors are grouped together according to class similarities or dissimilarities. During this clustering, human intervention should be minimized. Optimizing each of these steps is important for achieving the overall task.

Many different approaches have been proposed for texture feature extraction. Three popular methods are investigated and optimized for their supervised classification ability: cooccurrence probabilities, power spectrum, and Gabor filtering. This information is necessary to serve as a platform for image segmentation testing. The methods perform well under supervised classification experiments. The Gabor filters capture the same beneficial information as the power spectrum features, however, the Gabor filters have the ability to capture information with an effective window size that matches the desired frequency (multi-resolutional ability).

There exists only limited published research comparing different texture segmentation approaches. Cooccurrence features, although extensively used in classification studies, are rarely applied to full image segmentation. This is probably due

to the exceptional computational demands of the traditional approach (the grey level cooccurrence matrix or GLCM). To calculate cooccurrence texture features orders of magnitude faster than the GLCM, a linked list algorithm is designed and implemented.

Two clustering approaches are performed: one that assumes the number of classes *a priori* (mixture analysis) and one that automatically determines the number of classes (cluster analysis). Mixture analysis is performed by applying K-means (with the correct number of classes) and then iteratively applying the Fisher linear discriminant to improve the clustering (K-means Iterative Fisher or KIF). Cluster analysis is performed using the KIF in a binary divisive hierarchical tree where nodes branch or stop as a function of the cluster separability determined by the Fisher criteria. For operational purposes, the user indicates when the nodes branch or stop based on visual cues. The idea is to provide an easier method to segment the image to reduce operator fatigue and increase throughput.

When cooccurrence and Gabor features are applied to unsupervised image segmentation of Brodatz and SAR sea ice imagery they are demonstrated to have different abilities. Images containing regions with different textural resolutions are not segmented properly using cooccurrence features. Gabor filters, with inherent multi-resolutional ability, are able to perform this task. When all textures have the same resolution, both methods are able to distinguish the textures. While other issues such as tone and ice floe shapes must be included for a robust operational SAR sea ice identification system, incorporation of texture is a necessary component for successful image segmentation.

Acknowledgements

Supervisors are the guiding light of the grad student. For my thesis work, I had the enviable position of dual supervision, each supervisor tending to complement the other's background. Ed Jernigan's technical engineering knowledge coupled with the practical information provided by Ellsworth LeDrew provided a strong foundation for my thesis work. Their guidance and understanding helped to elevate and promote my work to a level that would not have been otherwise possible. Their assistance and inspiration are gratefully appreciated.

I would like to express my gratitude to the administrative assistants who generously helped me out. Thanks to Sue Gooding, Annette Dietrich, Faye Shultz, and Susie Bartlett.

Fellow grad students play an important role in my life, both socially and professionally. A list of all the UW grad students with whom I have interacted is quite lengthy, so I will just extend one big thanks to them all. Special acknowledgements are extended to my VIP and geography colleagues.

Funding is a crucial aspect of grad studies. Primary funding has been provided by the Natural Sciences and Engineering Research Council (Post Graduate Scholarship) and the Faculty of Engineering (Carl Pollock Graduate Scholarship). Funding for conferences has been provided by the Dean of Engineering, Graduate Student Office, Department of Systems Design, Ed, and the Sandford Fleming Foundation. Thanks to all the funding providers.

I thank my parents for the continous support of my endeavours. A bit of their teaching is always reflected in my work and I would never have reached so far without their encouragement.

Finally, special thanks to Mira. Her patience, love, and understanding were instrumental for successful completion of my dissertation.

Dedicated to the most
important people in my life:
Mira, Mom, and Dad.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 General

Interpretation of satellite radar images is an important ongoing research field. Scientists' efforts are directed towards improving calibration of radar platforms, participating in field validation programs to correlate ground measurements with the remotely sensed data, predictive modelling of the radar backscattering information, and developing automated algorithms for the identification of pertinent features observed in the satellite imagery. This thesis will focus on the development of texture segmentation strategies for the purpose of identifying different sea ice classes in SAR imagery.

Texture is an important part of image interpretation. As early as 1981, Shanmugan recognized the importance of texture for the interpretation of natural radar imagery [97]. Many different texture interpretation approaches have been considered, but a complete formal definition for texture has still not been developed. Texture measures are required that indicate some specific characteristic of the texture itself,

that is, a measurement is needed that captures some particular important relationship between a pixel and its neighbours. Texture interpretation is important for classification problems where individual image samples must be determined to be one class or another. Texture interpretation is also important for the segmentation problem where a single image contains multiple textures and the goal is not only to find the boundaries between the textures, but also to associate individual textures with a particular class.

The texture segmentation process can be broken down into two steps. First, suitable features (or characteristics) of each of the pixels must be calculated. This is referred to as feature extraction. Second, the pixels must be grouped together as a function of their features. This is the clustering aspect. Improving the quality of the features will ease the burden on the clustering scheme. A robust clustering methodology that successful identifies the pertinent clusters in the feature space is desirable.

1.2 Thesis Objectives

The following are the objectives of this thesis:

- Present an understanding of the relationship of the geophysical environment to the backscattering imagery obtained using synthetic aperture radar (SAR).
- Review research publications that have utilized cooccurrence texture features for classification of SAR sea ice imagery.
- Present a linked list implementation for determining cooccurrence probabilities. This method is considerably faster than the grey level cooccurrence matrix (GLCM) approach.

- Investigate the feasibility of using a multi-channel filtering approach for discriminating SAR sea ice imagery.
- Based on classification ability, determine the preferred parameters for three different texture feature extraction schemes: cooccurrence probabilities, power spectrum, and Gabor filtering.
- Apply the most promising feature extraction approaches to the more difficult problem of texture segmentation.
  - Design, implement, and demonstrate an unsupervised texture segmentation approach.
  - Design, implement, and demonstrate a texture segmentation approach that can be used for operational SAR sea ice segmentation.

1.3 Organization of Thesis

Following the background chapter (Chapter 2), an entire chapter (Chapter 3) is dedicated to presenting the application environment. Here, the SAR system parameters that influence the backscattering characteristics of the surface under analysis and the relationship of the backscattering as a function of the snow volume/ice surface are discussed.

In Chapter 4, various methods that have been used to generate texture features as well as techniques used to classify the feature vectors are investigated. In order to determine preferred methods for texture segmentation, classification studies are performed (Chapter 5). These preferred parameters are used to assist the design

process in Chapter 6 to perform texture segmentation. A summary and discussion of future research activities comprise the final chapter (Chapter 7).

Chapter 2

Background

Monitoring of sea ice is a major challenge to both scientific and operational communities. Many scientists are ultimately interested in issues dealing with global warming. According to Barber and LeDrew the sea ice acts as the Arctic's thermal regulator. "Radiative and energy balances are intimately linked with the ice cover as is the development and sustenance of ecosystems, ...ecological character and diversity of open water areas, particulary those that occur within the perennial ice cover, may act as sensitive indicators of climate variability and ecological change in the Arctic. [5, p.4]" That the difference between energy flux from open water and the surrounding ice may be two orders of magnitude is evidence that the ocean-ice-atmosphere relationship contains a wide range of sensitive energy balances [71]. Also, during the winter season, the Arctic sea ice covers twice the area as compared to the summer season [7]. Little is known about the global energy transfers required during such a transition, thus, research is conducted to determine the capability of remote sensing to measure pertinent variables related to the sea ice. Also, from an operational standpoint, ice poses severe hazards to ship and ice breaker movement

and offshore drilling activity [55, 85]. Knowledge of ice types, movement, and concentrations allow such activities to proceed safely and efficiently [106]. "Industrial development within ice-infested waters is premised on timely ice information for tactical navigation and hazard avoidance [7, p.46]."

The scientific study of the ice covered regions is difficult and challenging for several reasons. For example, the Arctic is a vast and isolated environment, making ground monitoring an intensive task. Local weather analyses do not necessarily represent the weather patterns over a larger region. Sea ice information should be recent (less than 24 hours old) to inform ship and ice breaker movements, especially those in the St. Lawrence seaway. A high resolution all-weather day or night remote sensing system that operates in real time is a necessity. Synthetic aperture radar (SAR) is capable of meeting these requirements. The only electromagnetic radiation capable of day, night, and all-weather monitoring is microwave sensing since selected microwave frequencies are capable of penetrating the atmosphere under virtually all conditions [37]. Remotely sensed aerial or satellite platforms regularly gather information over considerable land area at a sufficiently high resolution (on the order of metres or tens of metres).

That SAR provides critical information is emphasized by the number of SAR carrying satellites that have recently been launched: ERS-1, ERS-2, (European), JERS-1 (Japanese), and Radarsat (Canadian). Radarsat has been designed specifically with sea ice monitoring in mind: "In Canada, the Atmospheric Environment Services Ice Branch, responsible for the provision of ice information for all ice-infested waters in Canada, will be the single largest user of Radarsat imagery in the world, at least in the initial years of the satellite. [57, p.170]" The estimated required annual number of images is about 4000, 500km by 500km scenes. Since exceptional volumes of data will be transmitted from aerial or satellite SAR oper-

ations, rapid and accurate interpretation of such data is critical.

Typically, a human operator is clearly and rapidly able to visually distinguish different ice types and land in a SAR sea ice image, based not only on relative texture appearance of the different ice types, but on tonal and structural differences as well. Carsey [17, p.134] agrees since he states: "Classification of SAR images is difficult to do in the computational environment, even if it appears straightforward on sight." Developing a robust computer algorithm that is able to perform such a task has proven to be elusive. The ultimate goal is to develop an algorithm that has the ability to segment a SAR sea ice image into its salient ice categories, given a limited number of parameters. "The challenge is to develop computer algorithms that can consistently discriminate among the various ice types and conditions that are visually evident in the SAR imagery [3, p.112]."

A typical pattern recognition methodology (measurement, feature extraction, feature reduction, and class assignment) can be used to solve the sea ice recognition problem. The same approach is used to study classical pattern recognition problems such as character recognition and medical diagnosis. For the particular problem at hand, measurement refers to the acquisition of the SAR sea ice image. Parameters associated with this process are reviewed in Chapter 3. Feature extraction is the determination of unique and identifiable characteristics of the classes the user wishes to discriminate. For this research, the classes are the different categories found in the SAR sea ice imagery. Feature reduction reduces the dimensionality of the feature space to remove redundancy and accelerate the algorithmic speed. Since the number of features may be many and because there may exist inter-feature redundancy, feature reduction is important. Class assignment is a process that groups samples with similar features. Successful assignment assumes that the classes under consideration are fairly homogeneous and separable from

the other classes. Defining what constitutes a homogeneous region is problematic when dealing with natural imagery since such imagery is typically nonstationary. A recurring problem is finding a universal approach for all SAR sea ice data sets since discriminants derived from one data set are not always applicable to other data sets. A final note about the pattern recognition process is that expert knowledge is often required. In terms of sea ice classification, knowledge of the time of year, historical and prevailing weather conditions, knowledge of the geographical region, the size and shape of the ice floes, etc. can be of importance to assist the segmentation process.

Few numerical methods are used in practice to classify SAR sea ice imagery. The methodology and results of generating texture features using cooccurrence probabilities [52] have been clearly outlined in the research literature [4, 55, 98]. Kwok *et al.* have recently implemented a classification scheme at the Alaskan SAR Facility using only first order analysis. They recognize the limitations of this approach, given that other algorithms that include higher order information have performed better, however, "there are insufficient observations at this time to support the implementation of such feature extraction schemes in an operational system. [69, p.2392]" There are indications that, at a minimum, second-order information is required to properly classify SAR sea ice imagery. According to Barber *et al.* "texture measures consistently provided more robust classifications than features that encompassed tone... [3, p.117]". Ulaby *et al.* also obtained results that indicate second-order measures outperform tonal measures [103]. Texture identification is an important contribution to interpretation of SAR sea ice imagery.

To robustly perform SAR sea ice identification, additional texture feature extraction methods should be considered. Texture features derived from the power spectrum have potential to classify SAR sea ice imagery. Success using such meth-

ods for natural textures by D'Astous [27] and Liu and Jernigan [74] motivates consideration of this classification approach. The engineering research literature contains a host of potential segmentation schemes used in computer vision applications. Various methods are described in Chapter 4. Multi-channel filtering is a method that can identify characteristic features of an image by isolating pertinent information such as orientation and dominant frequency components. Based on wavelet theory and models of the human visual system (HVS), these multi-channel approaches show considerable promise for segmenting SAR sea ice imagery. Gabor functions, which can be implemented in a wavelet fashion, have several properties that make them quite attractive for the generation of texture features. The goal is to provide a feature set that contains separable clusters in the multidimensional feature space to ensure accurate texture segmentation.

Machine texture identification is an easily defined problem: identify different textures and classify them as well as the human observer. Considerable research to improve texture interpretation is found in the literature (Chapter 4; however, examples of specific real world applications are few. Identification of pertinent features in remotely sensed imagery is a well defined problem. In general image analysis, there are many different considerations to analyzing an image: tonal variations, object analysis, depth perception, noise removal, motion detection, etc. Texture is only one aspect of the general framework of computer vision. Scientists are still trying to grasp the basic framework of texture interpretation before connecting this understanding to other interpretation needs of machine assisted vision problems. As the integrating of these different needs progresses, being able to distinguish textures will be an essential aspect. For the time being, computer vision methodologies consider investigation of the entire scene for tonal or textural variations. Hence, textural investigation of SAR sea ice imagery is a practical and necessary step to

improving our understanding of image interpretation. Real world applications tend to be highly complex for machine implementation. Generic methodologies, that can be applied to a wide variety of circumstances, are essential.

Chapter 3

Application Environment

There are many different types of sensors used for remote sensing, each with their own capabilities. Interpretation of the data obtained from these sensors is an ongoing research field. My research will be directed to the interpretation of SAR imagery, primarily for the reasons outlined in the previous chapter. Note that the synergistic effect of a multitude of sensors can produce more productive data sets, but the potential of each sensor must first be fully explored. Also, although there are many different objectives for remote sensing activities, here the primary interest is the ability to identify different sea ice types using active microwave sensors.

This chapter contains two sections. The first section deals with system parameters, namely, those variables that pertain directly to operation of the SAR platform. The second section deals with the interaction of the microwaves with the sea ice surface.

3.1 SAR System Parameters

Radar is an acronym for *radio detection and ranging*. Real aperture radar (or brute force radar) systems have inherent difficulties producing high resolution data. Resolution in the azimuth direction (flight direction) is dependent on the antenna length. When attempting to achieve a high resolution, the antenna must have an unacceptably long physical length (the beamwidth of the radar beam is inversely proportional to the antenna length [37, 72]). SAR alleviates this problem by synthesizing the effect of a very long antenna using the Doppler history created by the relative velocity between the platform and the ground. Resolution in the range direction (direction of signal propagation) is dependent on the pulse length of the signal. Both real and synthetic aperture radars are called active sensors because they emit energy pulses and, based on the magnitude of the backscattering response to the antenna, interpret the surface characteristics. Passive sensors measure energy emitted or reflected from the surface of interest.

Many parameters influence the backscatter coefficient: wavelength of incident signal, incident angle, number of looks, polarization, and look direction. Whether a surface is considered smooth or rough can be indicated by the Rayleigh criterion:

$$h < \frac{\lambda}{8cos\theta}$$

where h is the root mean square (rms) height of the surface, λ is the signal wavelength, and θ is the angle of incidence with the surface. If the Rayleigh criterion holds true, then the surface is considered smooth and will reflect most of the signal, resulting in a low backscatter return. If the Rayleigh criterion is false, the surface is considered "rough" and scatters incident energy in all directions. Note that an increasing range distance causes the incidence angle to increase which modifies

the backscatter behaviour within a homogeneous scene. Applying an automated texture segmentation algorithm to SAR aerial imagery can be awkward because there is a large change in incidence angle across the image swath. A larger range of incidence angles leads to a perceived increase in variability of ice classes. Satellite SAR platforms do not exhibit this problem to the same extent because their significantly higher flight path altitude results in smaller changes in incidence angles over the same effective area. Thus, interpretation of the satellite data tends to be more consistent.

Wavelengths used to capture sea ice imagery are usually X-band (12.5-8.0GHz, 2.4-3.8cm) or C-band (8.0-4.0GHz, 3.8-7.5cm). Shorter wavelengths are not useful since they are significantly attenuated by the atmosphere. Longer wavelengths are advantageous because they have higher penetration depths into the ground. However, wavelengths such as L-band (2.0-1.0GHz, 15-30cm) are too long to discriminate the small scale surface roughness found on first year ice types impenetrable to radar. Thus, L-band is better suited to open ocean monitoring.

An important error produced by SAR is speckle, which is a grainy, salt-and-pepper appearance on the image. "Speckle arises from the coherent nature of radar waves, causing random constructive and destructive interference and, hence, random bright and dark areas in radar imagery [72, p.499]." Multiple look processing reduces the amount of speckle by averaging images of the same region. The number of images averaged is known as the number of looks. The amount of speckle is inversely proportional to the square root of the number of looks and the resolution cell size is directly proportional to the number of looks [72].

Another important SAR system parameter is polarization of the transmitted and received signals. Typically, the signal is transmitted/received in either the horizontal (H) or vertical (V) planes. Thus, four different designations are created:

HH, VV, HV, VH - the first letter refers to signal transmission and the second letter refers to signal reception. Radiation from natural sources, such as the sun, typically do not have any well defined polarization (random polarization). Different polarizations yield different types of backscatter returns.

3.2 SAR Relationship with Environmental Variables

3.2.1 Physical Properties of Sea Ice

Sea ice is a complex, dynamic, and anisotropic material that, under typical conditions, is a three phase medium. The solid phase consists of the ice crystals, the liquid phase comprises a brine solution, and air bubbles represent the gas phase. The initial composition of the sea ice is determined by many variables including air temperature, wind conditions, ocean currents, salinity, and rate of freezing.

The backscatter obtained using SAR may not necessarily describe the surface topography of the snow or ice. Note that microwave reflections bear no relationship to visible or thermal portions of the spectrum. Depending on the properties of the measured surface, the signal may be reflected, scattered, transmitted then reflected by a lower surface, transmitted then scattered by a lower surface, or attenuated. A dominant factor influencing backscatter is the dielectric constant of the material. For example, water is known to have a high dielectric constant. The presence of moisture in the volume under scrutiny can significantly increase radar reflectivity, distorting surface roughness measurements. Note that there is always some attenuation of the SAR signal by the atmosphere which is a function of the

New Ice: Recently formed ice which includes frazil ice, grease ice, slush, nilas, and shuga. These ice types consist of ice crystals which are weakly frozen together.

Young Ice: Ice in the transition stage between nilas and first year ice. Ice thickness generally ranges between 10-30 centimetres.

First Year Ice: Sea ice, developed from young ice, which has no more than one winter's growth. Thickness ranges from 30 centimetres up to 2 metres.

Old Ice: Ice which has survived at least one summer melt. Often sub-divided into either second year ice or multi-year ice.

Figure 3.1: World Meteorological Organization (WMO) sea ice categories (adapted from McKillop [81]).

cloud cover and current precipitation. Although this is not of serious consequence to the texture obtained in SAR imagery, it does have a direct bearing on tone.

The World Meteorological Organization (WMO) has a listing of sea ice categories summarized in Figure 3.1 (adapted from McKillop [81]). Researchers have a good understanding of the differences in backscattering characteristics of first year smooth, first year rough, and multi-year ice types [7, 81]. This information is important primarily because the ice type gives a strong indication of its total thickness, which is required for navigation purposes or climatological studies. Land identification is also of importance, however, land typically does not have a consistent, regular appearance within a SAR sea ice image making automated recognition difficult. This task can be performed by registration with existing land maps.

At the Ice Branch of Environment Canada (Ottawa, Ontario) individuals are hired specifically for the task of interpreting remotely sensed imagery of sea ice. These analysts record information in a consistent manner using a set procedure. Global attributes of the sea ice characteristics that are recorded include: total concentration (reported in tenths), partial concentrations of observed ice types, stage of developments of ice types (indicates ice thickness), and the predominant formation of the ice.

During the winter season, the backscattering behaviour of the snow and sea ice is fairly constant. At this time, the snow cover is essentially transparent at microwavelengths causing the backscatter to be dependent on the ice surface. Since first year ice has a high saline content, radar frequencies are easily reflected and surface scattering dominates [98]. Thus, first year smooth ice will reflect the radar signal away from the antenna. First year rough ice causes considerable and unpredictable scattering. Since multi-year ice has a much lower saline content (typically zero), backscatter return is influenced by both surface and volume scattering within the multi-year ice volume. The backscatter return of the multi-year ice is influenced by air bubble size and spacing, ice crystal grain size, and micro and macro-scale roughnesses. Whether or not new ice can be separated from first year smooth ice is uncertain, however, "New ice in leads has a sufficiently high proportion of well-oriented features to distinguish it uniquely from first year and multi-year ice [38, p.233]." New ice is not characterized by a discrete emissivity or narrow range of brightness temperatures, but often displays banded structures oriented parallel to the lead axis or polynya boundary. First year and multi-year ice are subject to more random growth than new ice.

3.2.2 Transitory Backscattering Characteristics

A fundamental difficulty with SAR sea ice image machine interpretation is the diurnal and seasonal variations that modify backscatter signatures. These variations are typically related to solar radiation heating of the surface causing fluctuations of the dielectric constant of the snow and ice due to changes in free water content and brine concentrations. Excessive free water content within or on the ice [75] causes a significant increase in magnitude of the dielectric constant which tends to mask backscatter signatures of first year versus multi-year ice types. Although this disrupts the ice class discrimination, it can also assist determination of the onset of melt. Periodic warming leads to recrystallization and air pocket creation which also affect backscatter return. The migration of brine out of the first year ice upward onto the ice surface and into the snowcover is another critical issue. This causes the dielectric constant of the snow to increase leading to volume scattering within the snow pack. Thus, first year smooth ice begins to appear rough in SAR imagery due to the volume scattering by the snow. Such seasonal transitions have a severe effect on backscattering characteristics of the cryosphere and are difficult to monitor. For the above reasons, Kwok *et al.* decided to "implement a classification procedure for seasons when the surface conditions remain fairly stable and a high confidence can be placed in the classification results [69, p.2392]."

Another process that leads to confusion for interpretation of SAR sea ice imagery is the occasional intrusion of sea water into multi-year floes (possibly due to flooding under high snow loading). The high brine content of the sea water tends to misrepresent the multi-year flow that has a very low brine content. It is difficult to monitor such a situation other than through tracking of floes in SAR images.

Another situation that varies backscatter characteristics is the incidence angle

of the sensor. Higher incidence angles are known to emphasize variations in surface roughness. For example, aerial imagery typically has high incidence angles and first year smooth ice and first year rough ice are quite discriminable. However, the surface variation (from the sensor's perspective) of the multi-year and first year rough types are similar and discrimination of these two ice types is more difficult. A sensor such as ERS-1 with a small incidence angle (23 degrees) discriminates between multi-year ice and first year ice types but has a difficult time discriminating between first year rough and first year smooth ice. The smaller incidence angle allows the signal to penetrate the multi-year ice more causing a significantly scattered response yet the first year ice types tend to reflect the signal away from the antenna creating a similar appearance.

3.2.3 The Role of SAR in Geophysical Interpretation

There are many research questions that SAR can help answer. As in most sciences, finding answers to these questions only seems to raise even more questions. Many applications require daily coverage and typically weekly coverage as a bare minimum [7]. Ice thickness and open water coverage are primary scientific interests and there is no well-established technique for their absolute determination [17]. Monitoring of ice edge ocean eddies using remotely sensed images may assist interpretation of their effects on the sea ice environment. Currently, the net significance of ice eddies is not fully understood. Dynamic thermodynamic modelling of sea ice cover is still an active research field. SAR can also be used to track ice motion. There are several reasons for this research interest: monitoring latent heat advection, oceanic surface stress, passive tracing of currents, open water production by way of ice divergence and shear, and interaction between ice breaker navigation and ice structure [68, 106]. There continues to be active ongoing research into theoret-

ical modelling, for example, trying to predict complex scattering processes within the ice sheet and overlying snow.

Much of the above research work requires field programs to provide validation studies. The Seasonal Sea Ice and Monitoring and Modelling Site (SIMMS) was such a program operated by the Earth Observation Laboratory (EOL) of the Institute for Space and Terrestrial Science (ISTS/University of Waterloo) in cooperation with many participating agencies. SIMMS was an annual multi-disciplinary field experiment that involved numerous different research projects [6]. This program has since moved under the direction of the Waterloo Laboratory for Earth Observations (WatLeo) and is now called the Collaborative Interdisciplinary Cryospheric Experiment (C-ICE) program.

Without any historical information (temperature, wind velocities, amount of precipitation, etc.), there does not seem to be a one-to-one mapping of the calibrated SAR pixel values to actual geophysical variables, but finding such a mapping for all seasons is an ultimate goal. Hence, researchers are first attempting to keep as many parameters constant as possible and note the effect of varying a limited number of them on their backscatter coefficient. Scientists and engineers narrow parameters down by looking only at a particular season or season change, looking at the results for sensors with different frequencies over the same region, altering the polarizations, etc. and comparing these backscatters for different ice types.

In closing, understanding the SAR platform parameters and the SAR relationship with the environment is important for proper interpretation of sea ice imagery. The visible differences in texture are clues to the segmentation of the imagery into appropriate ice types. The next chapter will investigate textural methods that can be used for sea ice identification.

Chapter 4

Methods for Texture Feature Analysis

No one has ever provided a definitive, formal description for texture. Many authors have put forth their own definition but they usually recognize that no single statement is complete. "Texture is concerned with the spatial (statistical) distribution of grey tones [52, p.611]." "Texture is the spatial distribution of intensities in image regions perceived by normal human observers to be homogeneous throughout the region [108, p.7]." "Texture is characterized by invariance of certain local measures of properties over an image region [59, p.1167]." "...image texture may be defined as a local arrangement of image irradiances projected from a surface patch of perceptually homogeneous radiances [13, p.55]".

Approaches to texture interpretation can be broken down into three categories: statistical, structural, and model based. Statistical approaches attempt to characterize textures in a probabilistic sense. Often the characterization is based on definitions of the local texture region such as smooth, coarse, grainy, regular, di-

rectional, etc., however, not all the statistical texture feature extraction methods provide statistics that follow such definitions. Statistically based methods may be either spatially or spectrally defined. Popular approaches include features derived from cooccurrence probabilities, from the power spectrum, and from multi-channel filter outputs. Statistical methods have had promising success, motivating their selection for this thesis.

Structural methods are distinctive from statistical approaches since they consider textures to have two fundamental components: a basic primitive that comprises the texture and the spatial organization of these primitives. The primitive is usually a fixed template that can be moved to match a certain image pattern. Since SAR sea ice types seem to have no predictable, consistent repeating pattern, structural methods are not considered here as a potential method for sea ice discrimination.

Model based methods provide another generic approach to texture analysis. By fitting some analytical function to the texture, the texture characteristics can be captured. The parameters of the model provide a specification for the texture under study. Typically the analytical function is based on a two-dimensional stochastic process or random field. Markov random field (MRF) models [9, 11, 19, 21, 24, 65, 66, 78] have been used successfully as well as fractal based approaches [86].

A distinction should be made between texture classification and segmentation. Classification is the assignment of feature vectors to a particular class. Segmentation subdivides an image into its constituent parts by locating boundaries between identifiable textures and by assigning similar texture regions to the same class. Thus, segmentation requires both texture discrimination (similarity measures) and boundary tracking (discontinuity measures). Segmentation approaches can be divided into region based and edge based approaches. Region based approaches are

more applicable for SAR sea ice segmentation since edge information is destroyed by speckle and there is often a mixture of ice types at distinct ice boundaries.

A supervised classification/segmentation approach requires user input to drive the algorithm properly. Typically, some familiarity with the data set and the class assignment procedure is essential for successful results. In the case of texture analysis, the feature vectors (for training or implementation) may be obtained from user-selected (assumed homogeneous) image samples. To select the training data that represents a unique textural class is difficult since the samples must contain sufficient characteristics to describe the class fully. Unsupervised segmentation, on the other hand, demands limited human intervention to encourage the algorithm to run in an automated fashion. No *a priori* selection of homogeneous class representations should be provided to the system. Clearly, an unsupervised scheme is desirable, however, these methods involve considerably more overhead and attention to algorithmic details. For example, determining the number of texture classes is a very difficult problem. Also, a feature extraction approach must be sufficiently robust to gather essential and separable class characteristics and allow for subtle variations within an assumed uniform texture without confusion with other distinct classes.

Texture analysis, with an emphasis on those methodologies used to interpret SAR sea ice imagery, is dealt with in the rest of Chapter 4. There exists an abundance of texture interpretation methods that could be used to study SAR data, but only a selected number of methods are considered here (Section 4.1). Methods to classify features are briefly reviewed in Section 4.2.

4.1 Texture Feature Extraction Methods

Methods for texture feature extraction, with an emphasis placed on SAR sea ice imagery, are described in this section. First order methods are discussed because these approaches provide a good first guess at segmentation of some natural textures. Cooccurrence texture features are the most popular method for analysis of remotely sensed sea ice imagery and this method is presented at some length. Evidence exists that suggests power spectrum approaches provide viable features so they are considered next. Promising, recently introduced methods for texture interpretation are based on multi-channel filtering. The section concludes with a description of these approaches in the context of wavelets and qualities of the human visual system (HVS).

4.1.1 First Order Texture Measures

As mentioned in Chapter 2, texture provides more robust classification information than tone. Tone is not able to consistently identify ice types and is not able to provide any directional information. Kwok *et al.* implement tonal analysis based on expected backscatter signatures of different ice types using look-up tables as a function of seasonal and meteorological conditions at the time of data acquisition [69]. The look-up tables contain data collected from field campaigns over the past ten years. Results using both aerial and simulated ERS-1 data are presented. The technique has been implemented in the geophysical processor system at the Alaska SAR Facility for classification of SAR sea ice data. Although texture should provide more information, the research group uses tonal analysis as a feasible first step to solving the problem of sea ice classification. The original resolution is 25 metres but the pixels are averaged to 200 metre resolution (eight by eight pixel

averaging). Calibration parameters are used to adjust backscatter coefficients to absolute calibration to maintain uniformity between data sets. The authors expect the user to have a high degree of familiarity with the type of data. For example, the clustering program requires the user to enter input parameters including the number of classes, the expected within-class scatter, and the between-class scatter. For their examples, the results are promising.

Filey and Rothbrock [42] used tonal analysis to distinguish between ice and open water and to resolve the details of deformation and motion of ice floes in Seasat SAR imagery. Both local average brightness and local variance are used. Ground resolution of each pixel is 25 metres. Each of the images used in the study is independently enhanced to remove the falloff of mean brightness with range and to stretch the contrast. The mean and variance of each adjacent 15x15 pixel window comprises the data set. This reduction in the resolution is required to minimize the computational effort. The group had success with distinguishing open water and ice as well as tracking the motions of ice floes. No effort was made to classify different ice types.

Identification of new ice using gradient operators applied to passive microwave data was performed by Eppler and Farmer [38]. They found that new ice has unique textural signatures that would assist in their identification. Banded structures are visible in aerial photographs and these same characteristics generate the observed radiometric signatures. First year ice and multi-year ice are subject to more random growth and lack such an organized structure.

4.1.2 Grey Level Cooccurrence Texture Features

Application of cooccurrence texture features for interpreting SAR sea ice imagery are detailed in this section. First, a description of how cooccurrence texture features are calculated is described. The traditional approach is to use a grey level cooccurrence matrix (GLCM). Then, in the context of several specific research papers, application of these features to SAR sea ice imagery is discussed.

Cooccurrence Features Defined

The cooccurrence probabilities provide a second-order method for generating texture features. A brief presentation of the GLCM method follows, but a more complete explanation is provided by Haralick [51, 52]. The matrix contains the conditional joint probabilities of all pairwise combinations of grey levels given two parameters: interpixel distance (δ) and interpixel orientation (θ) (measured counterclockwise from the horizontal). Following Barber and LeDrew [4], the probability measure can be defined by:

$$Pr(x) = \{C_{ij}|(\delta, \theta)\}$$

where C_{ij} (the GLCM) is defined by:

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

P_{ij} represents the number of occurrences of grey levels g_i and g_j and G is the total number of grey levels. The sum in the denominator represents the total number of grey level pairs within a window given a particular (δ,θ). A different

Maximum Probability (MAX)	$max\{C_{ij}\}\forall(i,j)$
Uniformity (UNI)	$\sum_{i=1}^{G}\sum_{j=1}^{G} C_{ij}^2$
Entropy (ENT)	$-\sum_{i=1}^{G}\sum_{j=1}^{G} C_{ij} log C_{ij}$
Dissimilarity (DIS)	$\sum_{i=1}^{G}\sum_{j=1}^{G} C_{ij}\|i-j\|$
Contrast (CON)	$\sum_{i=1}^{G}\sum_{j=1}^{G} C_{ij}(i-j)^2$
Inverse Difference (INV)	$\sum_{i=1}^{G}\sum_{j=1}^{G} \frac{1}{1+\|i-j\|/G} C_{ij}$
Inverse Difference Moment (IDM)	$\sum_{i=1}^{G}\sum_{j=1}^{G} \frac{1}{1+(i-j)^2/G^2} C_{ij}$
Correlation (COR)	$\sum_{i=1}^{G}\sum_{j=1}^{G} \frac{(i-\mu_x)(j-\mu_y)C_{ij}}{\sigma_x\sigma_y}$

where (μ_x,μ_y) and (σ_x,σ_y) are means and standard deviations of row i and column j

Table 4.1: GLCM texture statistics defined.

GLCM is required for each (δ,θ). Pixel separation distances (δ) are usually less than ten. Typically, only four orientations are used: 0, 45, 90, and 135 degrees. The orientation 180 degrees is redundant to 0 degrees, the orientation 225 degrees is redundant to 45 degrees, etc. The average of the four orientations may be used if invariance to spatial rotations is desired. Different statistical information can be determined from each GLCM, however, statistics that are grey level shift invariant are important so that the classification is not a function of tone. Eight such shift invariant statistics are presented in Table 4.1. Uniformity (UNI), Entropy (ENT), and maximum probability (MAX) are linear (scale and shift) invariant statistics.

The statistics extract three fundamental characteristics from the cooccurrence matrices. Moments about the main diagonal indicate the degree of smoothness of the texture. The closer the entries to the main diagonal, the smoother the texture. The statistics dissimilarity (DIS), contrast (CON), inverse difference (INV), and inverse difference moment (IDM) are statistics of this type. Note that DIS and

CON are inversely proportional to INV and IDM respectively. Another fundamental characteristic of the cooccurrence matrix is the uniformity of its entries. If the grey levels in the window tend to be homogeneous, then only a few grey level pairs represent the texture. Non-homogeneity generates many different pairs of grey levels. The statistics maximum probability (MAX), uniformity (UNI), and entropy (ENT) describe homogeneity. The final statistic, correlation (COR), describes the correlation between the grey level pair (g_i,g_j).

Two of the statistics (INV and IDM) are modified by normalizing the grey level difference $(i-j)$ by the number of grey levels (G). The normalization allows these two statistics to measure a specific characteristic of the texture, namely smoothness. Without normalization, these statistics are based on the sum of C_{ij} weighted by a numerical series ($\{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}...\}$ for INV or $\{1, \frac{1}{2}, \frac{1}{5}, \frac{1}{10}, ...\}$ for IDM). For a smooth texture these statistics are likely to sum C_{ij} values that are close to one while coarse textures tend to sum C_{ij} values close to zero. The outcome is a sparser cluster for a smooth texture relative to the coarse texture. This is opposite to the expected effect of these texture statistics when applied to natural imagery. Namely, texture features that use grey level differences should generate relatively tight clusters for smooth textures and sparse clusters for coarse textures. The normalization provides a series that is more linear in nature preventing this drawback. Without the normalization, it is uncertain what characteristic the statistic is determining about the texture. The normalized statistics consistently have a higher classification rate and larger inter-class distances than the unnormalized versions (based on the data sets used in Chapter 5) and hence, will be used throughout this thesis.

To demonstrate the extraction qualities of the cooccurrence data, a comparative example is presented in Figure 4.1. Here, a pair of 3x3 pixel subimages are presented: one represents a coarse texture (a) and the other represents a smooth

texture (b). The number of grey levels is restricted to four, θ is set to zero degrees (the horizontal), and δ is set to one. Note that symmetrical P_{ij} matrices (as prescribed by Haralick [51]) are determined. For a symmetrical GLCM, a pair of grey levels (g_i,g_j) oriented at 0 degrees is also considered as being oriented at 180 degrees so that entries are made at (i,j) and (j,i) in the GLCM.

There are a number of variables that the user sets when using the GLCM method. Since the dimension of the GLCM is indicated by the number of grey levels, attention must be given to grey level quantization. If the number of grey levels is large, the GLCM occupies considerable memory and the computations are burdensome. On the other hand, too few grey levels destroys the texture composition of the sample and has the potential to reduce the effectiveness of the texture features. The size of the sample window is another variable that is user-determined. In the case of slowly varying texture (smooth texture), large analysis windows should be used. High frequency texture (coarse texture) require relatively smaller windows. Better boundary identification is determined with smaller windows and large windows generate more consistent measurements across homogeneous regions. Selection of the interpixel distance does not have a well defined basis. For example, this parameter could be determined in the same manner as the window selection: small δ for fine textures and relatively larger δ for smooth textures. Also, with clearly indicated and placed texture primitives, δ could be related to the distance between the primitives. Selection of θ is data dependent as well since the user tries to select the direction that will provide the most discriminating information. Since SAR sea ice imagery is assumed, perhaps mistakenly [4], to have isotropic data, several orientations may be averaged together.

Texture Subimages:

$$\begin{bmatrix} 3 & 2 & 1 \\ 1 & 0 & 3 \\ 2 & 3 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 1 & 2 \\ 1 & 1 & 2 \\ 1 & 1 & 2 \end{bmatrix}$$

P_{ij} of Subimages:

	0	1	2	3
0	0	1	0	1
1	1	0	1	1
2	0	1	0	2
3	1	1	2	0

(a) coarse texture

	0	1	2	3
0	0	0	0	0
1	0	6	3	0
2	0	3	0	0
3	0	0	0	0

(b) smooth texture

	(a) coarse	(b) smooth
MAX	0.167	0.500
UNI	0.111	0.375
ENT	2.254	1.040
DIS	1.500	0.500
CON	2.833	0.500
HOM	0.431	0.750
IDM	0.383	0.750
COR	-0.193	-0.333

Figure 4.1: GLCM example ($\delta = 1, \theta = 0°, G = 4$).

Applications of Cooccurrence Texture Features to SAR Sea Ice Imagery

Many publications describe application of the GLCM for classification studies. Four papers that specifically use the GLCM for interpretation of SAR sea ice imagery will be discussed (Barber and LeDrew [4], Holmes *et al.* [55], Nystuen and Garcia [83], and Shokr [98]). Holmes *et al.* claim to be the first to publish texture classification results of SAR sea ice imagery. Barber and LeDrew are the only scientists to seriously consider alternative texture classification algorithms. The other researchers essentially default to the most common method (the GLCM). Although each of these papers uses the GLCM, there exist fundamental differences in the treatment of algorithm dependent variables and the data sets used.

Table 4.2 lists the system parameters for the indicated papers. Holmes *et al.* indicate that the L band data provides no additional information over the X-band data and that X_{HV} has greater dynamic range than X_{HH}. Shokr uses data from three sensors: X_{HH} (seven look), L_{HH} (four look), and C_{VV} and a variety of sensor angles.

The parameters used by these authors for generating the cooccurrence features are summarized in Table 4.3. Barber and LeDrew statistically prove that $\delta = 1$ produces a significantly superior classification when compared to $\delta = 5$ or $\delta = 9$. Holmes *et al.* indicate "Experimentation with the distance parameter led to our selection of $\delta = 2$ as the most appropriate" but do not give details of the experimental methodology or results. Shokr experimentally compares $\delta = \{ 1,2,3 \}$ and concludes that $\delta = 2$ is appropriate. Nystuen and Garcia use a variety of distances from one to ten and determine that the results are invariant for distances greater than four. When determining the window size, Shokr finds that the window dimensions of 5, 7, and 9 have no impact on the texture measures and selected the

Author(s)	Platform	Sensor	No. Looks	Resolution	Inc. Angle
Barber & LeDrew	aerial	X_{HH}	7	6m	69°-79°
Holmes *et al.*	aerial	X_{HV}	N.P.	3m	N.P.
Nystuen & Garcia	aerial	X_{HH}	7	16m	N.P.
Shokr	aerial	X_{HH}	7	6m	68°-78°
	satellite	L_{HH}	4	12.5m	N.P.
	aerial	C_{VV}	1	~2m	30°-70°

N.P. - information not provided by authors

Table 4.2: System parameters of selected GLCM studies.

smallest value to minimize computation time. Nystuen and Garcia used a variety of window sizes based on the size of the homogeneous subwindows in the original image. Holmes *et al.* and Shokr average the four different orientations (θ = 0, 45, 90, and 135 degrees) since they assume the texture measures are insensitive to the direction of the sensor. However, Barber and LeDrew's analysis demonstrated that the orientation in the look (range) direction of the sensor produces results that have greater statistical significance.

Preprocessing of the image will also influence the outcome of the texture classification (Table 4.4). Most of the studies reduce the number of grey levels to sixteen to reduce computation load. Reducing the resolution by a filtering scheme can reduce the effect of speckle.

Since different parameters are used by each study, conclusive comparative results are difficult to quantify. Holmes *et al.* classify multi-year and first year smooth ice types with an overall accuracy of 65%. The poor results are probably attributed to the use of only two texture features. Shokr obtains results ranging from 44 to

Author(s)	δ (pixels)	θ (degs.)	Window Size	GLCM Features
Barber & LeDrew	1*,5,10	0*,45,90	25	UNI,COR,ENT, DIS,CON
Holmes *et al.*	2	average	radius=5	ENT,CON
Nystuen & Garcia	4*-10	variety	mixture	COR,CON,HOM, ENT,clust.prom.
Shokr	1,2*,3	average	5*,7,9	CON,ENT,UNI, HOM,MAX

* indicates authors' preferred parameter value

Table 4.3: GLCM parameters of selected GLCM studies.

Author	Sampled Resolution	Grey Level (b - bits)	Filtering
Barber & LeDrew	6m	8b to 4b	adaptive compared to no filtering
Holmes *et al.*	6m	8b to 3b	N.P.
Nystuen & Garcia	N.P.	N.P.	N.P.
Shokr	36m(X-band) 50m(L-band) ~30m(C-band)	8b to 4b* and 8b to 5b	median filter, Frost filter (C-band only)

N.P. - information Not Provided by authors

* indicates authors' preferred parameter value

Table 4.4: Image preprocessing used in selected GLCM studies.

90% overall classification. He shows that ice types are uniquely identified by the mean value of any texture parameter, however, the variability is confusing. Barber and LeDrew performed multivariate analysis because they feel that "information contained within the GLCMs is not adequately captured by any single texture statistic [25]." K-hat measures [81] of 90% for training data and 70% for cross validation data are obtained. Barber and LeDrew admit that to make a universal statement from their results is premature and more studies must be completed. Nystuen and Garcia obtained 75% classification for their best set of parameters. They found the range of pixels within a sample window to be an informative texture feature.

4.1.3 Power Spectrum

Spectral statistical feature extraction generates different information compared to the spatial domain. For example, prominent peaks in the power spectrum indicate principal direction of texture patterns and location of the peaks gives the fundamental spatial period of the patterns. Typical noise processes tend to dramatically alter local spatial variation of intensity while having relatively uniform representation in spatial frequency. As a result, frequency domain measures should be less sensitive to such noise processes [74].

The discrete Fourier transform (DFT) $F(u,v)$ of a two dimensional signal $f(x,y)$ is defined by [48]:

$$F(u,v) \equiv \frac{1}{N^2} \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) \exp\{-j2\pi(ux+vy)/N\}$$

and the inverse DFT is defined by:

$$f(x,y) \equiv \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F(u,v) \exp\{j2\pi(ux+vy)/N\}$$

where N is the square image or subimage dimension. This transform can be performed quickly by using the fast Fourier transform algorithm.

There exists a one-to-one mapping between the DFT and its inverse. The power spectrum is simply $|F(u,v)|^2 = F(u,v)F^*(u,v)$, where * denotes the complex conjugate. The power spectrum of the image is the Fourier transform of the autocorrelation function. Often, the origin (which represents the D.C. component) is shifted to the centre of the image. The total energy excluding the D.C. component in the image is:

$$\phi \equiv \sum_{u=-N/2}^{N/2} \sum_{v=-N/2}^{N/2} |F(u,v)|^2 \quad (u,v \neq 0)$$

Since power spectrums of real images are symmetric, only $v \geq 0$ is considered.

One may also consider the power spectrum as a function of polar coordinates, $F(r,\alpha)$. Using this coordinate system, the radial distribution of the power spectrum is sensitive to texture coarseness. For example, a small bandwidth of low dominant frequency components (ie. high values close to the origin in the power spectrum) will indicate a smooth texture. A power spectrum with a high bandwidth (ie. power spectrum that is more uniformly placed over the entire domain) will tend to characterize fine textures. Texture features may be derived by summing the normalized energy as a function of a defined range of its radial component (r) over the entire range of its orientation component (α):

$$\phi_r \equiv \frac{1}{\phi} \int_{\alpha=0}^{\pi} \int_{r=r_1}^{r_2} |F(r,\alpha)|^2 dr d\alpha.$$

These features are referred to as "rings" or "annuli". The angular distribution of values in the power spectrum is sensitive to the directionality of the texture and may be defined by summing the entire radial range within a defined orientation range:

$$\phi_\alpha \equiv \frac{1}{\phi} \int_{\alpha=\alpha_1}^{\alpha_2} \int_{r=0}^{N/2} |F(r,\alpha)|^2 dr d\alpha.$$

Such features are referred to as "wedges". Intersections of rings and wedges will sum energies in particular regions in the power spectrum and these values are used as texture features. D'Astous [27] recognizes the shortcomings of only considering summed features as texture features from the power spectrum. For example, two textures with quite different spectra would be judged as the same as long as their respective ring/wedge summed energies were the same. She developed a group of power spectrum features that recognizes attributes such as the extent of regularity, directionality, linearity and coarseness, which are not reflected by traditional summed energy features. These new features include an entropy based approach as well as peak and shape characterizations. The entropy based approach did not seem to generate features that performed as well as cooccurrence features for natural textures, although the entropy measures reflect substantially different information about the texture characteristics than the summed energy feature [61].

Peak and shape characterizations are important power spectrum features. The degree of regularity in texture corresponds to peaks in the power spectrum. Location and concentration of the peaks relates to the coarseness and directionality in the texture. By isolating critical peaks, texture features may be obtained by measuring their particular characteristics. The D.C. component is not considered for peak analysis to generate features invariant to local mean grey level. Features are summarized in Table 4.5 where $k(u,v)$ denotes the energy at the highest point

Percentage Energy (PCT)	$k(u,v)/\phi * 100$
Laplacian (LAP)	$k(u+1,v) + k(u-1,v) +$ $k(u,v+1) + k(u,v-1) - 4k(u,v)$
Skew of Peak (SKW)	$\#\{$ neigbours of peak $> 0.5 * k(u,v)\}$
Distance of Peak from Origin (DIS)	$u^2 + v^2$
Relative Peak Angle (ANG)	$tan^{-1}(v/u)$

Table 4.5: Peak texture features based on power spectrum.

of a peak. Percentage energy of peak (PCT) represents the percentage of total energy of a peak. The Laplacian (LAP) takes on a small value for a flat peak and a large value for a pointed peak. Since a skewed peak and a symmetric peak are not discriminated under a Laplacian measure, counting the number of neighbours in the peak which contain at least fifty percent of the peak energy provides a discriminating measure (SKW). The distance (DIS) of the peak from the origin interprets coarseness (small distance) or fineness (large distance). The angle of the peak (ANG) takes on values between zero and π radians due to symmetry of the power spectrum. Absolute peak angles are sensitive to the rotation of the image, thus, the angle between two highest peaks is used as a texture feature. Typically, the highest and second highest peaks are used when determining the above features.

Shape features provide information concerning the overall shape of the power frequency spectrum. Highly directional texture corresponds to an elongated elliptical shape distribution, while the distribution of an isotropic texture is much more circular. A two-dimensional probability density function of the spatial frequency components is obtained by dividing the energy of each frequency component by ϕ. Each (u,v) in the power spectrum is then a random vector with associated probability $p(u,v)$. The mean of the distribution is (0,0) due to symmetry about the

origin. The covariance matrix is:

$$S = \begin{bmatrix} \sigma_{uu} & \sigma_{uv} \\ \sigma_{uv} & \sigma_{vv} \end{bmatrix}$$

where

$$\sigma_{uu} = \sum_{v=-N/2}^{N/2} \sum_{u=-N/2}^{N/2} u^2 p(u,v),$$

$$\sigma_{vv} = \sum_{v=-N/2}^{N/2} \sum_{u=-N/2}^{N/2} v^2 p(u,v),$$

and

$$\sigma_{uv} = \sum_{v=-N/2}^{N/2} \sum_{u=-N/2}^{N/2} uv p(u,v).$$

Based on these measurements, different features can be generated (Table 4.6). Isotropy (ISO) measures the elongation of the probability distribution and has a range of values from zero to one. For an isotropic feature, the feature value tends to zero. If the distribution only consists of straight lines ($\sigma_u = 0$ or $\sigma_v = 0$), then the measure tends to one. The spread of the spatial frequency component distribution (SPD) is determined using the square root of the two eigenvalues of the covariance matrix, λ_1 and λ_2. The circularity measure (CIR) compares the area of the distribution to the area of a circle with a radius equal to the length of the major axis. The fraction of the area of the distribution compared to the encompassing circle determines the texture feature.

D'Astous [27] demonstrates how different window sizes should not affect the various features that can be derived from the power spectrum. As long as at least one full period of the texture is contained within the window, power spectrum

Isotropy (ISO)	$\frac{\lvert\sigma_u - \sigma_v\rvert}{\sqrt{(\sigma_u+\sigma_v)^2 - 4\sigma_{uv}^2}}$
Spread of Distribution (SPD)	$\sqrt{\lambda_1}, \sqrt{\lambda_2}$
Circularity Feature (CIR)	$\frac{A_D}{A_C}$

Table 4.6: Shape texture features based on power spectrum.

features should be consistent. Using this concept, D'Astous describes an hierarchical split-and-merge approach to segment a full image. Like any other windowed feature extraction process, confusion occurs when multiple textures appear in the same subimage. This is the problem when the power spectrum of a full image is calculated. From the power spectrum, it is impossible to tell which dominant frequencies belong to which spatial coordinates since there is no one-to-one inverse transformation of the power spectrum back to the spatial domain. Subdividing the image using a split-and-merge approach circumvents this shortcoming.

Unlike the cooccurrence texture feature approach, no study has been found in the research literature that specifically investigates the potential of the different power spectrum texture features for application to SAR sea ice imagery. Power spectrum features have typically been developed and tested on Brodatz imagery.

4.1.4 Multi-Channel Filtering

Multi-channel filtering is a more recent approach to texture analysis. Processing the image using multiple resolution techniques, filter banks have the ability to decompose an image into relevant texture features that can then be classified accordingly. Multi-channel filtering mimics characteristics of the human visual system (HVS). This section will first describe how the HVS system processes information in a multi-channel sense. Then, the theory of wavelets is introduced followed by a de-

scription of a special wavelet function, the Gabor function. A discussion of methods to implement the Gabor functions for texture analysis conclude the section.

The Role of the HVS in Texture Interpretation

In order to better implement artificial vision systems, scientists have attempted to improve our understanding of the human visual system. Just as birds were once used as a natural system to help unlock the mystery of flight, scientists today use animal models to improve our understanding of biological vision. Implementation of artificial vision systems leads to an improved understanding and gives us a better appreciation of the abilities of the natural system. Granted, we do not design airplanes that have wings that flap, however, birds have given us the knowledge and inspiration to accomplish what might have been considered to be not possible. Determining how a wing works and how the brain operates involves different degrees of complexity. Little by little we have improved our understanding of how the HVS operates and we are learning how to better implement this knowledge effectively. Also, the HVS can also serve as a convenient basis to measure the performance of artificial vision systems. For example, two images are said to have the same texture if they are not effortlessly discriminable to the human observer [23].

Julesz presented paradigms for the HVS that have been widely publicized and discussed [62, 63]. He provided evidence that the HVS may not be sensitive to differences in probability distributions with orders greater than two. Recent work has indicated that sensitivity to higher order statistics is possible [64]. That second-order relationships are important provided impetus for development of the cooccurrence texture features. He also developed the concept of "textons" to explain preattentive aspects of human visual perception. Textons are suspected to provide a first-order cue to the HVS which directs the visual system to determine the posi-

tional relationship of the textons in the scene. In this manner, patterns in the scene are identified. Textons are structural in nature and, as a result, do not provide the type of discrimination needed for SAR sea ice imagery. Other research has provided more encouraging models for providing features for texture interpretation.

The human visual system (HVS) is extremely complex. A highly abbreviated explanation for the interconnections is as follows. Receptor cells are found on the back inner surface of the eyes. From there, electrical impulses are passed through various nervous pathways through the optic nerve and finally to the visual cortex (located posteriorly in the brain). Within the visual cortex, it is known that various cells perform different types of processing on the incoming signals. Hubel and Wiesel [56] tested responses of the visual cortex in cats and deduced that simple cells are tuned to specific orientations. That is, as a bar is rotated through a visual field, individual simple cells respond only when the bar is within a certain range of frequencies. The range is not more than 30° for a given cell. Campbell and Kulikowski [15] carried this concept further and demonstrated that humans have orientation sensitivity as well as a spatial frequency sensitivity. This led to their model [16, p.564] that the HVS is made up of "a number of independent detector mechanisms each preceded by a relatively narrow-band filter 'tuned' to a different frequency. Each filter and detector would constitute a separate 'channel' ...". Experiments have shown that the frequency bandwidth of simple cells in the visual cortex is about one octave (see Pollen and Ronner [87]). This HVS multi-channel filtering model agrees with a very popular generic approach to signal decomposition, namely, wavelet analysis.

The research by Rao and Lohse [90] provides additional insight to this idea of orientation and frequency sensitivity but from a human perception slant. Subjects were asked to classify pictures from the Brodatz album [14] based solely on their

own perception. Results indicate that people essentially use three high-level features for texture discrimination, namely: repetition, directionality, and complexity. Repetition and directionality are representative of frequency and orientation. Complexity relates to the consistency of the texture. For example, a purely sinusoidal texture would have low complexity and a texture without any well defined pattern would have high complexity. These three characteristics are important when investigating the ability of feature extraction methods to accurately identify different textures.

These experiments predict two fundamentally different theories of what is important for texture interpretation within the HVS. One is based on oriented frequency components and the other is based on structural recognition. The argument is similar to the historical debate over the wave versus particle theories of light. Which method is appropriate? Today we realize that a dual model is a better explanation of how light propagates. There is ample evidence to support that a combination of the above two models are a complete way to fully explain mechanisms of the HVS. However, which methods are most important when creating an artificial approach to mimic the characteristics of the HVS and implement a machine aided approach to texture segmentation?

Wavelets - What Are They?

This section serves as a brief introduction to the rapidly growing field of multichannel analysis. Further introductory information can be found in [77, 94, 100]. Wavelet analysis has the ability to mimic the frequency and orientation sensitivities characteristic of the HVS. Wavelets ("little waves") are a special group of signals that have been applied in a wide variety of signal processing applications, including digital image analysis. Essentially, they are bandpass filters designed with certain

characteristics [109]: they are oscillatory, they have a fast decay towards zero (in all directions), and they also integrate to zero. Sets of wavelets are created by modifying the original wavelet (often referred to as the mother wavelet) by scaling (dilation or compression) and translating (shifting along an axis). For example, if $g(x)$ represents the mother wavelet then:

$$\frac{1}{\sqrt{a}} g\left(\frac{x-b}{a}\right)$$

represents the set of wavelets as a function of scaling (a) and shifting (b). The $1/\sqrt{a}$ term is used for energy normalization. The wavelet transform of a one-dimensional signal actually produces a two-dimensional representation (as a function of a and b).

A continuous time wavelet transform (CTWT) is defined as an operation performed on the input signal $f(x)$ as a function of a and b with respect to a mother wavelet $g(x)$ [109]:

$$W_g[f(x)](a,b) = \frac{1}{\sqrt{a}} \int f(x) g^* \left(\frac{x-b}{a}\right) dx$$

.

$W_g[f(x)](a,b)$, a linear transformation, generates the wavelet coefficients. There exists a discrete wavelet transform (analogous to the Fourier series, not the discrete Fourier transform) known as the continuous time wavelet series (CTWS). Here, discrete refers to the transform domain parameter and not the independent variable of the function that is being transformed. Thus, the scales and translations are discrete, but the independent variable is continuous. The discrete time wavelet series (DTWS) is analogous to the discrete Fourier transform (DFT). Here, both the independent variable as well as the scales and translations are discrete. The DTWS is used for image analysis, the CTWS for continuous signal analysis, and the CTWT is often used to derive properties for the other two transforms.

By definition, a Fourier transform integrates the independent variable from $-\infty$ to $+\infty$. If only a finite range (a window) is used to generate the Fourier transform (as in Section 4.2.2), the transform is referred to as the short time Fourier transform (STFT). The STFT is an acceptable means of analysis for locally stationary signals, since dominant frequency components are identifiable and further processing can be based on these components. Wavelets provide preferable characteristics when the image or signal under consideration is non-stationary and wideband. SAR sea ice imagery is an example of such imagery.

The scaling and translation properties of the wavelet transform make it an attractive tool for a host of signal processing applications. The number of oscillations the wavelet experiences is independent of the scaling and translation. A scaling parameter less than one causes compression of the mother wavelet. In this case, more cycles occur in a shorter time frame which indicates higher frequencies and, inherently, the wavelet transform maps a shorter time interval to higher frequencies. Similarly, a longer time interval is mapped to lower frequencies. When attempting to segment a SAR sea ice image, this inherent ability has tremendous potential. A wavelet transform can be applied so that compressed wavelets identify multi-year ice types and dilated wavelets identify smooth ice types. The characteristic line patterns of pressure ridges would have high frequency oriented components that the wavelet analysis could be "tuned" to isolate. Generally, the wavelet transform can take a signal, break it down into component pieces, and the manipulation of these pieces can yield features that represent characteristics of the various textures that appear in the image. Such multi-resolutional filtering gives the opportunity to dissect an image and isolate the essential details necessary for segmentation.

Gabor Functions Implemented for Multi-Channel Filtering

Research has demonstrated that the HVS generates a multi-resolutional decomposition. Since wavelets are intrinsically multi-resolutional, they have been implemented successfully in texture analysis models. A popular multi-channel filter is the Gabor function. This function is not truly a wavelet (ie. in the mathematical sense), however, it can be implemented in such a manner as to mimic properties of a wavelet and it has properties that make it attractive for computer vision applications. These properties include its potential ability to mimic the behaviour of simple cells, appealing simplicity, and optimum joint spatial/spatial-frequency localization.

There is evidence to indicate that this pseudo-wavelet models the general nature of the two-dimensional receptive fields of simple cells in the visual cortex [28]. In linear systems language, the Gabor function acts as the impulse response for simple cells. Also, the Gabor function has the ability to isolate specific frequencies and orientations, well recognized properties of the HVS.

In terms of functionality, a Gabor function is a Gaussian modulated sinusoid. It is defined as a product of an elliptical Gaussian (aspect ratio σ_x/σ_y, centre (X, Y)) and a complex exponential (spatial frequency $F = \sqrt{U^2 + V^2}$, orientation $\theta = tan^{-1}(V/U)$). The *complete* Gabor function [28] can be represented spatially by:

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left\{-\frac{1}{2}\left(\frac{(x-X)^2}{\sigma_x^2} + \frac{(y-Y)^2}{\sigma_y^2}\right)\right\} \exp\{2\pi j[U(x-X)+V(y-Y)]\}$$

and its spatial-frequency domain representation is:

$$F(u,v) = \exp\left\{-2\pi^2\left[(u-U)^2\sigma_x^2 + (v-V)^2\sigma_y^2\right]\right\} \cdot \exp\{-2\pi j[X(u-U)+Y(v-V)]\}.$$

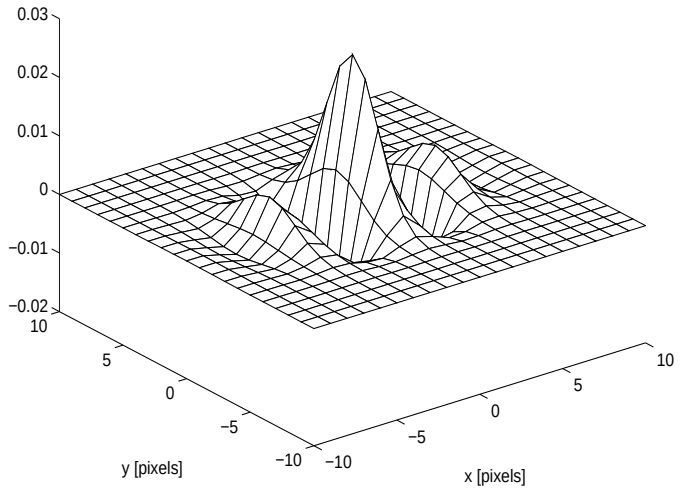


Figure 4.2: Even-symmetric Gabor function (spatial domain).

An example of the even-symmetric (real) part of apart of a Gabor filter is presented in Figure 4.2. Here, $X = Y = V = 0$, $U = \pm 0.2$ cpp (cycles per pixel), $\sigma_x = 3$ pixels, and $\sigma_y = 2$ pixels. The frequency domain representation is shown in Figure 4.3.

Selection of σ_x and σ_y determines the resolution in both the spatial and spatial-frequency domains. Low values of σ_x and σ_y favour spatial resolution and their high values favour spatial-frequency resolution. When segmenting an image, short spatial intervals are preferable because we wish to approximate the boundary between textures. However, smaller frequency bandwidths are preferable to make better distinctions among different textures. Unfortunately, the spatial extent and the spatial-frequency bandwidth have an inverse relationship. This tradeoff (conflicting goals of simultaneous spectral and spatial localization) is known as the uncertainty principle [12, 59]. A Gabor function has an important property: optimal joint resolution in the spatial and spatial frequency domains. If Δx represents the resolution

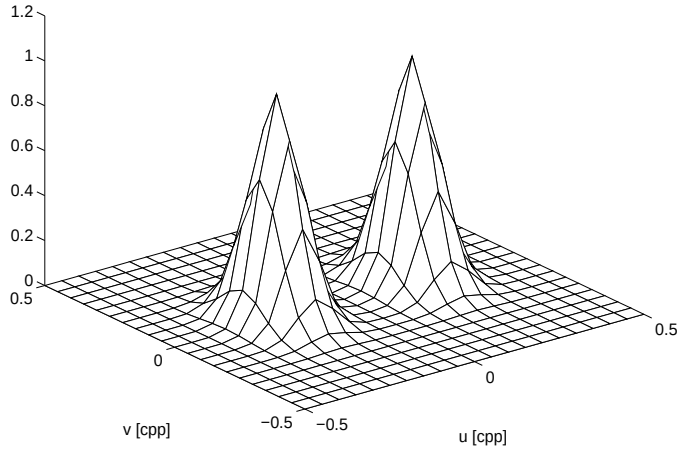


Figure 4.3: Even-symmetric Gabor function (spatial-frequency domain.)

in the spatial domain and Δu represents the resolution in the spatial-frequency domain, then the joint resolution is bounded by the inequality:

$$(\Delta x)(\Delta u) \geq \frac{1}{4\pi}$$

The one-dimensional Gabor function achieves the equality of this formula. A similar criterium is met in two-dimensions [28].

Implementation of Multi-channel Filters For Texture Interpretation

When describing the cooccurrence technique, there were several research papers that were used to describe different implementation approaches of the cooccurrence data to SAR sea ice imagery. It would be appropriate to demonstrate the implementation issues of Gabor and other multi-channel filters applied to SAR sea

ice imagery, but the literature is quite limited in this area. As a result, examples will be drawn from methods used for generic texture analysis.

When generating texture features using multi-channel filters, two primary issues must be addressed. The first issue deals with the functional characterization of the channels as well as their number, orientation, and spacing. The second issue deals with extracting significant features by data integration from the different channels. These two issues are discussed in this section in the context of Gabor filters.

Functional Characterization

The position (F,θ) and bandwidth (σ_x,σ_y) of Gabor filters in the frequency domain must be carefully set up to capture essential texture discrimination information. This is an important issue because centre frequencies of channel filters must agree very closely with the emergent texture frequencies or the channel frequency responses fall off too rapidly within a small neighbourhood of the centre frequencies. This assumes that some identifiable features for a particular texture are sufficiently spaced in frequency so that cross-channel interference does not occur. Since Gabor filters are not fully bandlimited, some aliasing will always occur regardless of the filter density. Also, since there are many combinations of (F, θ, σ_x, σ_y), minimizing the number of filters is highly desirable from a computational perspective. Numerous methods to characterize Gabor filters to perform texture interpretation will be discussed. These include: locating peaks in the image spectrum, using pyramidal schemes, optimizing separability between two textures, or sampling the frequency plane with a sufficiently dense set of channels.

Bovik [13] mentions three supervised approaches to selecting filter locations using empirical information based on the power spectrum characteristics of the in-

dividual textures. For strongly oriented textures, the most significant spectral peak along the dominant orientation direction is used as a filter location. Picking the lower fundamental frequency identifies periodic textures. Finally, for nonoriented textures, using the centre frequencies of the two largest maxima is recommended. Obviously, an automated method is far more appealing and practical. Also, identification usually requires multiple peaks for each texture in the SAR image, making this selection approach unwieldy. Identifying unique textures with small regions of support is also difficult.

Dunn and Higgins [35] develop a method to select optimal filter parameters based on known samples of the textures. The objective of this wholly supervised approach is to develop better methodologies to discriminate textures based on identifying textural boundaries using only a minimum number of filters. Only "the" particular filter which optimally separates two texture classes (in terms of their one-dimensional density functions) is used to partition an image. The "optimal" filter may reflect strong textural characteristics of one class but may actually express a lack of textural information (for that given frequency and orientation) of the other class. The other class is not being identified as having a particular characteristic, but lacking a characteristic of the other class. The output of the Gabor filters is modelled as a Rician probability density function. To use just a single filter to discriminate each texture pair may be difficult because each texture is subject to spatial variability and the texture may contain multiple dominant components. If multiple filters are used, how does one determine the number of filters to select?

Instead of trying to identify unique peaks that belong to unique textures, another approach to filter characterization is to spread the filters throughout the frequency domain field to capture salient information. By providing near uniform coverage of the spatial-frequency domain with Gabor filters, the issue of selection

of the centre frequencies is avoided. Wright [108] empirically derived preferred filter parameters in terms of peak frequencies and bandwidths based on preattentive visual discrimination experiments. His experiments provided evidence for dividing the orientation into at least seven uniformly spaced channels (ranging over 180°). In the frequency direction, the set of (1.7,2.8,4.3,7.1,10.3,13.0) cycles per degree and respective bandwidths of (2.0,1.9,1.7,1.6,1.5,1.3) octaves are recommended for texture feature extraction. Fernandes and Jernigan [41] applied Wright's specifications to discriminate textures in satellite forestry imagery.

When Gabor functions are applied as linear filters, the centre defined by (X, Y) in the spatial domain is unnecessary since the function will be implemented using convolution. In the spatial-frequency domain, these filters appear as simple Gaussians centred on (F, θ). Thus, Gabor filters are actually band pass filters. From a digital signal processing perspective, this type of analysis is quite straightforward to implement and utilize. A complete (even and odd components) Gabor filter can be represented spatially as:

$$h(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left\{-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right\} \exp(2\pi j F x).$$

It is assumed that the x-axis of the Gaussian has the same orientation (θ) as the frequency (F). Bovik agrees that this is the most convenient approach [13]. Rotation by θ in the x-y plane provides for any arbitrary orientation of the filter. The corresponding representation in the spatial-frequency domain is:

$$H(u, v) = \exp\{-2\pi^2[(u - F)^2\sigma_x^2 + v^2\sigma_y^2]\}$$

where again a rotation can be used to obtain any arbitrary θ in the u-v plane.

Jain and Farrokhnia [59] use a bank of even-symmetric Gabor filters to characterize the channels. The authors justify using only even-symmetric Gabor filters

on psychophysical grounds, but they provide no full explanation of this basis. If one were to use only the real (even-symmetric) component of a Gabor filter, then:

$$h(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left\{-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right)\right\} \cos(2\pi F x).$$

and the representation in the spatial-frequency domain becomes:

$$H(u,v) = \exp\{-2\pi^2[(u-F)^2\sigma_x^2 + v^2\sigma_y^2]\} + \exp\{-2\pi^2[(u+F)^2\sigma_x^2 + v^2\sigma_y^2]\}$$

Note that in the even-symmetric representation (such as in Figure 4.3), two Gaussians appear in the spatial frequency domain compared to only a single Gaussian in the complex case.

Placing the filters evenly across the spatial-frequency domain assists decisions about the filter parameters. There are six unknowns when implementing the filters: F, θ, σ_x, σ_y, B_F, and B_θ. The radial frequencies (F) can be selected by using psychovisual data. Jain and Farrokhnia implement F in octave format starting at $\sqrt{2}$ cycles per image (cpi). Since low frequencies are considered not to contribute to textural differences, they do not consider the two lowest frequencies (for images dimensioned to 256x256 pixels). Four equally spaced orientations are used: 0, 45, 90, and 135 degrees. In terms of psychovisual data, finer orientation selectivity may be required, but they provide no such analysis. Setting the variables B_F and B_θ (frequency and angular bandwidths) carefully will ensure proper coverage of the spatial-frequency domain. Since distances between centre frequencies are one octave, it makes sense to set B_F to 1 octave. Similarly, it makes sense to set B_θ to 45°.

In order to determine σ_x and σ_y, two equations are required. If B_F is determined as a function of position on the u-axis, then a representation for σ_x can be determined. Calculating the half-bandwidth (-6db) response when $v = 0$:

$$H(u,0) = \frac{1}{2} = \exp\{-2\pi^2\sigma^2(u-F)^2\sigma_x^2\}$$

and solving for u yields:

$$u = F \pm \frac{\sqrt{\ln 2}}{\sqrt{2}\pi\sigma_x}$$

Since the octave distance between two frequencies is simply the $\log_2$ of their ratios, then the radial bandwidth may be expressed as:

$$B_F = \log_2 \left[\frac{F\pi\sigma_x + \sqrt{ln(2)/2}}{F\pi\sigma_x - \sqrt{ln(2)/2}} \right]$$

Solving for σ_x yields:

$$\sigma_x = \frac{\sqrt{\ln 2}(2^{B_F}+1)}{\sqrt{2}\pi F(2^{B_F}-1)} \tag{4.1}$$

The second equation is determined by the angular half-bandwidth. Setting $u = F$:

$$H(F,v) = \frac{1}{2} = \exp\{-2\pi^2\sigma_y^2 v^2\}$$

and solving for v yields:

$$v = \pm\frac{\sqrt{\ln 2}}{\sqrt{2}\pi\sigma_y}$$

The tangent of the angular half-bandwidth is then found as:

$$\tan\left(\frac{B_\theta}{2}\right) = \left[\frac{\sqrt{\ln 2}}{F\sqrt{2}\pi\sigma_y} \right] \tag{4.2}$$

Solving for σ_y yields:

$$\sigma_y = \frac{\sqrt{\ln 2}}{F\pi\sqrt{2}\tan\left(B_\theta/2\right)}$$

A circular Gaussian may be desired to have a consistent spatial extent in all directions. In this case, σ_x is set to σ_y ($\sigma = \sigma_x = \sigma_y$) and B_θ is calculated. Equation 4.1 is used to determine σ and Equation 4.2 is used to determine B_θ. If a constant value of σ is selected for every (F, θ) pair then the decomposition would be that of a STFT. From a practical perspective, it makes sense to increase the bandwidth of the filter with increasing frequency since the power spectrum typically decreases in a logarithmic fashion and since this corresponds to narrower impulse responses for higher frequencies.

A problem with simultaneously capturing information across the entire spatial-frequency domain is that a considerable amount of data must be generated and manipulated. For example, given a 256x256 image and Wright's set of features, $256^2 \cdot 4bytes \cdot 42filters \cdot 2 = 22Mb$ of active memory (given float precision and the fact that real and imaginary components will be generated). In contrast, when information is transferred from the eye to the optic nerve, there is a considerable compression in the number of nerve fibres carrying sensory information. This leads to the possibility that the HVS is performing coding prior to the data reaching the visual cortex. However, this model of multi-channel filtering only multiplies the amount of data that must be processed.

Orthogonal wavelet implementations allow for rapid and straightforward generation of the wavelet coefficients using a matrix based filter bank. Since Gabor functions are not orthogonal filters, this can lead to redundant information being produced in the output of the filters. If the redundancy is low, there should be no

real concern for the texture interpretation schemes utilizing the Gabor function. Jain and Farrokhnia [59] perform a Gabor decomposition using near uniform coverage of the spatial-frequency domain. They demonstrate that reconstruction based on the filtered images generates an accurate reproduction of the original image. Mallat [77] prefers using basis functions that are orthogonal for multi-resolution analysis.

Filter Integration

There are a number of different methods to integrate the outputs from a bank of Gabor filters. Mallat [77] mentions that there is "no statistical model to combine the information provided by different channels". This is noted because considerable information is produced by the Gabor filter analysis and it is the efficient manipulation of this data that will provide appropriate texture features.

Scientists typically agree that a nonlinearity is essential - this matches biological functioning and generates computer implementation success. Texture identification is usually performed based on the magnitude of the output of the Gabor functions [13, 36]. In the case of a filter that "matches" the particular texture, the magnitude of the output is large, hopefully larger than any of the other filter outputs to enable identification. Filters that do not match the frequency components of dominant texture characteristics should have a negligible response and can be safely ignored as characteristic of that particular texture. Low responses to any filter may be important to identifying a texture if the responses are consistent. Instead of using the magnitude responses, Bigun and du Buf [10] use moments of the Gabor responses to reduce the number of texture features generated.

Malik and Perona [76] advocate the use of half-wave rectification, claiming that

full wave rectification and power spectrum responses destroy information necessary for texture discrimination. Their example looks at a Gabor-like function (strong central peak and negative side-lobes). Scaling the function by -1 leads to a texton that is visibly different. This will change the sign of the Gabor filter response and a magnitude calculation will destroy this information. Their approach of splitting the output into two half-wave rectifications, one for positive and one for negative, will recognize such a difference. Hence, their discrimination is actually based on tonal differences. Dunn *et al.* [36] suggests an alternative approach that maintains information indicated by differences in sign and phase.

Porat and Zeevi [88] attempt to identify the dominant local frequency and orientation component. By calculating moments of the responses, the preferred (F, θ) pair and their variances are determined. No consideration is given to the confusion that would arise in the common case of natural textures that have multiple dominant (F, θ) pairs. They concatenate their feature set by using local mean and variance of the grey levels as texture features.

The importance of phase has been the subject of some discussion, however, it has proven to be not as useful as the magnitude Bovik *et al.* [13] determine boundary identification using consistent textures with a phase shift generating a "break" in the homogeneity of the texture. The boundary identification being performed does not seem suitable for the needs of segmenting SAR sea ice imagery. Du Buf [30] discusses at length the relevance of local phase information, but recognizes that "... the quality of this information degrades if the textures are disturbed by a small amount of jitter or by band-limited additive noise [31, p.227]." Thus, such features would be very delicate if applied to natural imagery. That the phase of an image is more important to regenerating an image than its magnitude is noteworthy (see [73] for discussion), however, the contribution of phase to image texture segmentation

is still uncertain.

A nonlinear function applied to the wavelet response does not typically generate a sufficiently smooth signal to allow for a consistent segmentation. Jain and Farrokhnia [59] subject each filtered image to a nonlinear transformation followed by local smoothing. The nonlinear transformation is reminiscent of a sigmoidal activation function used in artificial neural networks and the window size used for the local smoothing is a function of the centre frequency of that particular channel. Bovik *et al.* [13] understand that textures which do not have sufficiently narrow bandwidths suffer from "leakage". They reduce the effects of leakage by postfiltering the channel amplitudes with Gaussian filters having the same shape as the corresponding channel filters but greater spatial extents. This approach can lead to erroneous boundary estimates.

Very little work has been done in the area of logically combining the outputs of the different Gabor filters to match observed biological vision characteristics. Manjunath and Chellappa [79] use competitive and cooperative processing of the Gabor filter outputs to help in noise suppression and to reduce the effects of illumination. The result is an approach that can identify edges, textures, and illusory contours. They also comment that complex cells in the visual cortex respond to complex patterns such as textures. Unlike simple cells, these complex cells do not contain any phase information, are less sensitive to precise location but are tuned to respond to different specific orientations and direction of movement. Modelling of these cells is performed by summing outputs of similar simple cells.

4.1.5 Other Texture Feature Extraction Methods

There are a wide assortment of additional texture feature extraction methods. These will be briefly summarized here.

Another method for studying texture identification is the use of texture synthesis. By assigning different parameters to a particular generator, different synthetic image textures are produced. If the texture of an image matches the texture of a synthetic image, this provides a better theoretical and visual means of understanding the texture. Several of these approaches are presented by du Buf *et al.* [32]. This method is commonly used to produce images to test texture analysis algorithms. Textural synthesis can also be used for the artificial generation of two textures so that the responses of all features under study are equal for the two fields. This is the method of choice for Wright [108]. Hall and Giannakis [50] use higher than second-order statistics to characterize texture images, which is counter to Julesz's conjecture regarding the adequacy of second-order statistics for visual discrimination of textures [62].

Markov modelling approaches that are used to produce texture synthesis are also used directly for texture feature extraction. In order to classify rotated and scaled texture images consistently, Cohen *et al.* utilized Gaussian Markov random fields (GMRF) [24]. Markov random field models were also used by Manjunath *et al.* for texture segmentation purposes [80]. They recognize that a search for an optimal solution is not possible because of the large dimensionality of the space, so alternative solution methods must be employed. Bello [9] applies a Markov random field to outputs from a wavelet filter bank. A combined multiple resolution and Markov random field analysis is performed by Bouman and Liu [11].

Laws [70] empirically developed what he called micromasks, local masks that

detect certain texture characteristics such as edges, lines, ripples, and spots. Bovik comments that some of these masks have Gabor-like representations [13]. Chen and Kundu [21] expand Laws approach into directional macromasks that determine a local estimate of the directional energy. A 2-D Wold-like decomposition [44, 45] is used by Francos *et al.* to break down an image into two mutually orthogonal parts: an indeterministic component (modelled by a 2-D autoregressive model) and a deterministic component (representing the harmonic nature of the texture). The representation of the deterministic component is similar to the simple texture model presented by Bovik [13]. Reed *et al.* use a Wigner distribution in a region growing context to perform texture segmentation [93]. McLean [82] uses a discrete cosine transformation (DCT) to generate texture features. This is the same transformation used for encoding in the JPEG and MPEG picture and moving picture formats.

An interesting formulation to model texture was proposed by Randen and Husϕy [89]. They perform texture feature extraction based on the design of an optimal finite impulse response (FIR) filter (optimal with respect to the Fisher's optimality criterion described in Section 4.2.2). They are able to find a closed form solution only if the filtered image is assumed to be a separable autoregressive process. They found that their method is quite successful for bipartite images. Although other methods exist that can easily accomplish the same or better results, the goal of optimal filters is preferred to reduce the amount of data.

Other multi-channel filtering methods have been used for texture interpretation. Malik and Perona [76] implement difference of offset Gaussians (DOOG) and comment that the selection of filters is not a critical choice as long as they have a good fit with the physiological measurements and they are computationally simple. Unser [105] uses a discrete wavelet frame (DWF) that he claims to be superior to standard wavelet transform feature extraction and this suggests that results are

better than cooccurrence probabilities. The wavelet decomposition he performs is quite successful for both classification and segmentation examples. Coggins and Jain [23] create ad hoc frequency domain filters that are based on frequency and orientation specific functions. Bello [9] uses a Haar wavelet to decompose images for texture segmentation purposes. Quadrature mirror filters (QMF), often used in image coding, can also be applied to the texture identification problem. Based on energies in the subspaces of the filtered images, Chen and Kundu [20] classify sets of Brodatz imagery. Krishnamachari and Chellappa [66] have designed a wavelet filter bank whose output will always be a Gaussian Markov random field (GMRF) given that the input is a GMRF. Using a supervised approach, they segment a 4-class Brodatz image that contains textural regions with a number of different shapes.

4.2 Pattern Classification

For the purposes of this thesis, traditional pattern classification techniques will be utilized. Only a brief discussion of the essential aspects are presented here.

4.2.1 Feature Reduction

The reduction of dimensionality (removal of redundancy) of feature space is an important part of a feature generation process. The feature selection process should provide a data set with only the essential information necessary for accurate classification. Computer implementations with high feature dimensions are typically unwieldly. Visualization of the feature space is awkward when the number of dimensions exceeds three. With increasing dimensionality of the texture feature space there is initially a corresponding increase in the classification accuracy [81]. Eventu-

ally, at some unknown point, the texture classification accuracy begins to decline. Prediction of the total number and type of features is difficult, so reducing the dimensionality given a full feature set is usually performed.

A popular method for feature reduction is the orthonormal whitening transformation (also referred to as the Hotelling transform, principal components analysis, or the Karhunen-Loeve transform). By removing the correlations among the feature dimensions, this method is used to find a lower dimensional representation. Multi-spectral imagery is often reduced using this approach [60].

4.2.2 Supervised Classification

There are many methods to perform supervised classification. The method of choice here is the Fisher linear discriminant (FLD) [43]. The FLD provides a method to reduce a multidimensional feature space down to a d-dimensional feature space. This is accomplished by finding the optimal projection to separate $d+1$ classes. We are typically concerned with setting $d = 1$ so that a 1-d maximum a priori (MAP) classifier can be utilized. In the case of C classes, $\begin{pmatrix} C \\ 2 \end{pmatrix}$ pairwise class discriminants are determined. The discriminant vector (w) is found by optimizing the Fisher criteria:

$$J(w) = \frac{w^T S_B w}{w^T S_W w} \tag{4.3}$$

where S_B and S_W are the between-class and within-class scatter matrices. In the $d = 1$ case, S_B and S_W are defined by:

$$S_B = (m_1 - m_2)(m_1 - m_2)^T$$

and

$$S_W = S_1 + S_2$$

where

$$S_i = \sum_{x \epsilon C_i} (x - m_i)(x - m_i)^T$$

and m_i represents the class mean.

By taking the derivative of $J(w)$, the optimal w is determined to be:

$$w = S_W^{-1}(m_1 - m_2).$$

Further discussion on the FLD is found in texts by Duda and Hart [33] and Schalkoff [95]. A simple improvement is to use covariance matrices in place of scatter matrices (simply divide each scatter matrix by the number of samples in the class). Using covariance matrices corrects for unequal class sizes:

$$S_i = \frac{1}{N_i} \sum_{x \epsilon C_i} (x - m_i)(x - m_i)^T$$

.

In order to classify a sample when more than two classes exist, the sample is classified according to each class pairwise discriminant. The class that is selected most often is the class to which the sample is assigned. The advantages of using this classifier include: low computational load, optimal reduction of a multidimensional space to a 1-dimensional space, and inherent normalization of the distance measures between the classes regardless of the scaling of the feature dimensions.

Also, normalized measurements of inter-class distances based on the Fisher criteria can be used to assist determination of preferred features.

By using pairwise discriminants in a multi-dimensional environment, it is possible for a sample to not have a definite class assignment when the number of classes is greater than two. For example, consider the case of three classes (A,B,C). If the sample x is assigned to class A for discriminant $A - B$, and class B for discriminant $B - C$, and to class C for discriminant $A - C$, then the class assignment is unresolved. For the classification studies, such a sample is considered erroneous. For the segmentation studies, unresolved pixels are typically found on the boundaries separating different assigned textures. These pixels are assigned to the class according to its nearest assigned pixel.

4.2.3 Unsupervised Classification

Clustering algorithms identify densely populated areas in the feature space. Given multidimensional vectors, the goal is to find the best partition of the feature space into these significant clusters. Duda and Hart [33] and Jain and Dubes [58] provide a detailed look at classical clustering techniques. Methods important to the implementations in Chapter 6 are reviewed here.

A common clustering approach is the K-means algorithm which iteratively assigns classes to the nearest class mean using the minimum Euclidean distance (MED) classifier. After each iteration, new class means are determined based on the most recent class assignments. A convenient stopping criteria is to determine when all samples cease switching to another class. The ISODATA algorithm extends the K-means approach providing user defined mechanisms for splitting and merging clusters. For both methods, scaling of each of the feature dimensions could affect

the success of the algorithms. Note that the K-means algorithm requires that the user indicate the number of classes and the ISODATA approach requires knowledge of the existing data set to produce an accurate clustering.

Hierarchical clustering techniques group/partition the feature space in an incremental fashion until the desired number of classes is reached. The grouping/partitioning is typically based on some distance measure between clusters. Two fundamental procedures are applied: divisive (begins with all samples in one cluster and adds classes until some criteria is met) and agglomerative (begins with each samples representing a different class and groups similar samples until some criteria is met).

Generally, determining the number of classes C is a difficult problem. Fukunaga [47] recognizes this and describes how, for a particular data set, different values of C can be used in conjunction with some optimization criteria to indicate when the number of classes is correct/incorrect. Methods to determine which clusterings are best are referred to as cluster validity indices by Jain and Dubes [58].

Chapter 5

Determining Preferred Texture Features

To attempt image texture segmentation, one must first determine preferred feature sets. "The development of computational formalisms for segmenting, discriminating, and recognizing image texture projected from visible surfaces are complex and interrelated problems. An important goal of any such formalism is the identification of easily computed and physically meaningful image features which can be used to effectively accomplish those tasks [12, p.2025]". In this chapter, three texture feature extraction approaches are studied to generate their individual preferred feature sets that will be used for texture segmentation. The emphasis of this chapter is to optimize the individual ability of each of the methods, not necessarily to compare them. Classification is the method of choice since calculations can be performed quickly (which provides the opportunity for a wide variety of tests) and quantitative error estimates can be determined. Features that perform better in a classification environment should have a stronger potential to perform better in an unsupervised

segmentation role. The three different texture analysis techniques analyzed are: the grey level cooccurrence features, the power spectrum features, and Gabor filter outputs. The final section of this chapter is devoted to summarizing the results with a look towards texture segmentation implementations.

Two data sets are used for the classification testing: textures obtained from a SAR aerial image (to gain an appreciation of the application environment) and textures obtained from the Brodatz album [14] (to observe the effects on a system with more yet better defined classes). The Limex SAR image (Figure 6.17, image details found on page 134), contains three different classes: brash ice, first year smooth ice, and open water. Sixty-four and 100 8x8 samples of each class are selected for training and testing, respectively. Brodatz imagery (illustrated in Appendix A) is undoubtedly the most common test imagery used in the texture interpretation literature. This imagery provides opportunity for testing using a variety of classes and for comparing results with other research. Also, the training and test image samples are assured to contain one class only; the same cannot hold absolutely true for the SAR image samples, although efforts are made to accomplish this. Eight different Brodatz textures are used: cloth (D19), cork (D4), cotton (D77), grass (D9), paper (D57), pigskin (D92), stone (D2), and wood (D68). (The 'D*' represents the numbering system assigned in the Brodatz album.) These textures are chosen since they have a noticeable but not necessarily regular textural pattern and several of the textures are similar in nature, somewhat mimicking textures found in SAR sea ice imagery. The cotton texture, used as a control, is the only one that has a well defined repeating pattern. Prior to classification, each 256x256 image is normalized to a full 256 grey levels. Training samples are selected by dividing the upper left hand quadrant into 64 16x16 images. Test samples are selected by dividing the bottom right hand quadrant into 64 16x16 images. It is interesting to note that

the 16x16 samples of Brodatz imagery are not easy to visually discriminate. This is also noted by McLean when performing classification on such imagery [82]. The class-pairwise Fisher linear discriminant is used for all classifications.

5.1 Cooccurrence Features

As described in Section 4.1.2, cooccurrence texture features are a popular method for classification of SAR sea ice imagery. Here, limitations of the cooccurrence texture feature extraction approach are described. Then, these implementational hurdles are overcome by doing the following:

- Instead of generating the cooccurrence features using the matrix approach (the GLCM), a linked list approach (grey level cooccurrence linked list or GLCLL) is described and implemented.
- Since many of the statistics generate redundant information, selection of pertinent statistics is performed. Reduction of the number of statistics is not only computationally favourable, but can assist the classification accuracy as well.

5.1.1 Limitations of the Matrix Approach

Although the cooccurrence method has been widely applied to remote sensing image interpretation, there are implementation restrictions of the approach as described by Haralick [52]. It is interesting to note that Haralick originally referred to the generation of GLCM features as "quickly computable" (p.618), however, the excessive computational demand is one reason why this method has been restricted

from implementation in operational settings. Another restriction is that the co-occurrence features have not demonstrated robustness across different scenes and images from different sensors.

Typically, a GLCM can be quite sparse. If the full dynamic range of a typical image is used, then each GLCM is 256x256 pixels (65536 entries). If the window size is 20x20, then at most (20)(19) = 380 pairs are possible, producing a very sparse matrix indeed! Applying statistics involves looping through each of the GLCMs, a very costly procedure given that most of the matrix entries are zero. In practice, these computational demands are reduced in a number of ways: quantizing the grey levels, limiting the number of features, and avoiding pixel-by-pixel segmentation.

- The image grey levels are quantized prior to GLCM implementation, typically from eight bits down to as few as four or five bits (Table 4.4). This reduces the size of the GLCMs and causes a dramatic decrease in computational time. Quantization has the potential to remove pertinent information from the image. What happens to the cooccurrence features if the full dynamic range is used?

- The number of statistics and/or the number of orientations for each cooccurrence matrix must be limited so that the features are calculated within a reasonable duration. Deciding which statistics are the most informative for remotely sensed imagery has been the focus of research efforts (Section 4.1.2). Some investigators search for the "best" statistics without considering what specific texture characteristics each is measuring [4, 83]. Shokr [98] and Baraldi and Parmiggiani [2] investigate the textural meaning of each statistic and give recommendations based on this analysis.

- Even with the above load reductions, performing a pixel-by-pixel segmen-

tation of an image is impractical for the GLCM approach, although some researchers have implemented it in this fashion [46, 55]. Non-overlapping windows generate texture features that are representative of the whole window compared to overlapping windows whose texture features are representative of the centre pixel. Thus, the resolution of the non-overlapping windows is the size of the windows, but the resolution of the fully overlapping windows is a single pixel. Baraldi and Parmiggiani [2] use non-overlapping windows and the result is a segmentation that is not visually realistic because boundaries between different textures are "blocky". Little research has been published to ascertain the ability of the cooccurrence features to perform pixel-by-pixel image analysis, an approach necessary to generate the detailed segmentation required for SAR sea ice imagery.

Unser [104] improved on some of these shortcomings by using sum and difference histograms as a substitute for GLCMs. Some of the features determined by Unser's method are quantitatively different than the GLCM features, however, they are supposedly similar in classification ability and require less computation time and memory storage to calculate. Like the GLCM, this method allocates memory for cooccurrences with zero probability.

5.1.2 Linked List Implementation (GLCLLs)

The subject of sparse matrices arises in many diverse fields, often for the purposes of solving linear systems of equations [34]. One method to improve sparse matrix performance is to use a linked list approach. No other linked list implementation for the generation of cooccurrence probabilities has been found in the research literature, but such an approach has been suggested [98]. Using a grey level cooccurrence

linked list (GLCLL) is very efficient because, unlike the GLCM approach, it does not allocate storage for those grey level pairs that have zero probability. The texture features that are generated using GLCLLs are identical to those generated by the GLCM approach.

The linked lists are set up in the following manner. Each node of the linked list is a structure containing the two cooccurring grey levels, their probability of cooccurrence, and a link to the next node on the list. The linked list is kept sorted based on indices provided by the cooccurring grey levels (g_i, g_j). An example of such a sorted list would be {(1,2), (1,4), (1,5), (3,4) (3,5), (5,5), (6,7), ...}, where $(g_i \leq g_j)$. In order to include a new grey level pair (g_i, g_j) in a linked list, a search is performed. This is done by finding the first instance of g_i and then proceeding from that point to find g_j. If the pair is found, then its probability is incremented; otherwise, a new node is inserted at the location where the search expected to find the node for (g_i, g_j).

An unsorted linked list *could* have been implemented by simply searching simultaneously on both grey levels. If (g_i, g_j) is not found, a new node would be added to the end of the list. However, maintaining a sorted list is more efficient because only one integer comparison for each node is performed when searching and an exhaustive search is unnecessary in the case that a node for grey level pair does not exist. Since the length of the linked lists dictates the algorithm speed, using sorted linked lists is critical.

In the traditional GLCM approach, the matrix is symmetric, however, this is detrimental to the computational advantages of the linked list approach. For example, given a grey level pair (5,10), the cooccurrence matrix would update the probabilities at (5,10) and (10,5) ("symmetric" implementation), however, the linked list enters a probability for the pair (5,10) only ("asymmetric" implementa-

tion; $g_i \leq g_j$). The shorter the linked list, the faster the feature extraction. For a feature such as COR which inherently requires symmetry, the feature extraction routines can compensate and calculate features as though symmetry existed. The features MAX, UNI, and ENT change their values under the asymmetrical linked list approach, however, they still represent the same type of feature and, if desired, they can be calculated in a manner that assumes a symmetric cooccurrence matrix. Comparative testing has indicated that the symmetrical calculation for MAX, UNI, and ENT provides better classification accuracy than the asymmetrical calculation. Thus, the symmetrical feature calculation will be used throughout this thesis, although the probabilities are still stored in an asymmetrical format.

When pixel-by-pixel features are required from an image, the GLCLLs can be implemented in the following manner. The GLCLLs are first created when the window is at the top left hand corner of the image. After the features are calculated, the window is moved one column to the right. Instead of recalculating entire GLCLLs, the current GLCLLs are updated to reflect the new information. The pairs of grey levels introduced by the new column are inserted into the GLCLLs. The pairs associated with the column that the window just passed over are subtracted from the GLCLLs. If the subtraction causes the grey level pair to have a zero probability, then that particular pair should not exist in the current linked list and that node is removed.

When the window reaches the end of the image row, it just slides down a single row. Here, it updates the GLCLLs by including the new pairs from the row that the window has just moved on top of and subtracting the pairs from the row that the window has just moved beyond (in the same manner as updating a column). The window then moves towards the opposite edge and continues to move in this zig-zag fashion until the entire image has been covered. This method for generating

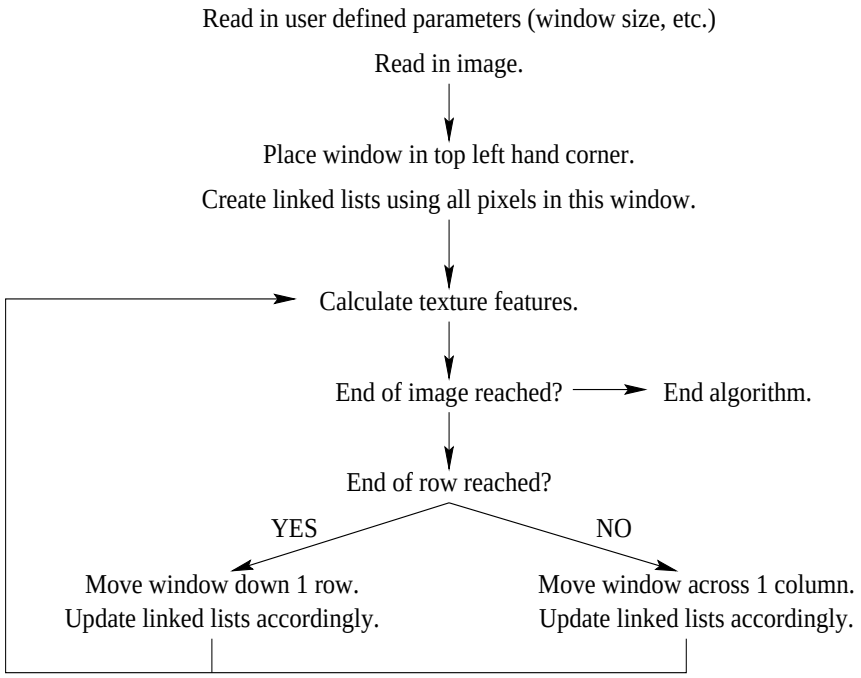


Figure 5.1: Algorithm for linked list generation of cooccurrence data.

the cooccurrence information will be referred to as "updating". This algorithm is summarized in Figure 5.1. Franklin and Peddle used a similar updating scheme to determine GLCMs [46].

5.1.3 Computational Efficiency of GLCLLs

Four different scenarios are compared from both a theoretical and applied perspective. These scenarios are: (1) the traditional GLCM without updating, (2) the GLCM with updating, (3) the GLCLL approach without updating, and (4) the GLCLL approach with updating.

Comparisons of the computational complexities are presented in Table 5.1. For each texture feature, whether GLCM or GLCLL based, computational speeds are

Case	Determine Probabilities	Determine Features
GLCM w/o updating	n^2	$s\rho G^2$
GLCM w/ updating	n	$s\rho G^2$
GLCLL w/o updating	$n^2 L$	$s\rho L$
GLCLL w/ updating	nL	$s\rho L$

Table 5.1: Computational requirements of cooccurrence texture features.

dependent on the window dimension (n), the number of statistics (s), and the cost of determining each statistic (ρ). Since the calculation of most statistics involve a different but approximately similar number of operations, only a generic cost of ρ is assumed. In addition, the GLCM methods are dependent on the number of grey levels (G) and the GLCLLs are dependent on the length of the linked lists (L) which is equal to the total number of distinct grey level pairs found in the window. The value of L is not only dependent on G but also on texture characteristics. The computational requirements are split into two aspects: the generation of the cooccurrence probabilities and the calculation of the statistics based on the cooccurrence probabilities. In the case of the traditional GLCM approach, generating each GLCM is dependent on the number of pairs in the window (n^2). Then, each GLCM must be looped through once to generate each statistic ($s\rho G^2$). In the other extreme, using an updated linked list requires updating the columns (n) and each update requires an average linked list search of L. A total of ($s\rho L$) operations is required to calculate the statistics for each GLCLL. The other cases are easily derived from these examples.

In order to compare the different scenarios, a 32x32 image of brash ice is extracted from the Limex image. Window sizes of 5, 10, and 20 and quantized grey levels of 32, 64, 128, and the full dynamic range are used to determine the cooc-

currence data. In the case of GLCM approaches, the full dynamic range is set to a fixed value of 256 grey levels. A total of 28 texture features are determined based on the set $\{\delta = 1;\ \theta = 0, 45, 90, 135$ degrees; statistics = MAX, UNI, ENT, DIS, CON, INV, IDM$\}$. This set deliberately excludes the COR statistic which has a different theoretical order than the other statistics. Testing is performed using an IBM RISC System/6000 Model 43P (64Mb RAM, 100MHz, SPECint92 - 128.1, SPECint92 - 120.2).

The increase in speed is impressive, as presented in Table 5.2 where the time per window sample (in μs) is recorded for each test case. The time per sample of the linked list approaches are always a fraction of the matrix approaches. For example, given 128 grey levels and a window size of 5, the traditional GLCM approach requires 330 μs per window to calculate the same set of features that the GLCLL with updating takes 0.65 μs per window to determine. For an image the size of a typical remotely sensed image, the computational savings are substantial.

The results of the matrix approaches match the theoretical orders well. At a fixed window size, doubling the number of grey levels increases the completion time of the GLCM approaches by a factor of four. Calculating the GLCM with updating only improves the computational speed slightly, thus the computational speed of the GLCM approaches is, as expected, highly dependent on determining the statistics ($s\rho G^2$). Changing the window size for the GLCM approaches has little effect on the computational requirements.

In contrast, increasing the window size for a fixed grey level increases the GLCLL completion time because the larger windows have more distinct grey level pairs and this increases the length of the linked lists (L). Reducing the number of grey levels reduces the completion times of the GLCLL approach since quantized grey levels shorten the linked lists. The symmetrical sorted lists (ie. $g_i \leq g_j$) have

Window Size	Number of Grey Levels	GLCM w/o update	GLCM w/ update	GLCLL w/o update	GLCLL w/ update
5	full	1300	1300	1.0	0.63
10	full	1300	1300	5.7	2.7
20	full	1300	1300	44	15
5	128	330	330	0.98	0.65
10	128	330	330	5.4	2.6
20	128	340	330	39	14
5	64	83	82	0.97	0.64
10	64	84	83	4.8	2.4
20	64	86	84	29	9.1
5	32	21	21	0.92	0.60
10	32	22	21	3.6	1.7
20	32	23	22	15	4.1

Table 5.2: Completion times (μ-seconds per window sample) to calculate statistics.

faster completion times than asymmetrical sorted lists. Results for the same test set using symmetrically sorted linked lists calculated on a DEC station 5000 are presented in [22]. Finally, updating the GLCLLs is advantageous, especially with larger window sizes.

5.1.4 Cooccurrence Parameter and Statistic Selection

Parameter Selection

There are a number of different parameters that must be indicated in order to generate cooccurrence data: window size (n), orientation (θ), pixel separation distance (δ), and the number of quantized grey levels (G). Limiting the number of parameters will prevent generation of an unnecessarily high dimensional feature

space.

For classification studies, the window size is often fixed and assumes that sufficient area of the texture is represented to capture an appropriate measurement. Typically, θ is set to {0, 45, 90, 135 degrees} since this is easiest to implement. To program the generation of cooccurrence data for arbitrary orientations is not reasonable. Short pixel separation distances have typically achieved the best success (Section 4.1.2) so $\delta = 1$ will always be used. Various values of δ can yield different features, however, it is unknown if there is any *a priori* method for selecting its value based on the given image characteristics. Since the ultimate goal is to develop segmentation methods that are unsupervised, δ will be kept at its "best" guess of 1. The rest of this section will investigate reduction of the number of statistics by removing redundant information. Simultaneously, the effect of varying grey level quantization will be studied.

Results for Entire Feature Set

Thirty-two texture features are selected based on the set $\{\delta = 1;\ \theta = 0, 45, 90, 135$ degrees; statistics = MAX, UNI, ENT, DIS, CON, INV, IDM, COR$\}$. In order to determine the effect of grey level quantization on the classification, 128, 64, 32, and 16 grey levels are used as well as the full dynamic range. The classification accuracy of both the Brodatz and Limex testing is presented in Table 5.3. Both sets of training data have classifications that are quite successful and consistent across all grey levels. Classification of test data is strongest at full dynamic range and decreases inconsistently with increased quantization. For all cases, there is a large discrepancy between the classification accuracies for training and test data. Limex imagery has a much higher classification rate then the Brodatz imagery, probably because the Limex data set has only three classes compared to the Brodatz eight

	Brodatz		Limex	
Grey Levels	Training	Test	Training	Test
full	98.6	87.5	99.5	90.7
128	97.7	86.5	99.0	84.7
64	97.1	83.8	99.5	83.3
32	96.7	86.7	99.5	87.0
16	97.9	82.8	98.4	87.3

Table 5.3: Classification accuracy (%) using all cooccurrence features.

classes.

Results for Individual Statistics

In order to provide further insight into the ability of the cooccurrence data each individual statistic is used to classify both data sets. The results are presented in Table 5.4 (Brodatz) and Table 5.5 (Limex) given $\{\delta = 1;\ \theta = 0, 45, 90, 135$ degrees$\}$. The system is trained using the selected statistic with the training data and classification accuracies are only presented for classification of the test data. For both data sets, the degree of grey level quantization had an unusual effect on the homogeneity statistics {MAX, ENT, UNI}; namely, they have significantly increasing classification accuracy with coarser quantization. One would expect increasing the grey level quantization would reduce the textural information and thus, reduce the classification accuracy. However, using the full dynamic range, few grey level pairs are repeated within the same window and a high state of entropy exists in the cooccurrence data for each of the classes. As a result, discrimination is difficult since all classes tend to a near maximum state of entropy generating clusters that overlap in the feature space (see MAX feature axis in Figure 5.2). Thus, these statistics are intrinsically sensitive to grey level quantization and actually rely on

Grey Levels	MAX	UNI	ENT	DIS	CON	INV	IDM	COR
full	29.7	53.5	55.9	81.5	80.1	80.9	80.3	77.0
128	36.7	57.8	61.3	81.5	80.1	81.3	80.5	76.4
64	37.1	66.2	69.3	81.5	79.9	82.0	80.3	74.6
32	43.6	69.5	75.4	82.0	79.9	81.5	80.1	72.5
16	48.6	70.7	74.6	81.6	81.3	81.1	81.5	69.1

Table 5.4: Classification accuracy (%) of individual cooccurrence statistics for Brodatz test imagery.

Grey Levels	MAX	UNI	ENT	DIS	CON	INV	IDM	COR
full	58.0	74.3	75.7	91.3	90.7	91.3	90.7	52.7
128	69.0	80.0	82.0	91.3	90.7	91.3	91.0	55.7
64	79.3	88.3	89.0	91.0	90.7	91.0	91.0	54.0
32	84.0	88.7	90.7	91.0	89.0	91.3	88.7	59.3
16	86.0	90.3	90.7	91.3	84.0	90.7	85.3	68.3

Table 5.5: Classification accuracy (%) of individual cooccurrence statistics for Limex test imagery.

the quantization to be effective.

The smoothness statistics {DIS, CON, INV, IDM} have stronger classifications than the homogeneity statistics. DIS and INV generate consistent and accurate results across both data sets. CON and IDM, the squared distance measures, maintain a steady classification accuracy for the Brodatz imagery across all grey levels but decrease their ability to distinguish Limex imagery with increased quantization. Quantization tends to smooth the Limex data preventing the smoothness statistics from performing optimally. The Brodatz textures are more structured than the Limex textures. Thus, quantization does not significantly alter the general appearance of the Brodatz textures and the smoothness statistics are able to maintain a consistent classification accuracy. Note that the individual accuracies of

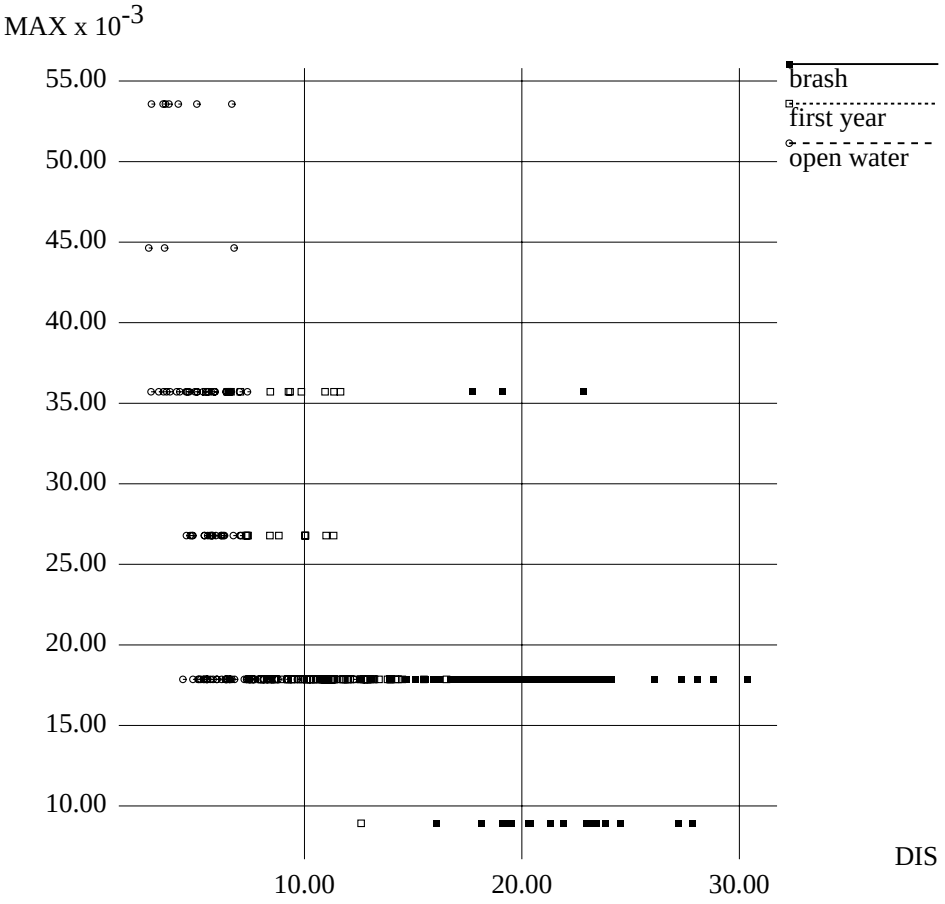


Figure 5.2: Texture feature plot of DIS versus MAX ($\delta = 1$; $\theta = 0°$; Limex data).

the DIS and INV statistics for the Limex data are better than the accuracies using the entire feature set.

In the Limex data set, COR is quite ineffective when compared to the other statistics, but tends to improve with coarser quantization. On the other hand, COR is more successful classifying the more difficult Brodatz imagery, probably due to the relatively higher regularity found in these images.

Selection of Redundant Statistics

There is evidence to indicate that certain cooccurrence texture features are highly correlated [4, 98]. Given $\{\delta = 1; \theta = 0°;$ G=16$\}$, the correlations of the statistics are presented in Table 5.6 for brash ice and Table 5.7 for the cork texture. The high correlation between the statistics {DIS, CON, INV, IDM} is expected since they are all functionally similar. An example of this high correlation in the feature space is given in Figure 5.3, where DIS and INV are compared for all three Limex classes. The statistics {MAX, UNI, ENT} also display highly correlated behaviour. A supporting example is found in Barber *et al.* [8, Figure 9] where UNI and ENT are plotted against each other and display a strong inverse correlation.

For the brash ice, the correlation statistic (COR) has a fairly strong correlation with the homogeneity statistics, but is poorly correlated with the smoothness statistics. The cork texture, on the other hand, has a stronger correlation between COR and the smoothness statistics. For the Brodatz textures, the COR statistic has a higher classification rate. This supports the notion that cooccurrence features extract one true characteristic of the texture and this characteristic is a function of the directional smoothness. Between the homogeneity and smoothness sets there are strong correlations (all >0.44). This is expected due to the nature of the imagery: smooth textures have fewer grey level pairs than rough textures, causing the cooccurrence probabilities for the smooth textures to have a lower entropy than the rough textures. Thus, the smoothness statistics and the homogeneity statistics are well correlated.

Given the same orientation and pixel distance, the cooccurrence features display a very high correlation for any number of grey levels, but definitely become increasingly correlated with coarser quantization, as indicated by the average inter-

	MAX	UNI	ENT	DIS	CON	INV	IDM	COR
MAX	1.00	0.79	-0.68	-0.55	-0.44	0.57	0.45	-0.28
UNI	-	1.00	-0.95	-0.75	-0.66	0.76	0.68	-0.29
ENT	-	-	1.00	0.84	0.80	-0.83	-0.81	0.31
DIS	-	-	-	1.00	0.95	-1.00	-0.95	-0.05
CON	-	-	-	-	1.00	-0.92	-1.00	-0.05
INV	-	-	-	-	-	1.00	0.93	0.05
IDM	-	-	-	-	-	-	1.00	0.05
COR	-	-	-	-	-	-	-	1.00

Table 5.6: Correlation of features for brash ice (θ=0 degrees; δ=1; G=16).

	MAX	UNI	ENT	DIS	CON	INV	IDM	COR
MAX	1.00	0.88	-0.79	-0.62	-0.54	0.66	0.66	-0.06
UNI	-	1.00	-0.95	-0.73	-0.66	0.74	0.74	-0.23
ENT	-	-	1.00	0.83	0.82	-0.79	-0.80	0.14
DIS	-	-	-	1.00	0.97	-0.96	-0.97	-0.33
CON	-	-	-	-	1.00	-0.86	-0.81	-0.34
INV	-	-	-	-	-	1.00	1.00	0.31
IDM	-	-	-	-	-	-	1.00	0.31
COR	-	-	-	-	-	-	-	1.00

Table 5.7: Correlation of features for cork (θ=0 degrees; δ=1; G=16).

correlation across all selected texture features for each class in Tables 5.8 (Brodatz) and 5.9 (Limex).

The high correlation of the entire coccurrence data set can be analysed using principal components analysis. The average class variances and ranges for the first principal component are presented in Table 5.10. Given that there are 40 different features, yet the first principal component captures most of the variance for each data set, it is obvious that a tremendous amount of redundancy exists in the cooccurrence texture features.

Grey Levels	Cloth (D19)	Cork (D4)	Cotton (D77)	Grass (D9)	Paper (D57)	Pigskin (D92)	Stone (D2)	Wood (D68)
full	0.55	0.49	0.30	0.51	0.37	0.53	0.47	0.50
128	0.59	0.49	0.30	0.52	0.37	0.53	0.49	0.50
64	0.62	0.53	0.34	0.55	0.41	0.53	0.51	0.54
32	0.64	0.56	0.39	0.57	0.45	0.59	0.56	0.56
16	0.67	0.58	0.42	0.59	0.46	0.62	0.59	0.57

Table 5.8: Average inter-feature correlations for Brodatz cooccurrence features.

Grey Levels	Brash	First Year Smooth	Open Water
full	0.33	0.33	0.30
128	0.38	0.37	0.37
64	0.46	0.47	0.42
32	0.54	0.51	0.48
16	0.59	0.57	0.56

Table 5.9: Average inter-feature correlations for Limex cooccurrence features.

	Brodatz		Limex	
Grey Levels	Average	Range	Average	Range
16	60	43-70	61	59-64
256	49	29-59	35	31-39

Table 5.10: Average and ranges of class variances (%) for the first principal component.

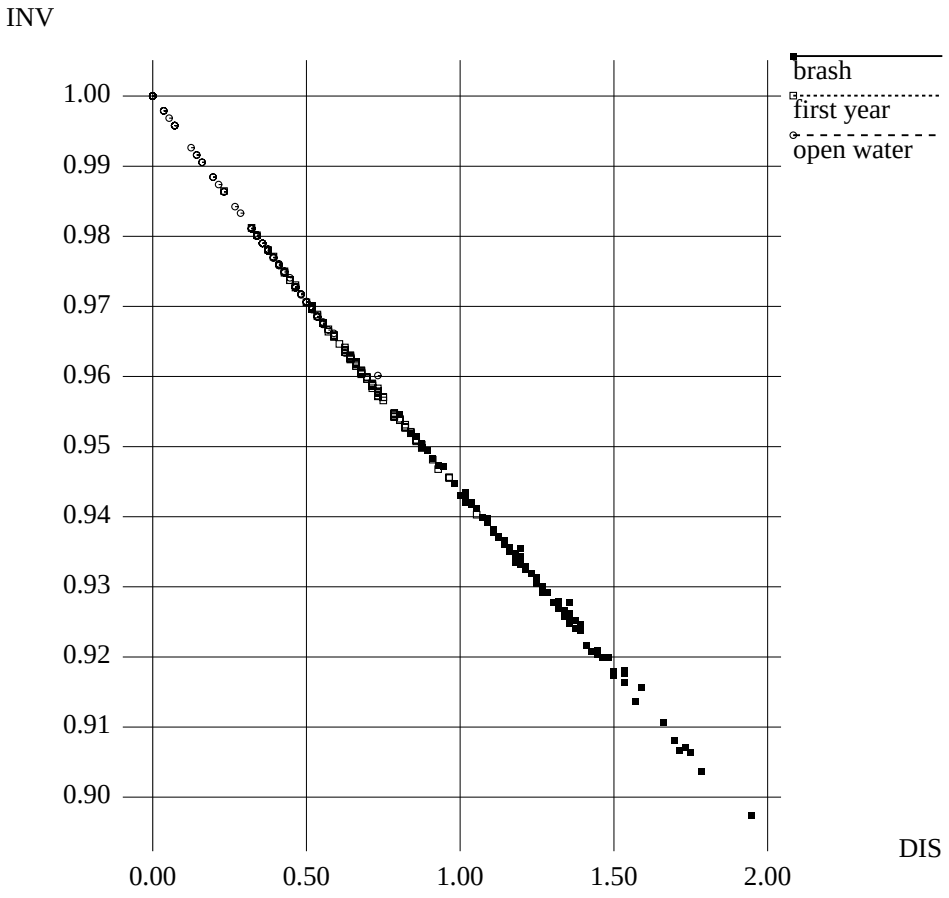


Figure 5.3: Texture feature plot of DIS and INV ($\delta = 1$; $\theta = 0°$; Limex data).

Of all the homogeneity statistics, ENT generates the best overall accuracy with the optimum occurring at 16 grey levels. Of all the smoothness statistics, DIS and INV generate the best classification, but the optimum occurs at full dynamic range. There is a very strong correlation between the ENT features (16 grey levels) and DIS features (256 grey levels) given the same orientation (Tables 5.11 and 5.12). All the entries in these tables demonstrate that a very strong correlation exists between the feature pairs. When the correlation reaches a minimum ('wood' at 90° is 0.49 and

θ	Cloth (D19)	Cork (D4)	Cotton (D77)	Grass (D9)	Paper (D57)	Pigskin (D92)	Stone (D2)	Wood (D68)
0°	0.89	0.84	0.79	0.87	0.78	0.84	0.81	0.83
45°	0.77	0.85	0.67	0.83	0.75	0.83	0.84	0.77
90°	0.70	0.85	0.74	0.82	0.75	0.82	0.74	0.49
135°	0.81	0.85	0.79	0.87	0.74	0.84	0.76	0.80

Table 5.11: Correlations between DIS (full dynamic range) and ENT (16 grey levels) for Brodatz imagery.

θ	Brash	First Year Smooth	Open Water
0°	0.83	0.74	0.66
45°	0.77	0.77	0.54
90°	0.80	0.69	0.40
135°	0.84	0.62	0.53

Table 5.12: Correlations between DIS (full dynamic range) and ENT (16 grey levels) for Limex imagery.

'open water' at 90° is 0.40), this is representative of the texture features not finding some measurable characteristic of the texture. There is no repeating characteristic in the wood texture oriented at 90°. Open water has a dominant characteristic in the 0° direction (due to the wind causing a ripple effect) and no dominant characteristic in the 90° direction. Although it is generally believed that smoothness and homogeneity statistics measure different textural characteristics, this strong correlation between DIS and ENT support the concept that these statistics actually measure the same dominant characteristic of texture.

5.2 Power Spectrum Features

Power spectrum features were introduced in Section 4.1.4. Due to the difference in sample image sizes, different texture feature sets are derived for each of the Brodatz and Limex data sets. The Limex imagery uses peak statistics for the highest peak (Peak1) {PCT1, LAP1, SKW1, DIS1}, shape statistics {ISO, SPD, CIR}, and summed features defined by the intersection of the angular regions $\{0-44°, 45-89°, 90-134°, 135-179°\}$ and pixel radial distances {[1,2), [2,4)}. Since the Brodatz imagery has a larger window dimension (16x16 compared to 8x8), a second peak may be relevant, thus, this imagery is analyzed using both the highest and second highest peak (AllPeak) {PCT1, LAP1, SKW1, DIS1, PCT2, LAP2, SKW2, DIS2, ANG}, the shape statistics {ISO, SPD, CIR}, and summed features based on regions defined by the angular regions $\{0-44°, 45-89°, 90-134°, 135-179°\}$ and radial regions {[2,4),[4,8)} in pixel units away from the centre of the power spectrum.

The computational complexity for the power spectrum features can be determined in the following manner. The FFT operation requires approximately $n^2 \log_2 n$ operations [48]. Calculating the power spectrum requires an additional n^2 multiplications. For each statistic, the calculation of the statistics is negligible compared to determining the power spectrum. Thus, the total complexity may be approximated by $n^2 \log_2 n + n^2$.

The results of the classification testing are presented in Table 5.13. These results are not, in general, as successful as the cooccurrence results (Table 5.3) for both sets of data. Only classification of the Limex test imagery is more successful for the power spectrum features compared to the cooccurrence features. Note that classification rates for the power spectrum features remain fairly constant across dif-

	Brodatz		Limex	
Grey Levels	Training	Test	Training	Test
full	86.5	77.2	97.9	90.7
128	86.7	77.7	97.4	89.7
64	87.1	76.2	97.4	90.3
32	88.5	74.8	96.4	90.7
16	87.1	74.6	95.8	85.7

Table 5.13: Classification accuracy (%) for all power spectrum features.

ferent grey level quantizations. This is expected because relative spatial-frequency data should not be altered unless the grey level quantization is to such an extent that it destroys salient texture characteristics causing different classes to appear similar. The effect of significant grey level quantization might be the cause of the Limex classification rate dropping dramatically from around 90.7% success to 85.7% when quantized to 16 grey levels. The Brodatz data undergoes a gradual decline in classification success with increased quantization. Here, the different textures are independently sensitive to the number of grey levels so the gradual decline is expected. Also, with increased grey level quantization, the correlation between the features stays consistently low (Tables 5.14 and 5.15), unlike the cooccurrence features.

How do the individual types of power spectrum features fare (Tables 5.16 and 5.17)? Summed features perform the best for both data sets. Shape features performed well for the Limex data but not for the Brodatz imagery. The peak features for the Limex case seem to be strong, however, following further investigation, it was discovered that the Laplacian of the peak is the only strong feature. The success of this feature is a result of the energy in the spectrum, not the energy of the peak, since the sample size is so small. Peak statistics are expected to be

Grey Levels	Cloth (D19)	Cork (D4)	Cotton (D77)	Grass (D9)	Paper (D57)	Pigskin (D92)	Stone (D2)	Wood (D68)
full	0.21	0.15	0.17	0.15	0.20	0.17	0.22	0.23
128	0.21	0.16	0.17	0.16	0.19	0.18	0.22	0.23
64	0.22	0.16	0.17	0.16	0.21	0.18	0.21	0.24
32	0.21	0.16	0.17	0.17	0.21	0.18	0.22	0.24
16	0.21	0.17	0.17	0.17	0.21	0.17	0.21	0.21

Table 5.14: Average inter-feature correlations for Brodatz data using power spectrum features.

Grey Levels	Brash	First Year Smooth	Open Water
full	0.15	0.16	0.14
128	0.15	0.16	0.14
64	0.15	0.16	0.16
32	0.16	0.17	0.15
16	0.17	0.17	0.15

Table 5.15: Average inter-feature correlations for Limex data using power spectrum features.

helpful in the Brodatz imagery, where regular repeating patterns are more apparent, however, they do not perform well using this imagery. This probably results from the window sizes (16x16) being too small for the autocorrelation function to capture sufficient periods of the regular repeating texture to make an accurate classification. D'Astous [27] agrees, implementing her algorithms on 32x32 windows of Brodatz imagery.

Grey Levels	Peak1	Peak2	AllPeak	Shape	[2,4)	[4,8)	[2,4),[4,8)
full	46.5	39.5	48.8	46.7	68.8	57.4	75.2
128	46.3	37.7	48.4	47.3	69.3	57.6	75.4
64	46.7	37.5	50.2	42.2	69.5	56.8	74.6
32	46.5	36.5	50.2	43.7	69.3	57.8	74.4
16	47.1	39.1	50.4	46.1	69.1	56.5	76.0

Table 5.16: Classification (%) of individual power spectrum features for Brodatz imagery.

Grey Levels	Peak1	Shape	[1,2)	[2,4)	[1,2),[2,4)
full	84.7	52.0	82.0	90.7	90.1
128	82.7	63.7	79.3	87.0	88.7
64	81.7	75.7	79.7	87.0	89.0
32	80.7	85.0	79.3	87.0	88.0
16	78.0	72.7	80.7	85.0	88.3

Table 5.17: Classification (%) of individual power spectrum features for Limex imagery.

5.3 Gabor Filter Features

Gabor filters were introduced in Chapter 4. Here, Gabor filters will be first studied for functional characterization (selection of orientations, centre frequencies, and bandwidths). Then, using these preferred filter parameters, a variety of feature extraction methods will attempt improvement of the feature set (Section 5.3.2).

Instead of using all 100 Limex test samples, only the first 64 are used. From an implementational perspective, it is easier to amalgamate the 64 8x8 samples into a single 64x64 image and perform the filtering on the entire image using the spatial-frequency domain. Features based on pixels located at centres of the 8x8 samples are used for the classification purposes. The feature vectors are dimensioned to the

number of Gabor filters used.

5.3.1 Functional Characterization

Jain and Farrokhnia [59] implemented Gabor filters for texture segmentation using an octave frequency bandwidth of one, an angular bandwidth (B_θ) of 45°, and angular spacing (S_θ) of 45°. Unit octave spacing matches the experimentally determined HVS ability [87], however, 45° angular bandwidths are not in agreement. Smaller bandwidths of about 30° are more characteristic of the HVS [56]. The filter implementation outlined in Section 4.1.4 will be used. To provide minimal overlap and still maintain reasonable coverage in the spatial-frequency domain, B_θ is set equal to S_θ. If B_θ is greater than S_θ redundancies can occur which could deteriorate the feature set. The Gabor filters can be set up using one of two methods: provide B_θ and determine σ_y or allow σ_y to equal σ_x and calculate B_θ.

Different filter configurations have been tested on the same data sets used for the cooccurrence and power spectrum features. Frequency radii (F) of $\{\sqrt{2}, 2\sqrt{2}, 4\sqrt{2}$ cpi$\}$ are used as centre frequencies to filter the Brodatz imagery and $\{\sqrt{2}, 2\sqrt{2}$ cpi$\}$ are used to filter the smaller Limex samples. Thus, when $S_\theta = 45°$, there are twelve features for the Brodatz imagery and eight features for the Limex imagery. For $S_\theta = 30°$, there are eighteen and twelve features respectively. For the lower frequency bands, some spatial filter overlap occurs with adjacent samples (since the individual samples are amalgamated into a larger image). Although this is not important in terms of evaluating the ability of the Gabor filters, it does restrict direct comparison with the cooccurrence and power spectrum results. The DC component is set to zero to prevent contribution of the mean local grey level. Since the Gabor filters capture information similar to that captured by the summed power

		Brodatz		Limex	
S_θ	Bandwidth	Train(%)	Test(%)	Train(%)	Test(%)
45°	$B_\theta = 45°$	72.9	68.2	84.4	86.5
30°	$B_\theta = 45°$	77.3	71.1	87.0	85.4
30°	$B_\theta = 30°$	81.6	77.5	91.2	87.5
45°	$\sigma_x = \sigma_y$	75.8	72.1	87.0	85.4
30°	$\sigma_x = \sigma_y$	79.7	74.6	88.5	87.5

Table 5.18: Classification accuracies (%) using Gabor filter outputs for different functional characterizations.

spectrum features, testing different grey level quantizations is not necessary. The classification results will decrease with increasing quantizations and the decrease is dictated by the discriminability of the textures at the quantized levels. Thus, using the full dynamic range should generate better classification accuracy without any effect on the computational speed.

Results for Gabor filter features based on different functional characterizations are displayed in Table 5.18. Better results are obtained with the finer orientation spacing of $S_\theta = 30°$. Note that the performance of the classification is poorer for $B_\theta = 45°$ than $B_\theta = 30°$ given the same spacing of $S_\theta = 30°$. One might consider that the additional overlap in the spatial-frequency would be conducive to improving the classification rate, however, here the redundancy actually decreases the effectiveness of the features. Marginal improvements are made by setting $B_\theta = S_\theta$ as opposed to setting $\sigma_x = \sigma_y$.

5.3.2 Feature Extraction of Filtered Outputs

Methods

There are a number of different ways to manipulate the Gabor filter outputs to attempt improving the performance of the texture features. These include:

- applying spatial smoothing,
- using the power spectrum,
- using only the real part,
- rectifying the complex output,
- creating moments based on the spatial-frequency plane,
- determining a measure of homogeneity, and
- implementing a syntactic approach to characterize the filter outputs.

Gaussian smoothing seems to be essential for improving the performance of Gabor filters for texture segmentation [13, 59]. Bovik uses Gaussian filters with the same Gaussian parameters as the corresponding filters, but with larger extents. The extent is controlled by γ. If $g(x, y)$ is the Gaussian function of the Gabor filter, ie.:

$$h(x, y) = g(x, y) \exp(2\pi j F x)$$

then the function that smoothes the filter's magnitude response is $g(\gamma x, \gamma y)$. Bovik recommended setting γ to $\frac{2}{3}$. The smaller the value of γ, the greater the smoothing. To match the smoothing of the Gabor outputs with the Gaussian of the Gabor

filter makes sense since this should yield results that are still spatially well localized. There exists a physiological reason for utilizing smoothing in that it mimics characteristics of the HVS. Hall and Hall [49] describe the existence of sustained channels in the visual system, indicating that the HVS not only considers pixels in the field of view, but also pixels in the vicinity.

Instead of attempting to improve all the filter configurations available in Table 5.18, only the most successful filter configuration $\{S_\theta = B_\theta = 30°\}$ will be post-processed. Post-processing results have been calculated for the next best case $\{S_\theta = 30°;\ \sigma_x = \sigma_y\ \}$ and these are always slightly poorer than post-processing $\{S_\theta = B_\theta = 30°\}$. These particular data sets may be better tuned to the filter configuration $\{S_\theta = B_\theta = 30°\}$. Another explanation is that the spatial bandwidth is greater in the y-direction for $S_\theta = 30°$ than for $S_\theta = 45°$. The additional spatial information might increase the discriminability.

Jain and Farrokhnia [59] implemented the filters using only the real part of the Gabor filter. No attempt was made to compare results based on the magnitude response. (Their segmentation approach is described in more detail in Chapter 6, page 110.) Full wave rectification (summing the absolute value of the real and imaginary responses) is another method that can be used to process the complex filter outputs.

Bigun and du Buf [10] use moments based on the squared magnitude Gabor filter responses. They refer to the squared magnitude response as the "local power spectrum". Geometric moments are defined by:

$$m_{pq} = \sum_{i,j} \omega_i^p \omega_j^q |g(\omega_i, \omega_j)|^2$$

where p and q are integers which represent the moment order. The authors use $p + q < 5$ to reduce their feature set from 30 filter outputs down to 15 features.

The parameters (i, j) index the power spectrum. The filter response is $g(\omega_i, \omega_j)$, defined in terms of Cartesian spatial frequency coordinates. Central moments are defined by:

$$\mu_{pq} = \sum_{i,j} (\omega_i - \omega_x)^p (\omega_j - \omega_y)^q |g(\omega_i, \omega_j)|^2$$

where $\omega_x = m_{10}/m_{00}$ and $\omega_y = m_{01}/m_{00}$. Bigun and du Buf provide no analysis comparing the Gabor filter magnitude responses with these moment features. Only central moments are employed in their segmentation studies, without any explanation for the preference. In this thesis, central and geometric moments generated by both magnitude and local power spectrum filter outputs are compared. When smoothing is used on the filter outputs, the moments are applied to the smoothed outputs.

According to psychovisual experiments by Rao and Lohse [90], there are three important factors that the HVS uses to identify texture: directionality, repetition, and complexity. Directionality, or orientation, is accounted for in both the cooccurrence and Gabor filter methods. Repetition, or identification of the dominant frequencies, is accounted for in the cooccurrence scheme by the smoothness statistics and by the nature of the Gabor filter outputs. In the case where textures are not very regular, the complexity (also referred to as regularity or homogeneity) of the texture may be an important issue. Cooccurrence statistics attempt to measure complexity using the homogeneity features, but, as demonstrated previously, their applicability is questionable. Irregular or complex textures are not as detectable as regular textures by the Gabor filters because of filter leakage caused by spatial texture variability. Another feature may be necessary to complement the existing feature set to address this complexity issue.

A homogeneous texture has a consistent appearance. In terms of the Gabor filter responses the strongest filter responses should be consistent. A non-homogeneous texture will have dominant Gabor responses that fluctuate. Two different methods of determining homogeneity will be considered. By determining the local spatial variance of the filter response one can determine the degree to which the filter response is changing. Those textures with low variances have a high degree of homogeneity and those textures with high variances will not be homogeneous. Another method to determine homogeneity involves applying multiple Gabor filters and incrementally increasing the spatial bandwidth (keeping F and θ constant). If the response is consistent, then the texture is homogeneous since it has the same response over a wider spatial extent. The slope of the responses can be used as a texture feature. The more horizontal the slope, the more consistent the filter response, and thus, the more homogeneous the texture.

Both the local variance and slope measures of homogeneity have been estimated based only on the filter output with the highest magnitude for each pixel. This filter has the frequency and orientation $\hat{F}$ and $\hat{\theta}$ that is assumed to dominate the response of the HVS and thus controls the regularity that we observe and interpret. The local variance estimate of homogeneity has been implemented using 3x3 windows on the magnitude plane for filter ($\hat{F}$,$\hat{\theta}$). The slope feature is determined by increasing the filter bandwidth by 5% twice. The slope of the three values (determined by linear regression) is used to represent the new feature, which augments the existing feature set created by the Gabor filter bank. Different percentages have been used and results are consistently similar.

The final method for manipulating the Gabor filter outputs is a syntactic approach. If all the filter outputs are ranked in order of magnitude, then one would expect that the ranking is consistent for the same texture. By assigning each filter

its own symbol (an integer value), a syntactic string is created. There are existing methods for string recognition that can be implemented for cluster analysis [102, Ch. 9]. Here, the strings representing the filter outputs are unique since they are all of the same length and each symbol is only represented once. A "distance" between two strings $A = a_1 a_2 \ldots a_M$ and $B = b_1 b_2 \ldots b_M$ can be determined by finding j for which $a_i = b_j$, $i = 1$ to M. Then the distance measure can be found by summing all of the differences of i and j:

$$\mathrm{d(A,B)} = \sum |i - j| \text{ when } a_i = b_j \text{ for } i = 1 \text{ to } M$$

The squared distance has also been used, but this did not significantly affect the overall results. This type of measure only compares distances between individual strings. As a result, the k-nearest neighbour classifier is implemented. The computational demands of the k-nearest neighbour routine are much higher than the demands of linear classifiers.

Results

Results are presented in Table 5.19. This table is split into two sections: one for unsmoothed filter outputs and the other for smoothed filter outputs. Note that the magnitude response of the no smoothing case is also presented in Table 5.18.

The smoothed responses definitely yield results that are substantially better than their unsmoothed counterparts. Results for the real component without smoothing are much poorer than the magnitude response without smoothing. It makes sense that the real unsmoothed version performs poorly since a matched real filter response generates sinusoidal variations in the spatial domain. The sinusoidal variations are not condusive to discriminating the texture classes. In contrast, the magnitude response is flat, generating a consistent measurement that assists the

	Brodatz		Limex	
	Train	Test	Train	Test
No smoothing				
Magnitude response	81.6	77.5	91.2	87.5
Real component only	70.7	59.6	83.9	82.8
Full wave rectification	81.6	75.6	91.7	88.5
Geometric moments	79.1	76.4	91.2	88.5
Central moments	66.2	58.2	92.7	86.5
Homogeneity (local variance)	82.6	77.7	94.8	88.0
Homogeneity (slope)	82.6	76.7	91.7	88.5
Smoothing ($\gamma = \frac{2}{3}$)				
Magnitude response	98.1	95.3	100	98.4
Real component only	98.2	95.3	100	99.5
Full wave rectification	98.1	94.5	100	98.4
Geometric moments	96.5	95.5	100	99.5
Central moments	87.5	76.6	100	92.7
Homogeneity (local variance)	96.3	92.6	100	98.4
Homogeneity (slope)	98.4	95.5	100	95.8

Table 5.19: Classification accuracies (%) of manipulating Gabor filter outputs using $S_\theta = B_\theta = 30°$ (excludes syntactic approach - see text).

classification. Smoothing with a wider Gaussian than the Gaussian in the Gabor filter smoothes the sinusoidal variation sufficiently to generate a comparable classification result. The full wave rectification results are similar to the magnitude response results.

Results for both geometric and central moments are not as successful as results using the magnitude response since classification accuracies are generally lower. The moment order was set to $p + q <= 3$, reducing the number of features to 10. Additional testing was performed for the Brodatz imagery using $p + q <= 4$ and similar results were obtained. Note that the presented results have been calculated

as a function of the magnitude response, not the local power spectrum response as prescribed by Bigun and du Buf. The classification results using the local power spectrum moments were also calculated and proved to be poorer than those for the magnitude moments of the Brodatz imagery and about the same or poorer for the Limex imagery. For brevity, only the results based on magnitude moments are presented. Although the moment features utilize a smaller dimensional feature space, their poorer classification ability does not seem to warrant their usage. Bigun and du Buf did generate successful segmentation, however, the Brodatz textures they used are quite distinct. It is also interesting to note that they did not mention anything about smoothing the filter outputs. They did use a novel feature reduction method that might have exploited some characteristic of the moment generated data.

Homogeneity results provide a minor increase in classification success. However, when Gabor homogeneity features are plotted, they do not reveal additional classification ability (Figure 5.4). Since cotton is a relatively regular texture, its slope homogeneity is expected to have a higher value than the other textures. Unfortunately, the slopes are steeper. This is because the highest filter response is strong and the stronger Gabor filter responses tend to have higher variability, generating steeper slopes. Perhaps "complexity" deals with the number of dominant multiple frequency/orientation components. These dominant frequency/orientation components may be spatially transient preventing accurate spatial-frequency identification due to filter leakage. Using this definition, the complexity would already be represented in the feature set of properly filtered Gabor outputs.

Note that the Gabor smoothed slope homogeneity features tend to improve the classification of the Brodatz imagery but decrease the accuracy of the Limex imagery, just like the cooccurrence homogeneity features. Therefore, these homo-

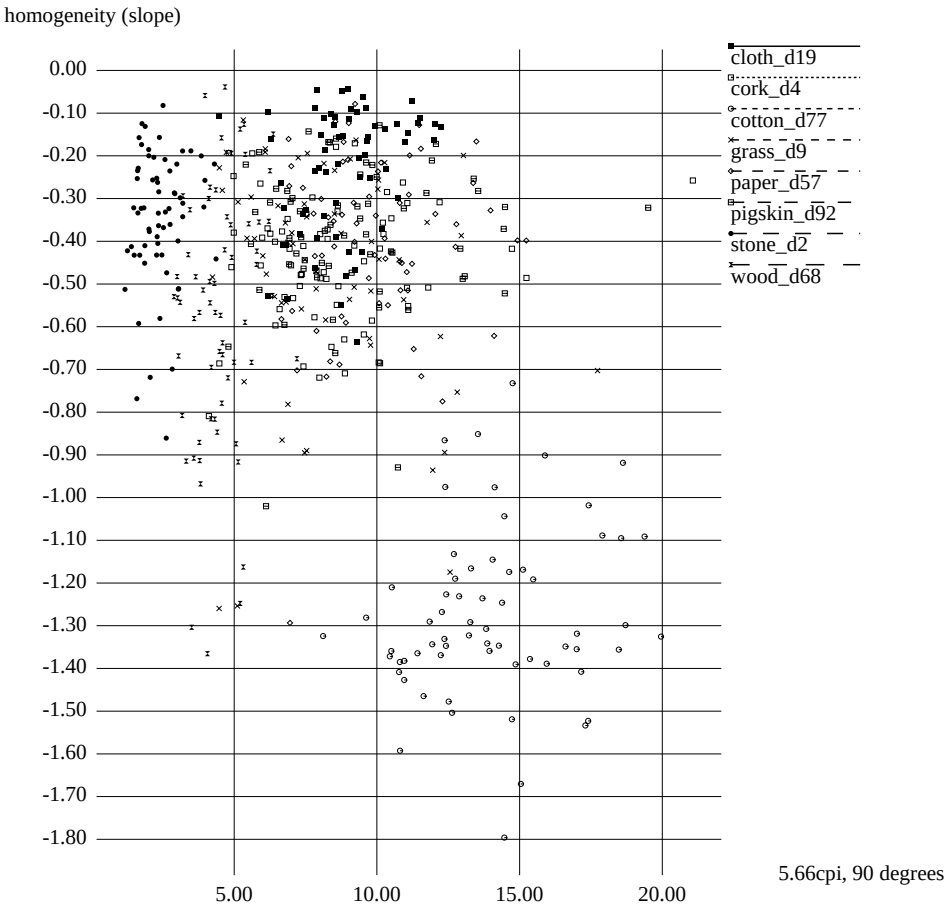


Figure 5.4: Inconsistency of Gabor homogeneity features for all Brodatz classes.

geneity statistics may be helpful for distinguishing textures containing significant structural information.

Results using the syntactic approach are poor ($< 60\%$ for Brodatz and $< 50\%$ for Limex; results only for smoothed training data sets). When the filter orientations are spaced by 45°, contrary to the other testing, classification accuracy of the Brodatz imagery increases, but the results are still below 80% (Limex results remain below 50%). For brevity, these results are not presented in Table 5.19. The syntactic

approach does not seem appropriate for textures that are not very distinguishable, however, in the case of distinct textures, they can be easily used to distinguish and code a texture in a very compact manner. For example, none of the cotton samples are misclassified using the syntactic technique.

Order of Operations

The computational complexity for implementing the complex Gabor filtering is summarized in Table 5.20. Feature extraction requires a 2-d FFT ($N^2 \log_2 N$ multiplications), where N represents the full image dimension. A total of s filters are applied. Recall that the image plane must be rotated to accommodate the various orientations of the filter (order R), so the Gabor filtering requires not sN^2 multiplications but actually sRN^2 multiplications. The inverse FFT must be performed once for each filtered image ($sN^2 \log_2 N$ multiplications) and then the magnitude envelope must be determined (sN^2 multiplications). Combining all of these sources generates a total of $sN^2 \log_2 N + RN^2$ multiplications. Smoothing of the magnitude envelope may be performed in the spatial domain or in the spatial-frequency domain. If performed in the spatial domain, the number of multiplications for a separable implementation is approximately sMN^2 where M represents the largest dimension of the mask. If performed in the spatial-frequency domain, the magnitude envelope is transformed ($sN^2 \log_2 N$ multiplications), then filtered (sN^2 multiplications), and then inverse transformed ($sN^2 \log_2 N$ multiplications). After combining these orders, the spatial-frequency filtering generates a total complexity of approximately $sN^2 \log_2 N$. Since M will have an average value of greater than 36 for a typical feature set (given $\pm 3\sigma$) and $\log_2 N^2$ is less than 20, it is better to perform the smoothing in the spatial-frequency domain. Also, mask truncation may occur during the spatial domain smoothing which could cause inaccurate calculations.

	Complexity (NxN image)
Gabor filtering	$sN^2 \log_2 N + sRN^2$
+ Spatial smoothing	sMN^2
+ Spatial-frequency smoothing	$sN^2 \log_2 N$

Table 5.20: Computational complexity of Gabor filter texture features.

5.4 Discussion of Classification Results

5.4.1 Summary of Classification Results

Linked List Implementation

The linked list approach for the generation of cooccurrence image data is greatly preferred to the traditional matrix approach since it calculates texture features orders of magnitude faster. The computational time of the matrix approach is more dependent on the number of grey levels than the linked list approach. Depending on the application, performing pixel-by-pixel segmentation can be computationally feasible using the linked list approach with quantized grey levels, however, generation of cooccurrence data will not always lend itself to a rapid feature extraction. The full dynamic range of the image can be utilized using GLCLLs, but only with an increase in the computation time. The degree of this increase is texture dependent.

Textures that have noticeable but subtle differences at full dynamic range may become statistically and visually similar under coarse quantization. This should be important when classifying remotely sensed data, especially sea ice imagery. A SAR sea ice image can contain many different types of ice types as well as transitions between these ice types. Quantizing the imagery can remove the subtle differences between two distinct similar ice classes that must be segmented. It should be much safer and more consistent if a feature extraction method utilizes the full dynamic

range or a quantization level that does not significantly attenuate the textural information.

Cooccurrence Features

The classification studies reveal that grey level quantization has an inconsistent effect on the classification accuracy across all the available statistics and the given data sets. Smoothness statistics {DIS, CON, INV, IDM} are most effective when the full dynamic range is used and provide more consistent texture measures for different grey level quantizations than the homogeneity statistics {MAX, ENT, UNI}. Homogeneity statistics improve classification ability with increased grey level quantization. This raises an interesting issue. Since poorer homogeneity features are extracted at full dynamic range and using only two grey levels, then between these two extremes there should exist a preferred grey level quantization. How can this optimum be selected? Can it be selected *a priori*? Must some knowledge of the grey level distributions for each class be known? Because of such variability, are these features useful for unsupervised applications?

For the Limex imagery, the classification accuracies generated by using the DIS statistic alone have been compared to three scenarios: all of the statistics, {DIS, ENT, COR}, and {DIS, ENT}. The results using DIS alone proved to be superior when classifying the test samples. For the Brodatz imagery, which contains textures that contain more structure, the results are a bit different. Adding COR to the DIS feature set increases the classification accuracy. Furthermore, {DIS, ENT, COR} generate a strong classification that is quite consistent across all grey level quantizations and comparable to the results using all statistics. For structured images such as Brodatz imagery, homogeneity features and COR are important. SAR sea ice imagery can contain textures that are well structured, necessitating

	Brodatz		Limex	
Grey Levels	Training (%)	Test (%)	Training (%)	Test (%)
full	94.5	85.9	99.0	90.0
128	93.2	84.0	97.9	89.0
64	93.0	85.7	97.4	89.7
32	94.0	86.7	96.9	91.3
16	94.1	84.2	95.3	90.7

Table 5.21: Classification accuracy across different grey level quantizations using statistics {ENT, DIS, COR}.

application of COR and homogeneity statistics (although well structured textures are not available in the Limex test set). Thus, these are the three statistics {DIS, ENT, COR} that will be used for the segmentation analysis in combination with the four primary orientations and a pixel spacing of one. This is essentially the same conclusion as Baraldi and Parmiggiani [2]. A grey level quantization of 32 will be used to achieve a compromise between accuracy and speed. Results using this selected parameter set on the Brodatz and Limex imagery are found in Table 5.21. The results of this data set are quite similar to that of using all the cooccurrence statistics (Table 5.3).

Power Spectrum Features

Power spectrum features do not generate sufficiently strong classification results to motivate their use for segmentation analysis. They are restricted because they require fixed window sizes. Perhaps this method requires larger windows so that sufficient periods of the texture are represented to calculate meaningful features. Another drawback of this method is that, although two dominant peaks are utilized, additional dominant frequencies may be required by the HVS to identify a given

texture. Underestimating or overestimating the number of peaks will prevent the feature set from performing optimally. Peaks provide appropriate measures for textures that have one or two dominant (F, θ) pairs. Generally, the summed features tend to give better results. These features are similar to those obtained by the Gabor features, which are advocated since they are able to capture multi-resolutional features. The power spectrum analysis did demonstrate that the spatial-frequency texture measures are independent of the grey level quantization (as long as the quantization does not destroy salient texture characteristics).

Gabor Filter Features

Since better results are obtained using $B_\theta = S_\theta = 30°$, this will be the choice for the texture segmentation studies. The filter bank for a 256x256 image using the four highest frequency bands is illustrated in Figure 5.5. There is some concern that the added dimensionality of 30° spacing may not be warranted due to the increased computational demands. That is, using 30° spacing instead of 45° spacing increases the feature set by 1.5 times and there is a corresponding increase in memory, disk space, completion times, and swap space to perform the analysis. For some test texture images, setting S_θ to 45° will generate a proper segmentation, however, to generate a more robust and universal feature set, $S_\theta = 30°$ is recommended. With higher smoothing (by decreasing γ), better classification results occur. This may be detrimental when textures representing small spatial regions are to be segmented.

Both the magnitude and full wave rectification responses are successful for classifying the data sets. Since the unsmoothed real responses are relatively much poorer, utilizing the real response is not recommended. The smoothed real responses do generate strong discriminations, however, if less smoothing is used, how these features will respond is uncertain. The Gabor homogeneity features do mildly

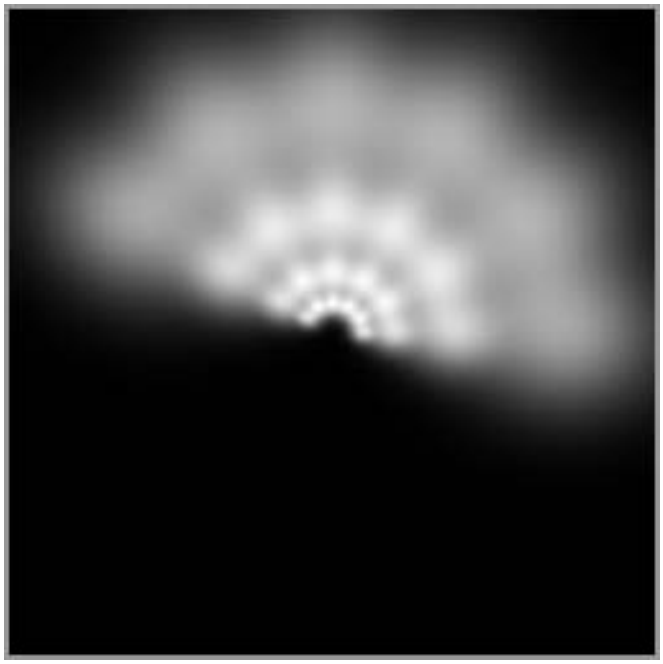


Figure 5.5: Frequency domain representation of Gabor filter bank for segmentation of 256x256 images.

increase the classifiability of the test data sets. However, given the added computation, their implementation may not be warranted. Generating moments based on the local spectrums consistently decrease the discriminability of the texture features so their implementation is not advocated.

In this thesis, successful results segmenting textured images using Gabor filters have been generated by:

- using a pseudo-wavelet implementation.
- providing full coverage of the spatial-frequency domain, without significant overlap between individual filters.
- setting the DC gain to zero to prevent classification based on tone.

- utilizing the magnitude of the Gabor response.
- smoothing outputs.
- using a class pairwise Fisher linear discriminant to classify the training and test sets.

5.4.2 Relationship Between Gabor and Cooccurrence Features

There is a high degree of correlation between the tuned Gabor filter outputs and corresponding cooccurrence measurements. Figure 5.6 compares the response for a Gabor filter at $(2\sqrt{2}cpi, 0°)$ (16x16 sample) and a cooccurrence measurement for (CON,0°) using the cotton and wood textures. Both of these textures have noticeable vertical components. The correlation for wood given these two features is 0.82; for cotton, 0.60. These are very strong correlations, especially given that the features are extracted using two completely different techniques. Such high correlations indicate that both methods tend to generate the same critical information about the particular texture.

5.4.3 Comparison of Results to Other Studies

There exist other texture classification comparative studies. These will be reviewed and contrasted with respect to the current study. To make direct comparisons is awkward due to differences in the test data and differences in the methodologies implemented. In some cases, the potential of the methods for image segmentation is evaluated.

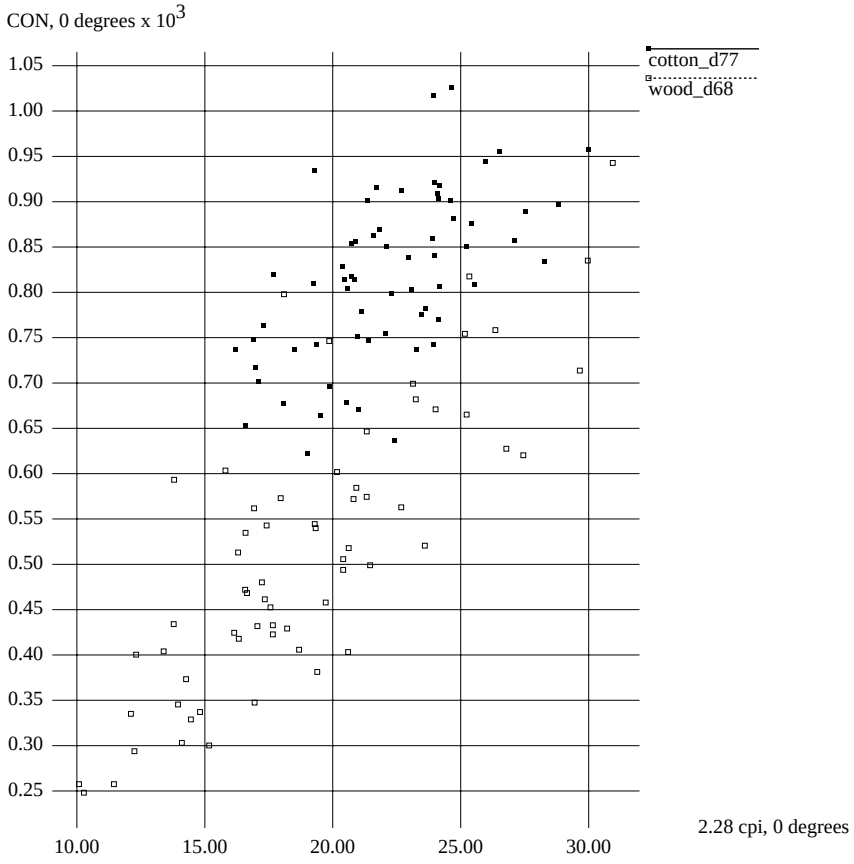


Figure 5.6: Demonstration of the correlation between Gabor filters and cooccurrence probabilities.

Comparing Cooccurrence and Power Spectrum Results

In one of the earliest studies, Weszka *et al.* [107] compare three standard approaches to texture classification and demonstrate that the summed power spectrum features perform the poorest and the cooccurrence method is comparable to the first-order statistics of grey level differences. This study is of limited applicability because of insufficient data and application of only single and pairwise features when classifying.

D'Astous and Jernigan [26] present the results of comparative texture classification of peak and shape features, summed features, and a lone cooccurrence texture measure (CON). Their results indicate that the peak and shape measures are superior to the cooccurrence measure which is superior to the summed features when 32x32 samples of Brodatz imagery are used. Liu and Jernigan [74] investigate the potential of using D'Astous' texture features in the presence of additive noise. They also introduce additional power spectrum texture features to compare a total of twenty-eight such features. Five Brodatz classes are successfully classified based on training/test data sets (about 95% accuracy). They indicate that their methodology is an improvement over the results obtained in D'Astous' thesis [27].

He *et al.* [53] compare cooccurrence features, peak power spectrum features, and summed power spectrum features. The IDM statistic proved to be very effective discriminating the Brodatz texture images. Contrary to D'Astous and Jernigan results, the summed energy features outperformed peak texture features. Overall, of the top seven texture features in their analysis, six are based on cooccurrence features.

McLean [82] compares a vector quantization (VQ) approach to cooccurrence and power spectrum texture features. He used a k-nearest neighbour approach to classify samples from seven Brodatz images. This classification strategy is appropriate for the vector quantization approach, however, it may not be a preferred method for classifying cooccurrence and power spectrum feature sets. For his study, the VQ results are superior. His classification results for using the cooccurrence technique (60.9%) seems quite low compared to the results presented earlier in this chapter using 16x16 windows of similar Brodatz imagery. The classification accuracy of the cooccurrence features consistently exceeds that of the power spectrum features by a significant margin. The VQ methodology runs into the same

difficulty as the cooccurrence and power spectrum approaches if implemented for segmentation purposes - selection of window size. In addition, increasing the window dimension in the VQ approach causes a quadratic increase in the feature space. Unsupervised approaches using VQ are awkward because classes are identified by multiple prototypes.

Essentially, these results support the testing presented earlier. Namely, the cooccurrence features tend to perform better than the power spectrum features. Under conditions where the textures are quite regular and window sizes are a sufficient size to represent a number of periods of the texture, power spectrum measures can be quite strong. When the window size is equal to or less than the period of the texture, the cooccurrence features are better able to handle such uncertainty.

Cooccurrence Features Compared to Gabor Filtering Results

Augusuteijn *et al.* [1] classified a Thematic Mapper (TM) set of images using a number of different methods, including cooccurrence matrices, power spectrum, and Gabor features. Power spectrum features (as defined in their paper) and Gabor filters are found to be preferred choices. Cooccurrence features also generate successful results, but their applicability is shunned by the authors due to the long execution times. Even for a single image band, the results for these three methods are all quite successful using 8x8 windows, suggesting that the samples are distinct. They even use the average of the four dominant angles to generate the cooccurrence features, indicating that the textures are isotropic. Their filter methodology may not be appropriate since they use a constant bandwidth for all frequencies and this may not generate suitable spatial-frequency domain coverage. The Gabor filters are not forced to have a zero mean impulse response, hence, the filters are using local mean grey levels to assist discrimination. This is substantiated by the fact that, to

achieve a strong classification based on the cooccurrence features, the mean grey level had to be included as a feature.

Kundu and Chen [67] compare the cooccurrence features with those obtained from a quadrature mirror filter (QMF) bank. Interestingly, they apply the cooccurrence features to the low-low output of the filter bank and derive other additional features from the high pass filtered data. Their results show that the combined QMF/cooccurrence features are preferred over the cooccurrence technique alone. Applying results from the QMF bank to image segmentation is awkward since the filtering methodology is set up in a block format. When compared with the Gabor wavelet scheme, not every pixel has its own individual response for every resolution and orientation.

Chang and Kuo [18] use a tree structured wavelet to perform classification. A number of methods are compared. The discrete cosine transform (DCT), discrete sine transform (DST), discrete Hadamard transform, and Laws masks all performed similarly and second best to their tree structured approach. For the same reasons as Kundu and Chen, this method is not conducive to image segmentation. The Gabor filtering seems to lag in performance. However, their implementation only used the real part of the Gabor filter without smoothing, which, as described in Section 5.3.2, is not expected to generate appropriate texture features.

Ohanian and Dubes [84] compare four different techniques for their classification ability: Markov random field parameters, Gabor filters, fractal based features, and cooccurrence features. Two syntactic (fractal) and two real (leather and painted surfaces) are used as test imagery. The cooccurrence features generate the best classification rate, determined by a K-nearest neighbour classifier and leave-one-out approach. However, the Gabor filtering is probably not performing optimally since a constant spatial-frequency bandwidth is used without any local smoothing.

Barber *et al.* [8] compare a number of different methods in the context of classifying samples of SAR sea ice imagery. The cooccurrence features are deemed better than the Gabor filtered data implemented using Wright's filter parameters (Section 4.1.4). The poor success of the filtered data can be explained in a number of ways. Wright's parameters are not organized in a wavelet fashion and the bandwidths are not set up to provide proper spatial-frequency coverage. Also, the filter output is unnecessarily normalized by its area covered in the spatial-frequency domain. In addition to these features, the filter outputs are manipulated using the approach that D'Astous used working with the power spectrum [27]. From the entire set of Gabor filter outputs and D'Astous features based on these outputs, only two features are selected. This limited amount of information may explain why the Gabor filter outputs are not able to perform optimally. It is interesting to note that each of the methods compared in this publication used different classifiers. Typically, in the other comparison papers, only one classifier is used so as to treat each texture analysis method equally.

Chapter 6

Texture Segmentation Study

This chapter draws on information accumulated thus far in this thesis to present an unsupervised texture segmentation methodology. Drawbacks of current approaches are summarized to motivate the design of a clustering scheme. This algorithm is applied to the segmentation of test and SAR sea ice imagery. At the same time, distinct abilities of the cooccurrence and Gabor texture features are identified. The end result is a theoretical approach for unsupervised texture segmentation and an operational approach for SAR sea ice image texture segmentation. A comparison of these results with other published studies is performed.

6.1 Design of Clustering Methodology

Both Gabor filters and cooccurrence probabilities are able to characterize textural information. However, an unsupervised method to group the feature vectors to create a segmented image is difficult to devise. Determining the boundaries between classes as well as grouping regions with similar characteristics is necessary. No

successful, robust method has been identified in the research literature. One of the more difficult tasks is determination of the number of distinct texture classes in the image. The goal is to design a generic texture segmentation methodology that works with a wide variety of imagery, not a segmentation approach that only works with two classes, very distinct textures, or straight boundaries. An algorithm that best mimics what the HVS is able to perform and considers to be distinctive is important. This section will present a method that is able to identify the pertinent textural regions and the number of classes in the image given that the classes are well identifed by the features. Keep in mind that the derivation and examples are based wholly on textural cues only, not tonal.

6.1.1 Drawbacks of Current Approaches for Clustering Gabor Texture Features

Dunn *et al.* [35, 36] (also discussed in Section 4.1.4) perform image segmentation in a very limited manner using Gabor filtered outputs: only two distinct textures are found in the image, each texture takes up one-half the image, the textures are separated by a single straight boundary, training samples are required (fully supervised), and only one filter is used to distinguish the two textures (although textures may have more than one dominant frequency/orientation pair). Dunn's current approach is not appropriate for the generic texture segmentation problem. His concept that only a limited number of filters is required to segment an image, however, is important from a computational perspective.

Jain and Farrokhnia have published a well-referenced texture segmentation algorithm [59]. This algorithm has significant restrictions.

- The methodology for determining the number of classes is not effective since

it works properly for only one of the provided eight examples [40]. They implement an incremental strategy using the modified Hubert index to indicate when the number of classes is exceeded. This requires that the entire system be solved at least as many times as the number of classes, which is computationally expensive.

- The ISODATA algorithm is applied to cluster the feature vectors. Since the number of samples per class is roughly the same for their examples, the ISODATA algorithm may be using this *a priori* knowledge to indicate minimum and maximum sizes for the clusters. The ISODATA algorithm is not appropriate for solving the generic texture segmentation problem and should only be used when considerable knowledge about the image is known. SAR sea ice imagery is an example where the number of samples per class cannot be assumed *a priori*.

- The spatial coordinates of the pixels are used as features. The example images have individual textures placed in a block-like fashion, where textures are located in only one part of the image and the areas covered by the textures are approximately the same. For an image like Figure 6.7(a), using spatial coordinates as features provides a tremendous advantage for segmentation. For an image such as Figure 6.5(a) or Figure 6.9(a), the spatial coordinates will only confuse the classification since areas represented by each texture are different and some textures are located in more than one region of the image. Thus, inclusion of the spatial coordinates in the feature set is not appropriate for solving the generic texture segmentation problem. The smoothing of the feature maps in the Gabor scheme provides this type of close neighbour information.

6.1.2 Clustering Implementation

To solve the general texture segmentation problem, the classes should be represented by fairly well defined clusters in the feature space. The efforts of Chapter 5 are directed towards achieving this goal. There are several other constraints which significantly compound the problem:

- Each of the clusters may be represented by a different covariance matrix.
- No assumptions can be made about the number of samples that belong to each class.
- The total number of classes is not known at the onset.

The goal is the ability to identify dense groupings of samples in the feature space. To find the centroids of the clusters and the proper class variances associated with them, a two step approach is used that requires procedures outlined in Section 4.2.1. The K-means algorithm is used to find the centroids of each of the classes. This, of course, assumes that the number of classes is known. Once the centroids are found, the samples can be assigned and this represents a rough approximation of the true segmentation. If the clusters are well separated, then the centroids should be found and a general description of the segmentation generated. The clusters can then be processed using a class-pairwise Fisher linear discriminant algorithm (FLD). The *a priori* probabilities as estimated by the cluster sizes are used. The Fisher discriminant will correct for the class variances, but it is not certain that the variances have been fully corrected following only one pass. Thus, the clusters are improved incrementally by applying the Fisher discriminant algorithm iteratively (iterative FLD or iFLD). Applying K-means first followed by iFLD will be referred to as the KIF (K-means iterative Fisher) algorithm.

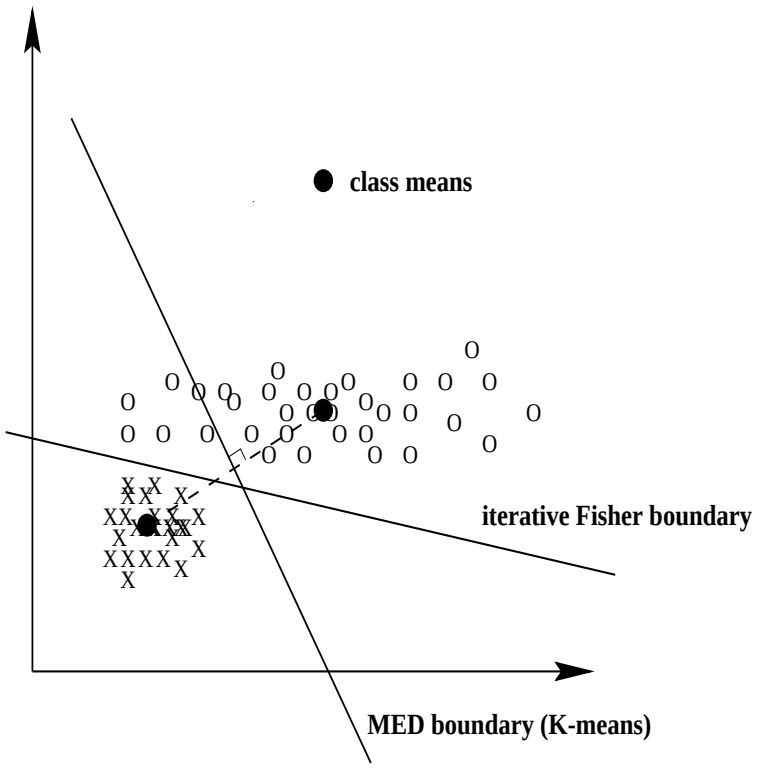


Figure 6.1: Clustering accuracy of K-means versus KIF.

The advantage of KIF is illustrated in Figure 6.1. The K-means algorithm finds the means of the two class clusters, but, the Euclidean distance classifier may assign many samples to the wrong class. The iterative Fisher linear discriminant (iFLD) generates a more appropriate separation of the classes by utilizing the class covariances.

When do the iterations cease? As in any clustering problem, to ascertain when the optimal separation occurs is difficult. Here, the average of the inter-class Fisher distances is used as an indicator. If this average decreases following an iteration, then the algorithm ends. However, in my experience only a few iterations are required to correct for the class variances, so the maximum number of iterations

is set to five. If this Fisher approach is not able to improve the segmentation following five iterations, then the shortcoming of the algorithm is that K-means did not identify the prototypes, probably because the clusters are not well separated in the feature space. This is either due to the features not accurately distinguishing the classes or due to the similarity of the textures in the image. The K-means algorithm makes the inherent assumption that each of the classes forms a distinct Gaussian-like cluster in the feature space.

The feature space is actually quite sparse. If the image is 256x256 then there are 65536 samples. Assuming only two classes in the image, one would encompass a minimum of 32768 samples. The cooccurrence approach will have a smaller dimensional space (12-d) than the Gabor approach (24-d). Assuming that ten samples are sufficient to describe the probability density function (pdf) in 1-d space, approximately 10^{12} samples are required to represent the 12-dimensional pdf for a single class. Obviously, the 32768 samples are inadequate to do this. Fortunately, the K-means algorithm is quite oblivious to the sparseness and only identifies dense regions as identified by the minimum Euclidean distance.

One must make sure that the feature space is properly scaled. For the cooccurrence texture features, each selected statistic is performing a different measure on the cooccurrence probabilities, thus, the feature dimensions must be normalized to have the same range. On the other hand, the Gabor filters are measuring the same qualitative aspect of the textures; namely, the responses to certain filters. Scaling these features will destroy the relative responses of the filter outputs, critical information for the segmentation. Thus, to scale the Gaborian feature space is unnecessary.

After K-means determines the class prototypes, the next task is to improve the segmentation using the iterative Fisher linear discriminant. This discriminant

projects the sample vectors from an n-dimensional space into a 1-dimensional space. The resultant 1-dimensional pdfs have sufficient samples to properly estimate their parameters. One could use a minimum inter-class distance metric for classifying samples, however, the quadratic boundary is not justified with this relatively sparse data set.

The seeds that are used for the K-means algorithm are the first samples found in the image (beginning at the top left hand corner). Given no other information about the image, this is a suitable place to begin. Due to the assumed periodicity created with discrete time forward and inverse FFTs (used to generate the Gabor features), these seeds have the potential to lie near a boundary between numerous distinct textures (from each corner of the image). For the cooccurrence measures, a boundary is set-up around the image so as to mimic periodicity. The boundary width is one-half the window dimension used to extract the cooccurrence features. This ensures that the Gabor and cooccurrence methods are treated equally.

6.1.3 Determining the Number of Classes Using a Divisive Hierarchical Tree

Determining the number of classes in a feature set is a difficult problem. The distinctiveness of a pair of clusters can be determined by an inter-cluster distance measure. The Fisher criteria (Equation 4.3) provides such a measure by determining the ratio of the Euclidean distance between the means and the cluster covariances. If the cluster projections onto a Fisher discriminant are always Gaussian, the error between two classes can be determined. The Gaussian assumption does not always hold which leads to inaccurate error estimates. Parzen pdf estimates of the projected data can be made and the error determined. This has not been attempted

due to the added computation. The Fisher distances offer accurate estimates of the separability of the clusters and will help determine the number of clusters.

A first try at solving the clustering problem was performed by incrementing the number of classes and solving the entire set of samples for each increment (like Jain and Farrokhnia as outlined on page 110). Instead of using the modified Hubert index, the inter-cluster Fisher distance was used as an indication of whether to continue incrementing the number of classes or to stop the algorithm. With each increment, one of the following should occur: a distinct class is identified from a larger grouping of clusters, a cluster of classes is split into a pair of subclusters or classes, or a single class is split. This method proved awkward because the most separable classes are not always the first classes identified by the K-means algorithm, so to determine when the algorithm should stop is difficult. Also, single complex textures might be split before two distinct yet similar classes are separated and a method to determine when this would happen is unknown. When multiple classes are being considered the average inter-class Fisher distance is not an appropriate method to indicate when the iterations should stop since, although the average may drop, the best segmentation may not have been achieved. Inter-class distances may be reduced for well defined class pairs, but those class pairs that are not as separable may not be properly separated when the average inter-class distance drops.

A better method to perform the clustering and avoid solving the entire feature set as many times as the number of classes, is a binary divisive hierarchical approach. This approach is best explained graphically. Consider the case of three distinct classes in a 2-d space (Figure 6.2). At the top of the binary hierarchical tree, two classes are assumed. Applying KIF to the entire set of samples generates clusters A and BC (clusters B and C combined). The Fisher discriminant combines

clusters B and C so that the joint projection resembles a single class on the A-BC discriminant, generating a large Fisher distance. The large distance will allow the algorithm to proceed by assuming that A and BC are distinct. Then, K-means and the iterative Fisher scheme are applied to each of the two clusters, A and BC, to attempt to split them up. The splitting of the single cluster A generates a low Fisher distance so cluster A is left as an individual class. The BC cluster is easily separated. Here, the tree can proceed to the next level and try separating the two clusters B and C. As when A was considered, the splitting of B and C generates low Fisher distances and they are left as individual classes. The final clustering identifies A, B, and C as individual classes. This method is also able to determine if there is only one texture class in the image.

If the clusters are well separated in the feature space, there will exist a range of acceptable Fisher distances that can be used to generate an accurate unsupervised segmentation. Let τ denote the threshold value. If τ is set below the minimum of the range, then classes that should not be split will be split. If τ is set above the maximum of the range, then clusters that should be split into classes or subclusters will not be split. By setting τ to any value in the correct range, the segmentation will proceed properly. If all textures in the image are highly separable and well clustered, then τ has a large range. Setting τ is a function of the feature extraction method and the type of imagery being analysed.

6.2 Algorithm Testing

Testing will be performed on both test and SAR images. Brodatz images are the primary test images used and these assist investigation of theoretical concerns dealing with unsupervised segmentation. This information will be carried to the SAR

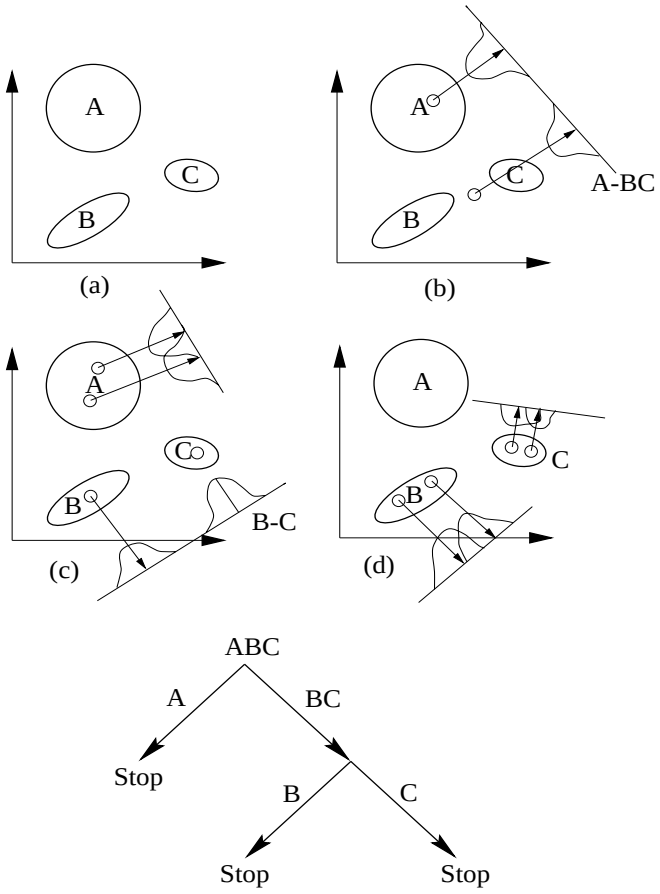


Figure 6.2: Divisive hierarchical clustering strategy

images where a practical implementation for image segmentation will be described and presented. Comparisons between Gabor and cooccurrence texture features will be performed.

6.2.1 Segmentation of Test Imagery

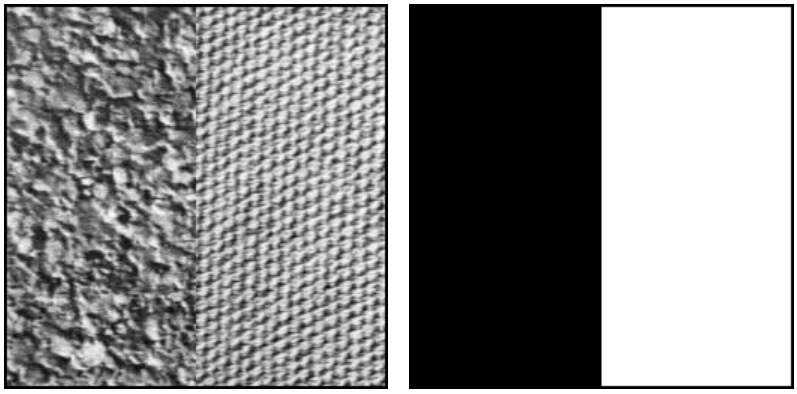
Two segmentation approaches will be used in this section. If the only information provided to the clustering problem is the number of classes, this is known as mixture resolving [58]. Unsupervised cluster analysis refers to clustering without any *a*

priori information whatsoever. Both the mixture resolving and cluster analysis approaches are presented for each image using both cooccurrence and Gabor feature sets. Mixture analysis is performed by giving the correct number of classes to the KIF algorithm. Cluster analysis uses the hierarchical divisive approach.

The first image is a simple 256x256 bipartite image (Figure 6.3). The original image (a) contains cork (D4) and cotton (D77) and (b) represents the true segmentation. Both the Gabor (c) and cooccurrence (d) segmentations have similar results and are easily able to segment the images in an unsupervised fashion. The range of τ is [9.11, 31.3] for the Gabor approach and [4.22, 23.6] for the cooccurrence approach. The Gabor approach tends to generate clusters that are more separable.

The next image (Figure 6.4) contains the same textures as Figure 6.3 but the sinusoidal boundary makes it more difficult to distinguish the two textures. The segmentation results are again about the same, however, some changes in the ranges for τ occurs. The Gabor segmentation requires a range of τ = [7.44, 13.1] and the cooccurrence segmentation requires τ = [8.58, 10.5]. The upper limits of τ=13.1 and τ=10.5 demonstrate that the two clusters representing the two classes are not as separable as in the case of the straight boundary (τ=31.3 and τ=23.6).

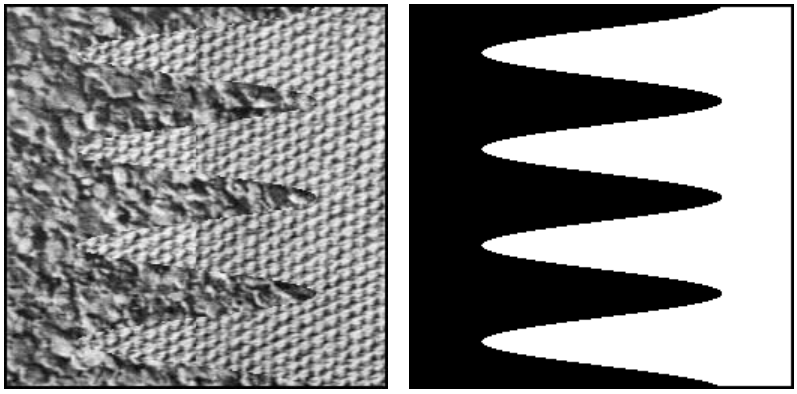
The next step in complexity involves adding textures to the test image. Figure 6.5 represents segmentation of an image with four Brodatz textures: calf (D24), grass (D9), wool (D19), and wood (D68). This image was originally published by Krishnamachari and Chellappa [66]. Note that each of the textures have approximately the same textural resolution. Both the Gabor and cooccurrence methods are able to segment the image, using mixture resolving ((b) and (c)) and unsupervised cluster analysis ((d) and (e)). The ranges of τ required to perform the segmentation unsupervised are τ=[8.4, 10.1] for the Gabor approach and τ=[4.81,9.21] for the cooccurrence approach. The Gabor approach has a tendency to want to subdivide



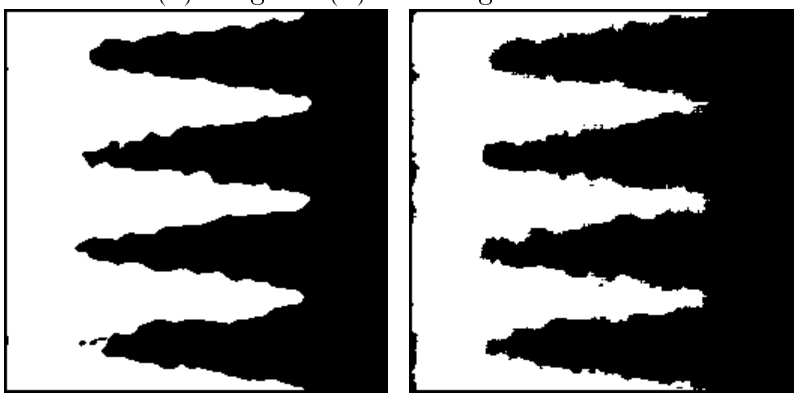
(a) Original (b) True segmentation

Unsupervised: (c) Gabor result (d) Cooccurrence result

Figure 6.3: Segmentation of bipartite Brodatz image separated by straight boundary.



(a) Original (b) True segmentation



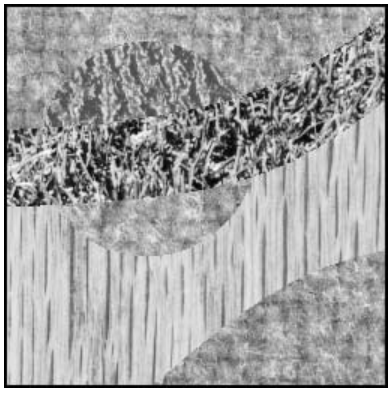
Unsupervised (c) Gabor result (d) Cooccurrence result

Figure 6.4: Segmentation of bipartite Brodatz image separated by sinusoidal boundary.

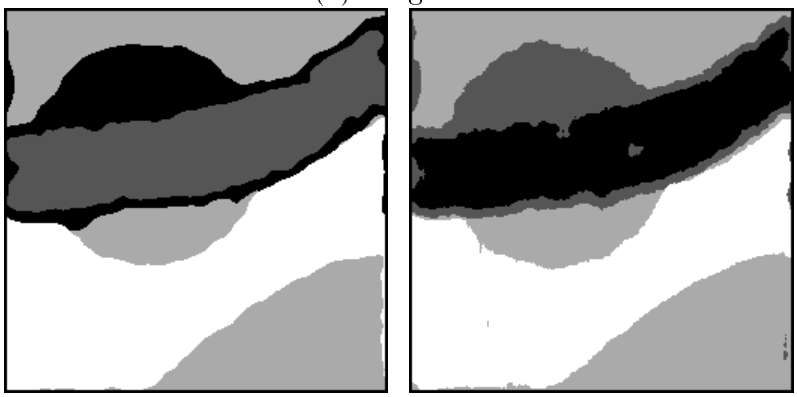
the grass texture (τ=8.4 compared to τ=4.81 for the cooccurrence method). The grass is not very consistent and represents a texture that is not easy to characterize. The mixture analysis in (b) and (c) increases the error along the lower boundary of the grass texture for both methods. The hierarchical approach reduces this error by considering regions individually.

As an example of how the iterative Fisher linear discriminant improves on the K-means clustering, Figure 6.6 shows the clustering at the top of the hierarchical tree. The system tries to divide all the Gabor feature vectors for the image into two clusters. K-means is able to produce a fairly accurate segmentation. The iterative Fisher approach significantly reduces the error so that the four classes are more accurately divided into two clusters of two.

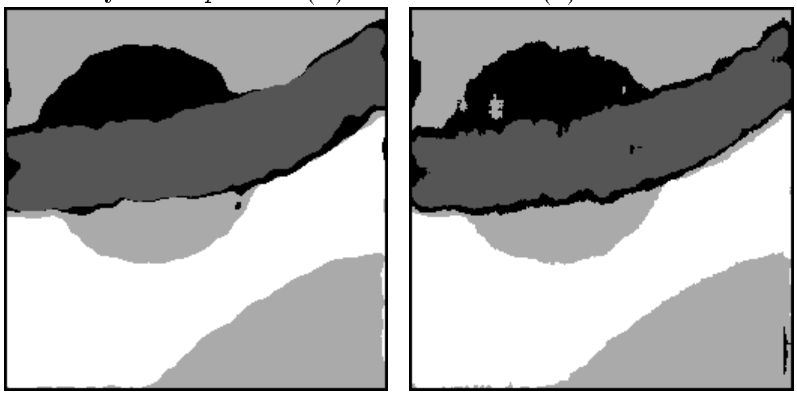
The next test image (Figure 6.7) also contains four Brodatz textures, however, these textures (cork - D4, cotton - D77, paper - D57, and raiffa - D84) have different textural resolutions. The cork and paper are poorly defined, especially compared to the cotton and raiffa. Whether mixture analysis or the hierarchical clustering approach is used, the textures are clearly discriminated using the Gabor features. The range for τ is [11.7, 12.3], indicating that an unsupervised segmentation is possible. The cooccurrence features cannot discriminate between the cork and the paper textures. Using the unsupervised protocol, the cooccurrence segmentation would have stopped after the first level in the hierarchical tree. Here, it would isolate the cotton texture, but then find that the cotton texture is easier (based on Fisher distances) to separate into two clusters than to break down the remaining three classes into two clusters. However, if the image is forced to split into four clusters, the clustered image in (f) results. Even when the proper number of classes is provided, the cork and paper textures cannot be distinguished using the cooccurrence features (d).



(a) Original

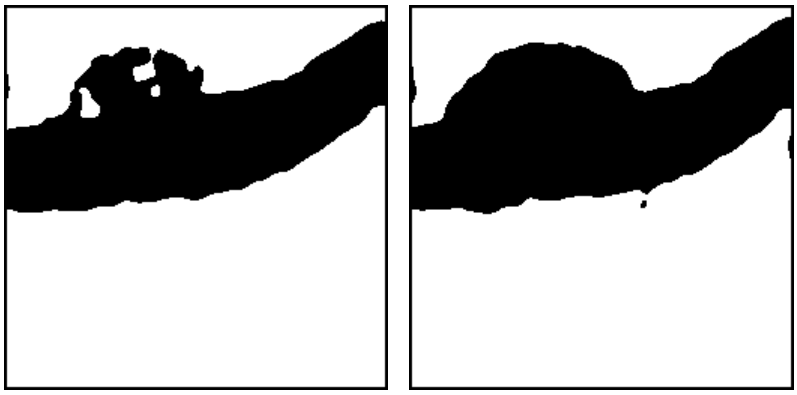


Mixture analysis: *a priori*: (b) Gabor result (c) Cooccurrence result



Unsupervised: (d) Gabor result (e) Cooccurrence result

Figure 6.5: Segmentation of image with four Brodatz textures (obtained from [66]).

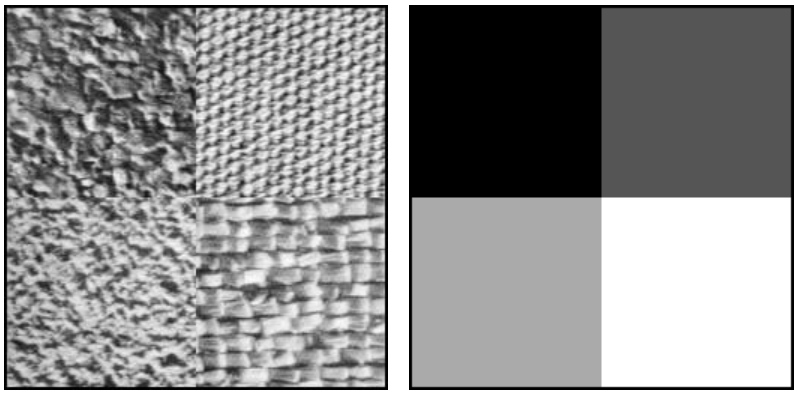


(a) Following K-means (b) Following K-means and iterative FLD

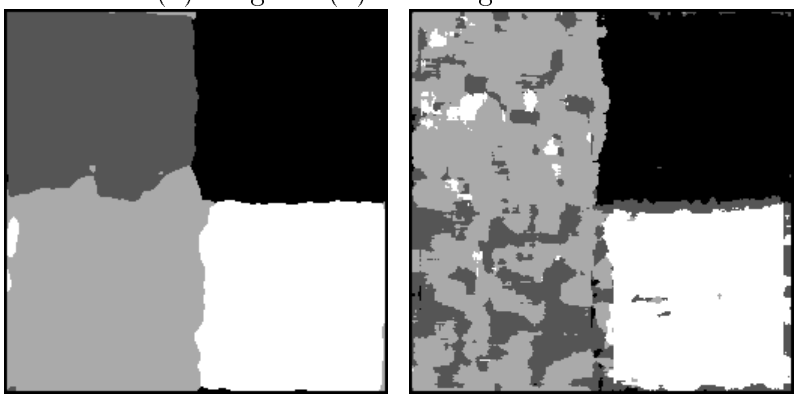
Figure 6.6: Demonstration of the ability of KIF to generate better clusters than K-means alone.

If a larger window size is used to capture the cooccurrence features, the cork and paper textures may be better discriminated, however, the boundaries between the textures will be more erroneous. If a smaller window size is used, the boundaries are better approximated, but the regions have poorer estimates. The Gabor technique is better suited to the multi-resolutional problem since different frequencies are tuned to different textural resolutions found in the image.

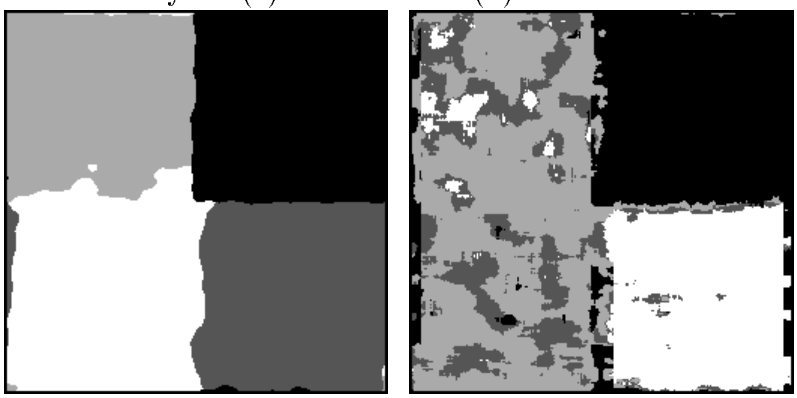
What happens if Figure 6.7 is created with sinusoidal instead of straight boundaries (Figure 6.8)? The Gabor features are not able to distinguish between cork and paper for the mixture analysis case. Since the boundaries occupy a significant area of the image, the blurred textural boundaries are identified as a separate class. Using less smoothing could alleviate this problem. A segmentation was performed using the average local grey level as an additional feature. Using this feature set, the mixture analysis is able to distinguish between cork and paper and generate a result similar to (e). The hierarchical approach (e) does manage to separate the two classes (τ has a range [9.4, 13.8]) without using the average grey level as a



(a) Original (b) True segmentation



Mixture analysis: (c) Gabor result (d) Cooccurrence result



(e) Gabor result (unsupervised) (f) Cooccurrence result (forced)

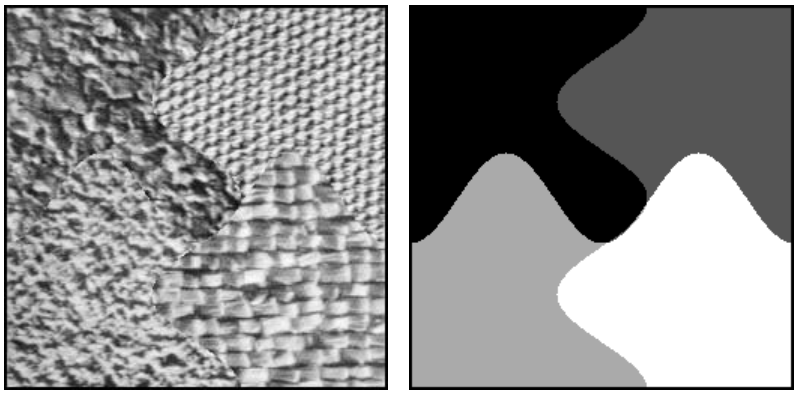
Figure 6.7: Segmentation of image with four Brodatz textures separated by straight boundaries.

feature. The lobe at the centre of the image that should be assigned to the cork class is erroneously assigned to the paper class. In this local region, insufficient spatial extent is provided to accurately identify the paper texture. The Gaussian smoothing blurs the boundary between the paper and cork causing the cork to be assigned to the paper class. Regions are estimated well but boundaries are not accurately estimated.

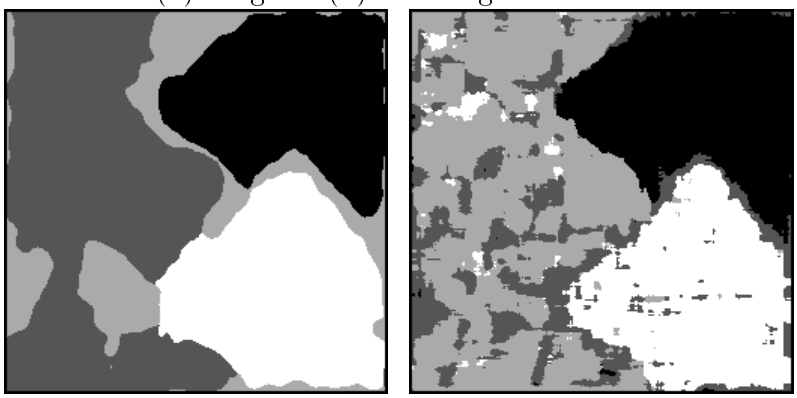
The cooccurrence features are again unable to discriminate between the paper and the cork textures. Here, the unsupervised algorithm would have again stopped after finding only two classes. Image (f) presents the results if the clusters are forced to divide properly into four classes.

The image represented in Figure 6.9 was originally published by Bigun and du Buf [10]. Seven Brodatz textures are combined in a 16 patch mosaic so that each texture's boundary touches each of the others at least once. The textures are visibly distinct from each other. One texture is noticeably complex and the rest have regular patterns. Different textural resolutions are apparent. Horizontal boundaries are straight, but the vertical boundaries are generated by a random walk procedure.

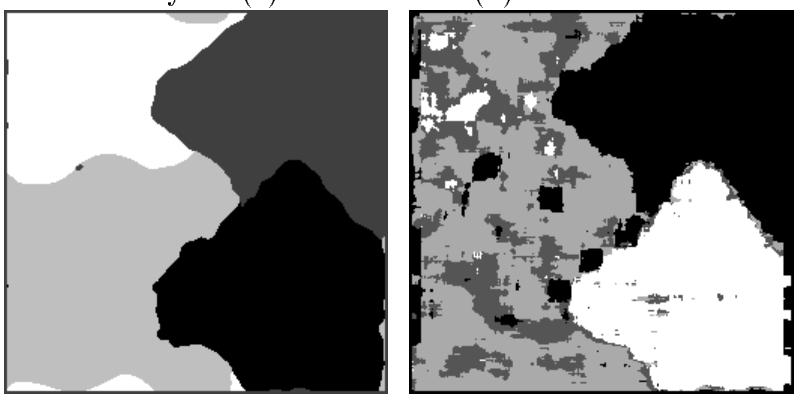
The Gabor features easily segment the image into the seven classes. The hierarchical unsupervised approach (d) generates a segmentation that is marginally better than the mixture resolving approach (b). For the unsupervised approach, τ has a large range of [9.11, 14.2]. The final seven clusters that are determined are passed through the iterative Fisher linear discriminant to improve the segmentation. The cooccurrence features are not able to segment this image effectively. For the mixture analysis, the cooccurrence features are not clusterable. The unsupervised hierarchical approach can only identify five of the seven textures and the boundaries between the textures are not well localized. Manually forcing the clusters to separate into seven classes, the textures are still poorly identified (e).



(a) Original (b) True segmentation



Mixture analysis: (c) Gabor result (d) Cooccurrence result



(e) Gabor result (unsupervised) (f) Cooccurrence result (forced)

Figure 6.8: Segmentation of image with four Brodatz textures separated by sinusoidal boundaries.

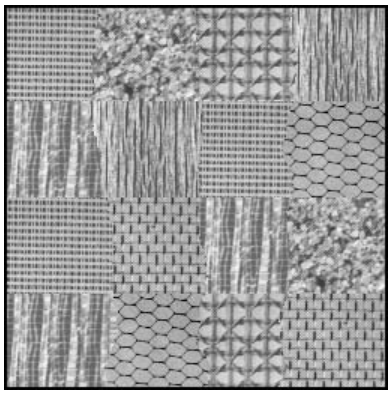
Again, if a larger window size is used, the regional identification will improve, but this will generate more errors at the textural boundaries. Smaller windows will do the opposite. For this type of image, the cooccurrence measures are ineffective.

The final image (Figure 6.10(a)) for this section is provided by Unser [105]. In this 128x128 test image, there are four distinct texture regions each with equalized histograms so that step edges between the textures are not apparent. The textural resolution is about the same for each of the textures. Both the Gabor and cooccurrence approaches are successful using the unsupervised approach (τ is [10.1, 12.8] for the Gabor features and [5.41, 10.3] for the cooccurrence features). The cooccurrence segmentation is sensitive to the window size. Smaller windows (8x8) have been used to obtain a better boundary approximation, but regional accuracies suffer noticeably.

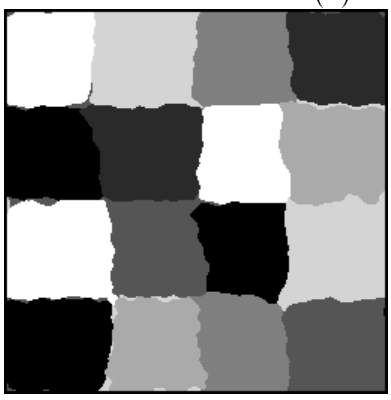
6.2.2 Segmentation of SAR Sea Ice Imagery

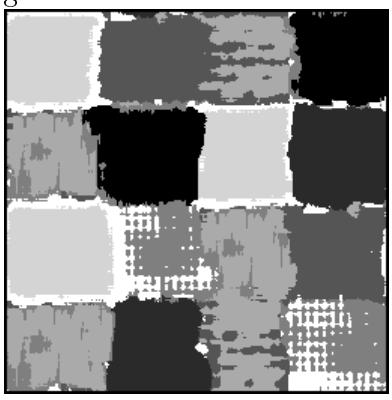
Operational Implementation

There is a noticeable lack of textural SAR sea ice image *segmentation* publications and none have been found that describe automated approaches. This is probably due to the requirement that unsupervised segmentation must satisfy additional demands: boundary identification and "vague" texture identification. Boundary identification is difficult since textural edges are highly variable, both in terms of directionality and edge definitiveness. Defining what constitutes a homogeneous region is problematic when dealing with natural imagery since there is considerable and unpredictable within-class variation. For example, multi-year ice is not always identical from image to image or even within the same scene but its general appearance is often the same. It is this general appearance that must be character-

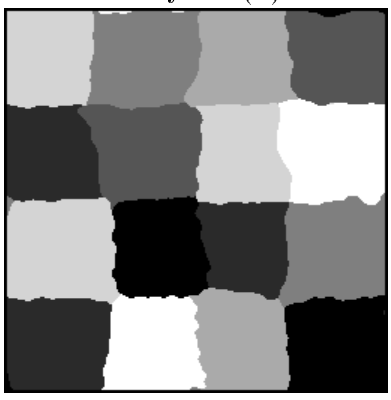


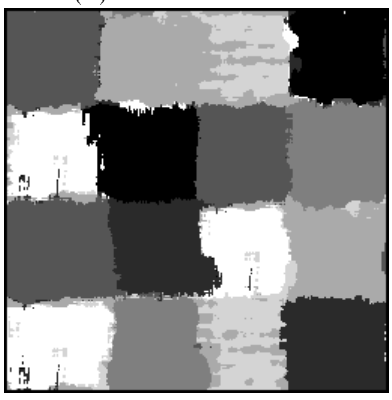
(a) Original





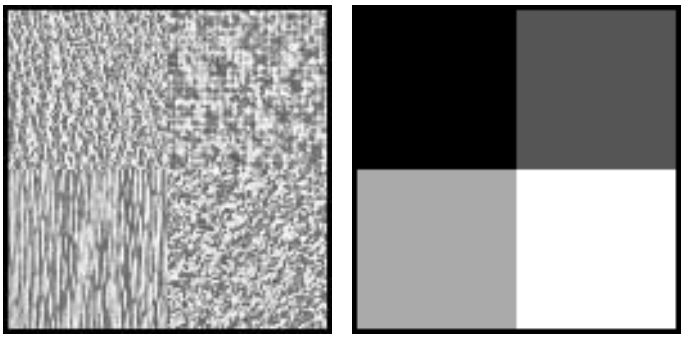
Mixture analysis: (b) Gabor result (c) Cooccurrence result



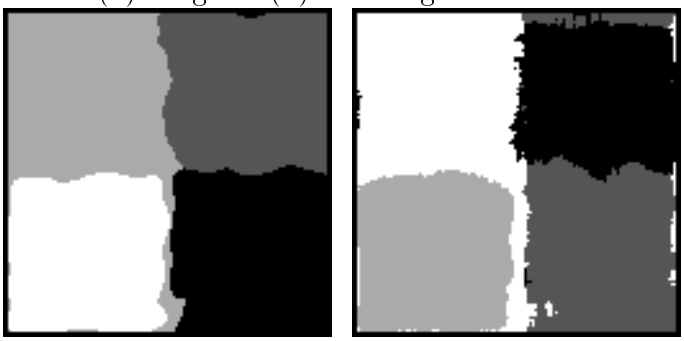


(d) Gabor result (unsupervised) (e) Cooccurrence result (forced)

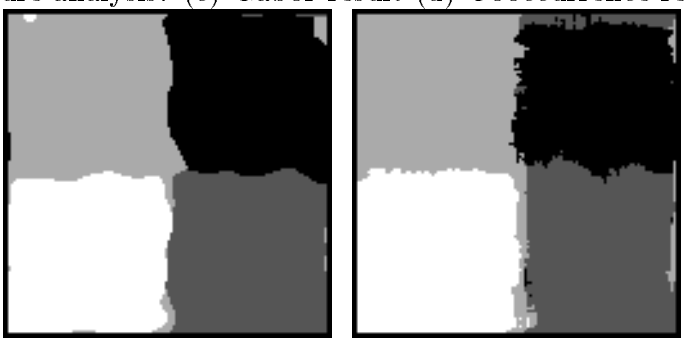
Figure 6.9: Segmentation of image with seven Brodatz textures (obtained from [10]).



(a) Original (b) True segmentation



Mixture analysis: (c) Gabor result (d) Cooccurrence result



Unsupervised: (e) Gabor result (f) Cooccurrence result

Figure 6.10: Four class image. Image obtained from [105].

ized. Recall that many different variables influence the outcome of the backscatter coefficient, further disrupting classification procedures. These identification problems exist even though the ice types are quite discriminable by a human observer. Because of this extensive variability, segmentation of SAR sea ice imagery should be performed in an unsupervised manner. Supervised approaches will not generate consistent results since selected class samples are insufficient to describe all the class variability throughout the image. Sometimes it is difficult to provide *a priori* an exact number of classes. A chief concern is finding a universal approach for all SAR sea ice data sets since texture features that are found to be suitable for one data set are not fully extendable to other data sets.

There are some generic characteristics that are required from a SAR sea ice segmentation system. Preferably, no preprocessing should be applied. This saves computation time and prevents destroying any pertinent textural information. Methods should be semi-automated to process the considerable amount of information. Also, the method should be robust and universal. It should be able to classify a particular scene, as well as adjust for seasonal and diurnal variations. End user requirements are also important. For navigational purposes, the user is particularly concerned about identification of multi-year ice. A risk factor that weights probabilistic error in favour of selecting multi-year ice types would help to avoid missing any multi-year floes. Collection of ice type concentrations throughout the Arctic region would require different risk factors depending on the goals of the study.

One of the goals of this thesis is to provide a practical tool for operators to generate end products for users of remotely sensed imagery. The wholly unsupervised approach is effective, but an incorrect value of τ could prevent segmenting classes that are visibly distinct. Also, the expertise to take into account the other visible cues that the operater requires has not yet been developed. A significant degree of

supervision is still required when interpreting such imagery.

The hierarchical segmentation approach is an appropriate technique to assist the user. The user starts with the full image. Using existing maps, the land is masked out. Land typically does not have a regular appearance and there is no known methodology for its automated recognition. Then, the KIF algorithm is applied to generate two clusters. The user chooses to continue subdividing either clustered region in the same manner until a satisfactory visible segmentation is achieved. Based on texture alone, the segmentation will probably not proceed beyond four or five textures. Nuances in the regions will subdivide the image further, but the human operator is probably better able to do this on their own. The computer is able to save the operator considerable effort subdividing the larger constant texture regions of the image. This is the approach used to segment the SAR images in this section.

Since the textures have a finer resolution than what appears in the Brodatz imagery, the feature extraction parameters have been modified accordingly. The Gabor features only use the three highest frequency bands and γ is set to 1.2. Cooccurrence features use a window size of 8x8. Computations for feature extraction are faster because of the reduced feature dimension (in the Gabor case) and the smaller window size (in the cooccurrence case). This is important because the image sizes are all 512x512.

Testing

Since the SAR sea ice images often contain information other than textural, careful attention has been paid to use test imagery that contain textural cues. To test algorithms on SAR imagery that require tonal features would lead to erroneous

results since both the Gabor and cooccurrence approaches have been implemented to be independent of local average grey level.

The images in Figures 6.11 and 6.14 were obtained from Barber *et al.* [8] and were captured using a STAR-1 platform (X band, HH polarization). The resolution is 6 metres (7 look). Each image is captured using a different look direction. Multi-year, first year smooth, and first year rough ice types are apparent. Both the Gabor results (in Figures 6.12 and 6.15) and the cooccurrence results (in Figures 6.13 and 6.16) generate similar segmentations. Since the Gabor method uses smoothing of the feature images, textural boundaries tend to be a bit more blurred, however, regional errors are fewer. Both methods generate accurate segmentations by identifying the pertinent ice types.

Considerable segmentation information available in the images are attributed to tone, however, the pure textural information captured by both the Gabor and cooccurrence feature sets provide quite accurate image segmentation. Multi-year ice types have the highest dominant frequencies followed by first year rough and then first year smooth ice types. Visually, we include tone in the discrimination. However, the brighter backscatter responses often have higher frequency components and this information is used by the cooccurrence and Gabor features for segmentation. The tonal difference between the multi-year ice and the first year ice types create an edge that generates strong responses from the high frequency Gabor filters. This response is blurred causing the segmentation to encroach from the multi-year ice types into the first year ice types. This is not something that would be expected from the HVS since boundaries should stay localized at the step boundaries between the ice types. Inclusion of tonal information should improve the boundary localization.

Barber *et al.* used the images to perform supervised texture classification com-

parisons. Classification ability always dropped when applying discriminants created from the calibration samples to samples obtained from the validation image, regardless of the technique used. Here, it is demonstrated that unsupervised segmentation is probably a better means of performing texture identification. Class samples used in the supervised approach can never be assured to represent all the variability for that class throughout the entire image. Unsupervised approaches do not require selection of samples to generate a discriminant.

The final image is presented in Figure 6.17. This C-band HH image was obtained during the Labrador Ice Margin Experiment (Limex) during springtime 1989 (8 degrees Celsius) [54]. The image has pixels with 100 metre spacing and was taken in nadir mode at incidence angles of between 50 and 70 degrees. This image contains brash ice (top left hand corner), open water (dark region in centre; with a wind blowing left to right), and first year ice (right and left hand sides).

A manual segmentation (performed by myself) is presented in Figure 6.18 for comparison to the Gabor and cooccurrence segmentations. These two methods capture different information about this image. The Gabor features are able to identify the ice floes at the edge of the first year ice as being distinct from the rest of the first year ice (Figure 6.19). These floes tend to be smaller than the floes found in the rest of the first year ice. The cooccurrence features are unable to do this because of the small window size used (Figure 6.20). The cooccurrence features are able to distinguish the regions of the open water with and without waves, something that the Gabor features do not capture. Both methods approximate the classification percentages of the manual segmentation as presented in Table 6.1.

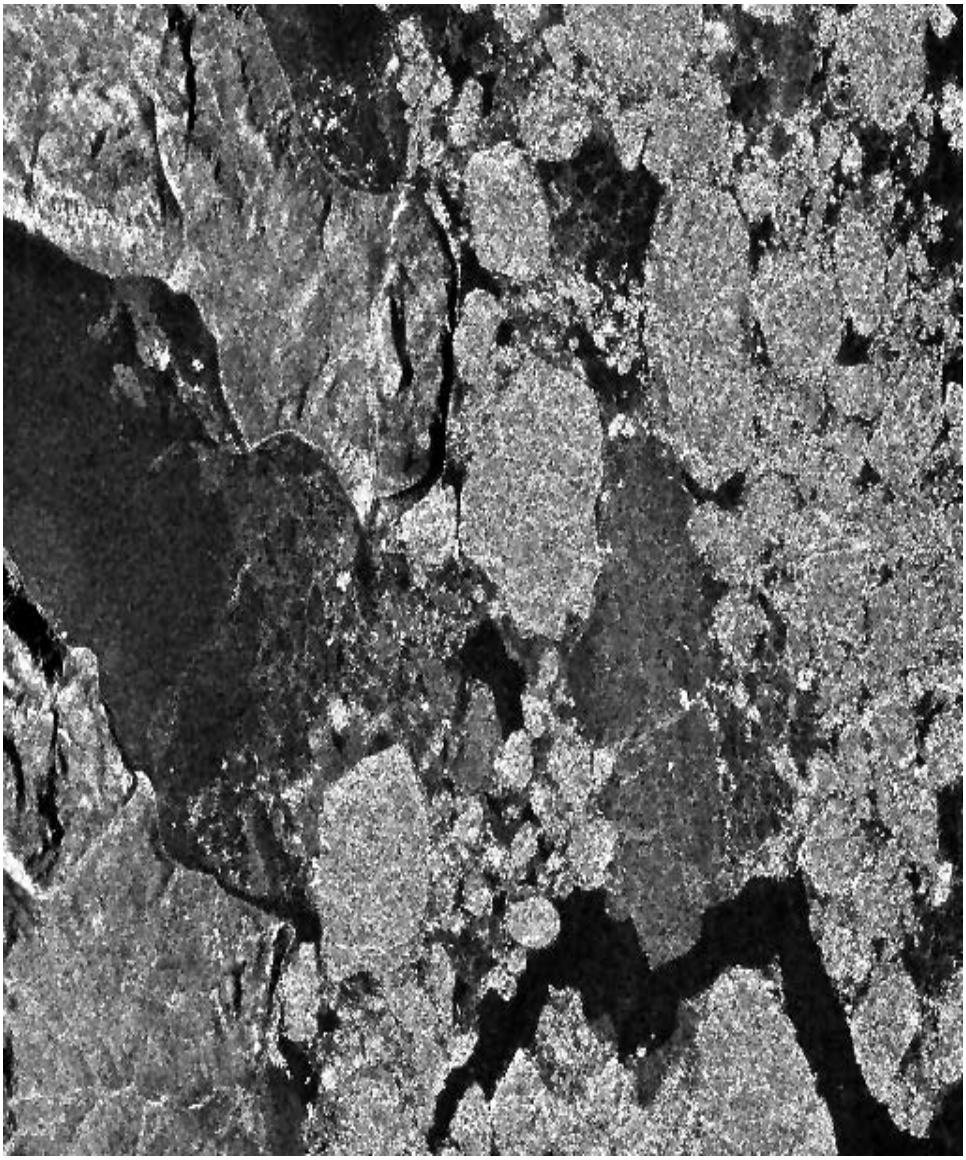


Figure 6.11: Aerial SAR image (subimage of Figure 2(a) in [8]).

	Open Water	First Year	Brash
Manually	54	34	12
Gabor features	52	35	13
Cooccurrence features	46	44	10

Table 6.1: Class assignment percentages following segmentation of Figure 6.17.

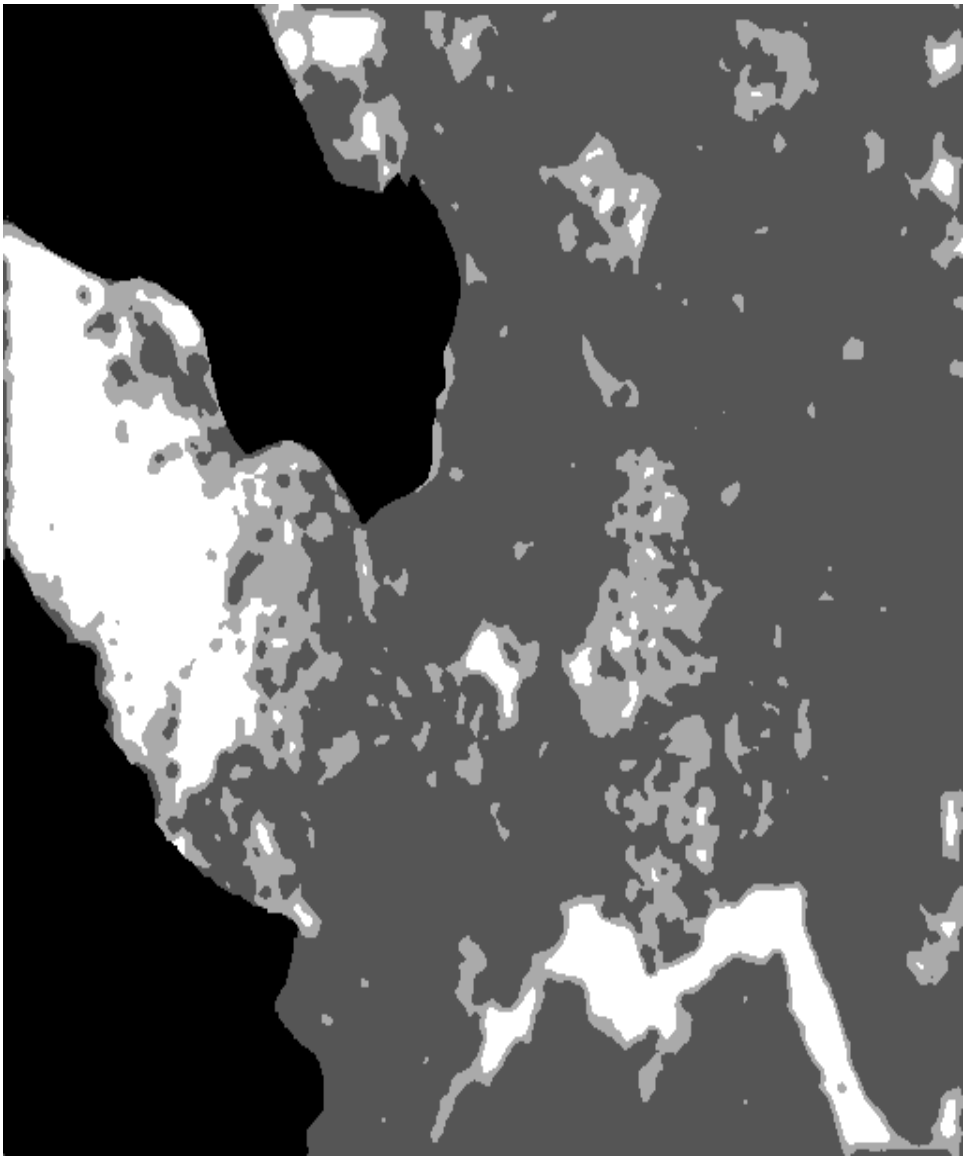


Figure 6.12: Segmented Figure 6.11 based on Gabor features.

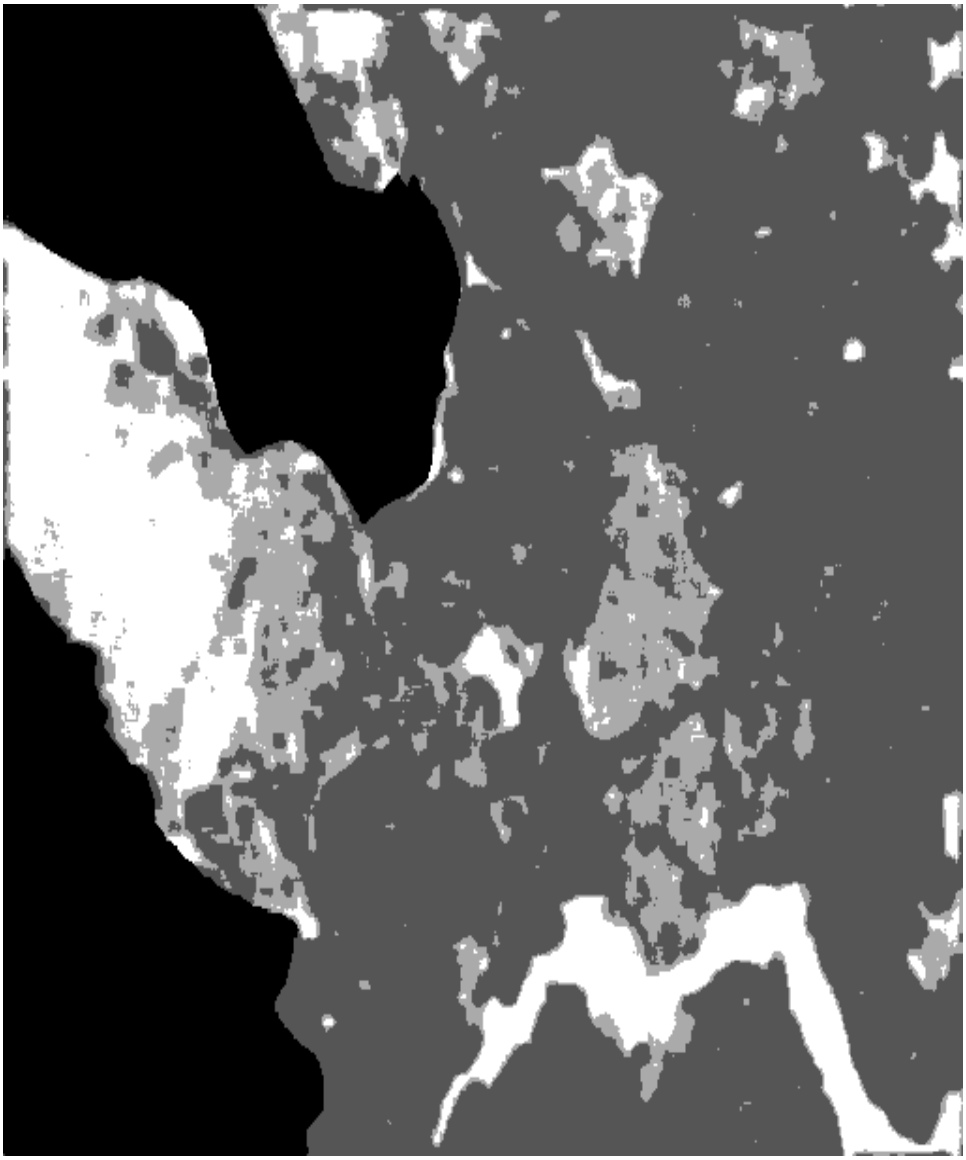


Figure 6.13: Segmented Figure 6.11 based on cooccurrence features.

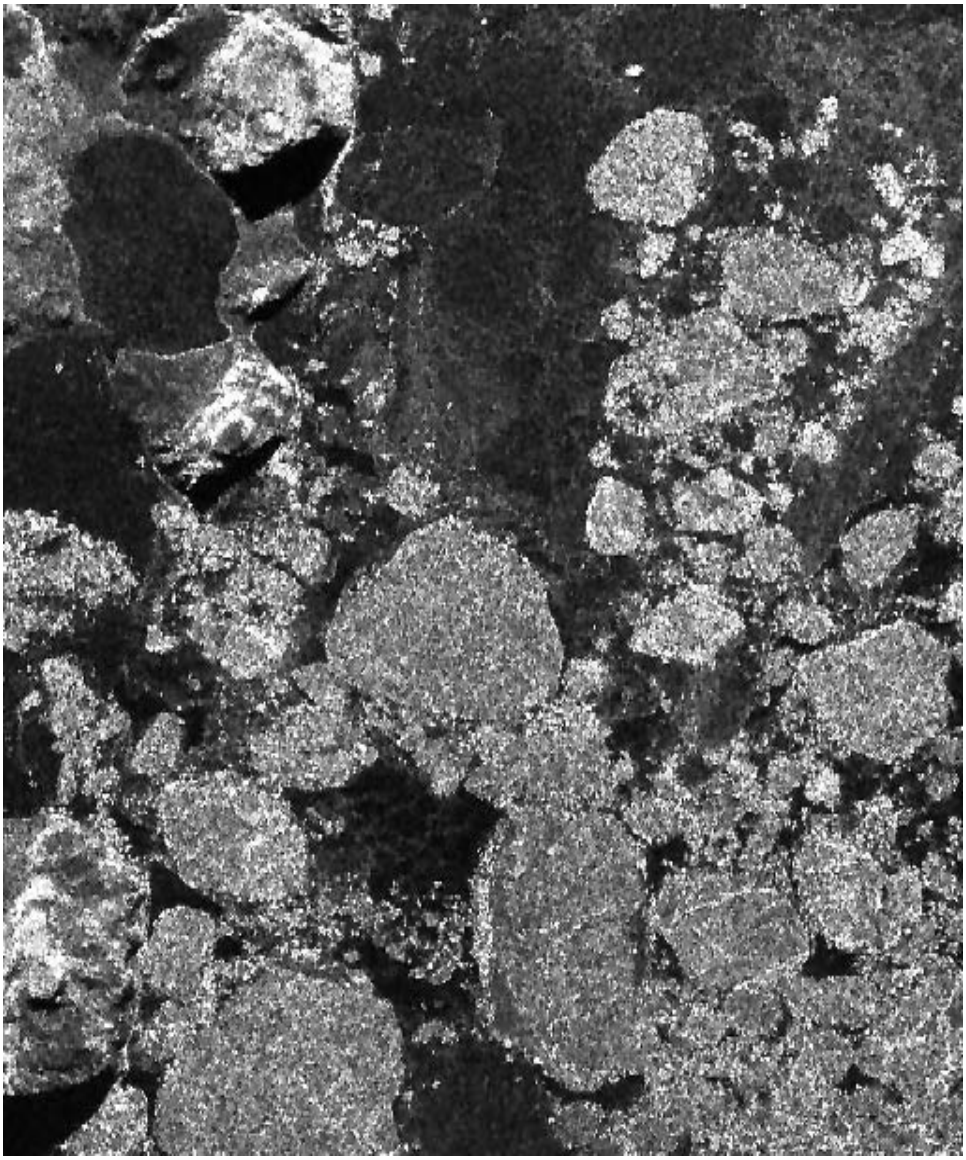


Figure 6.14: Aerial SAR image (subimage of Figure 2(b) in [8])

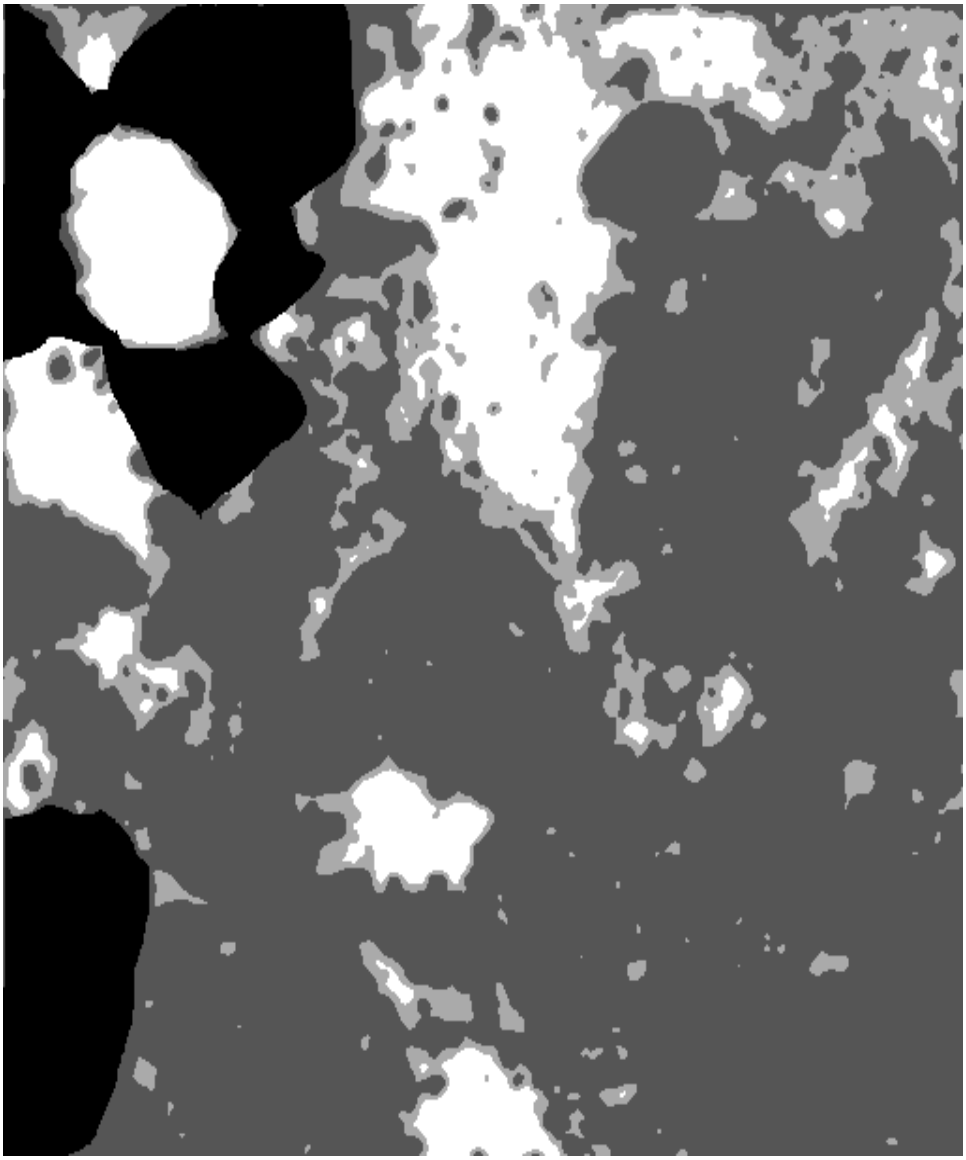


Figure 6.15: Segmented Figure 6.14 based on Gabor features.

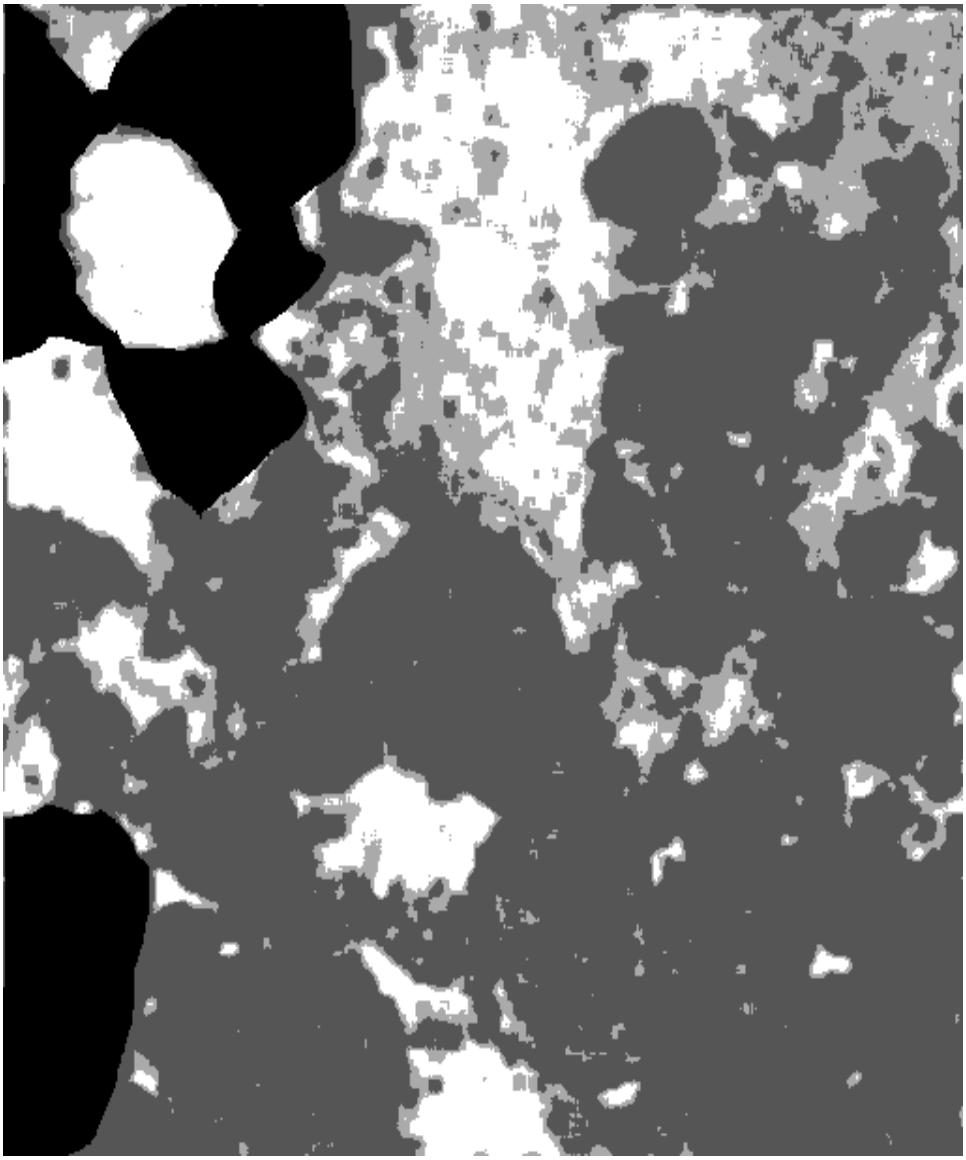


Figure 6.16: Segmented Figure 6.14 based on cooccurrence features.

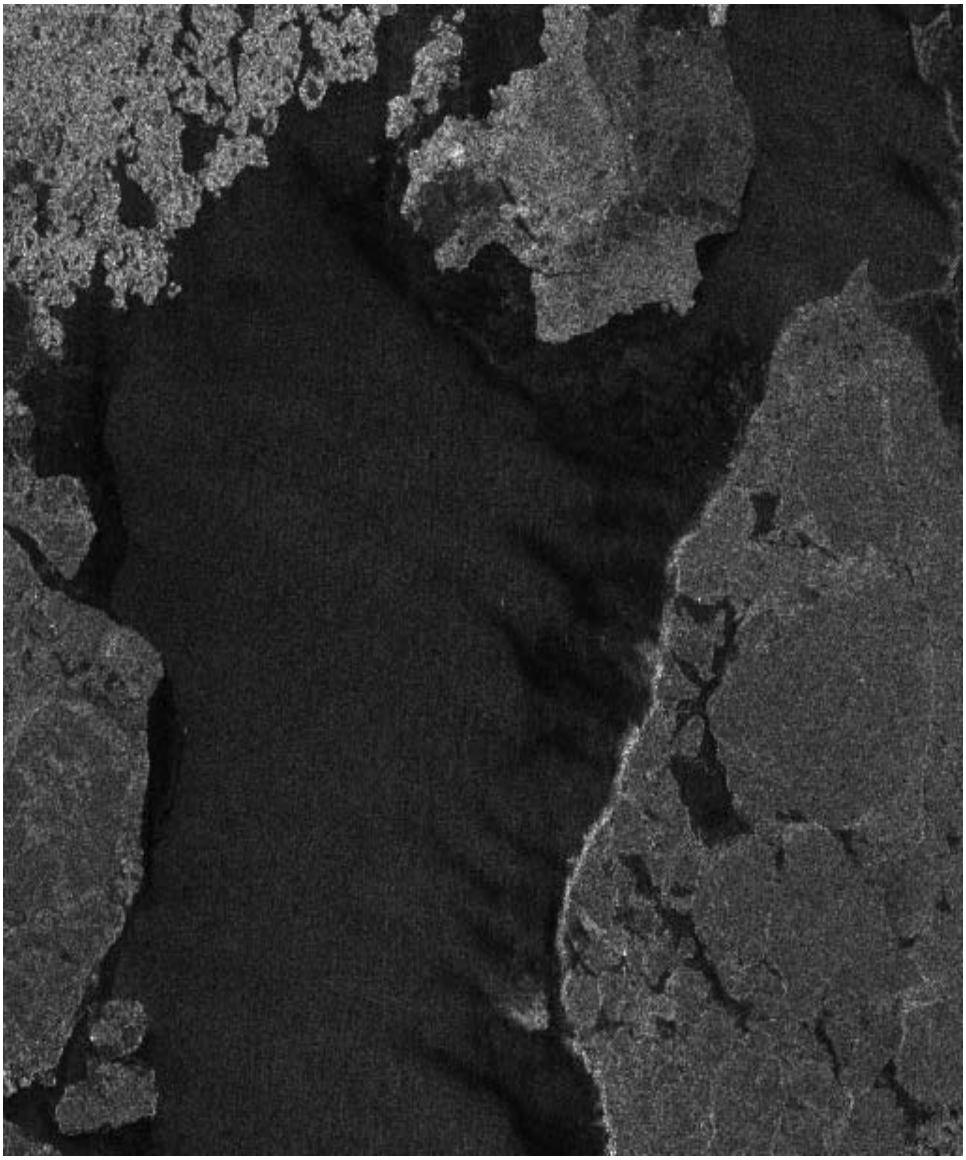


Figure 6.17: Aerial SAR image obtained from [54].

Figure 6.18: Manual segmentation of Figure 6.17.

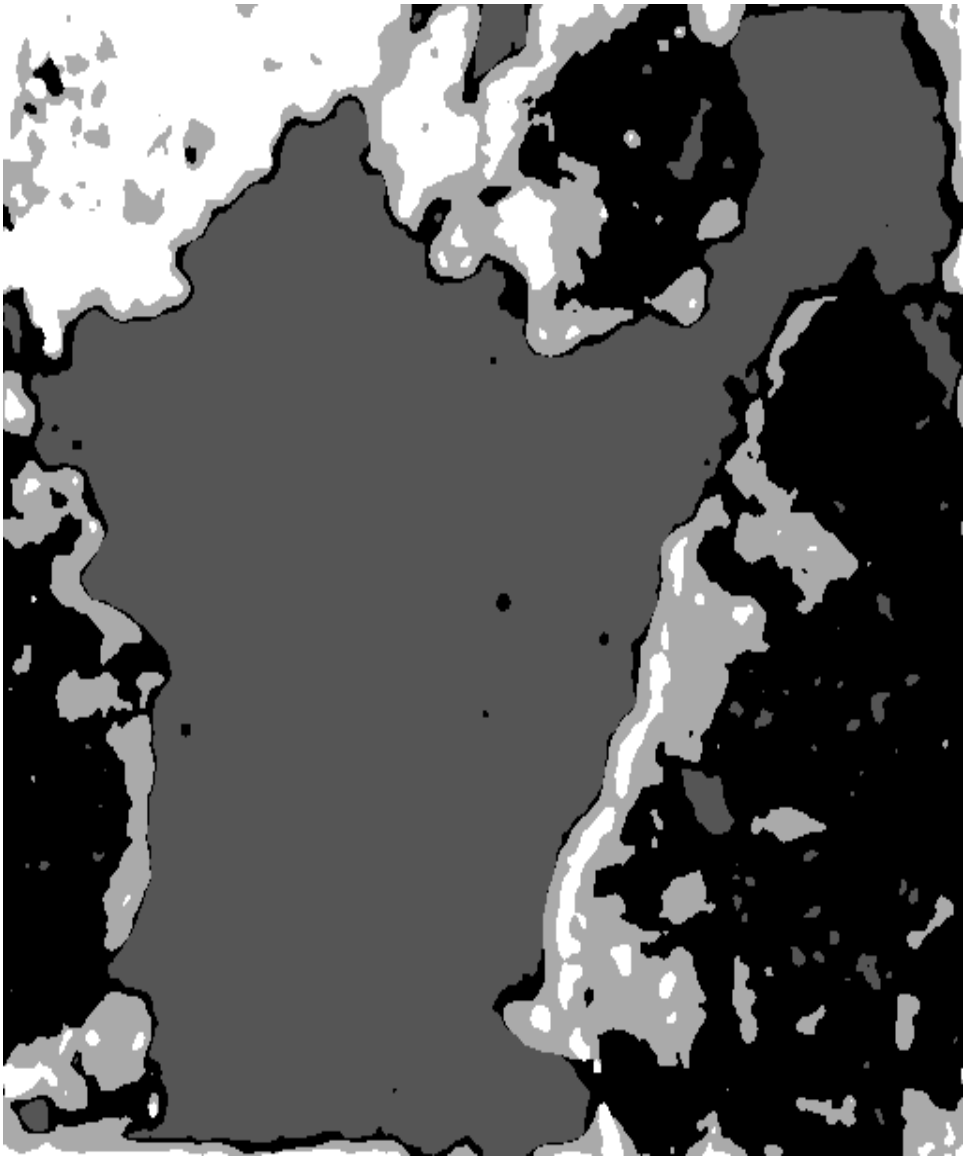


Figure 6.19: Segmented Figure 6.17 based on Gabor features.

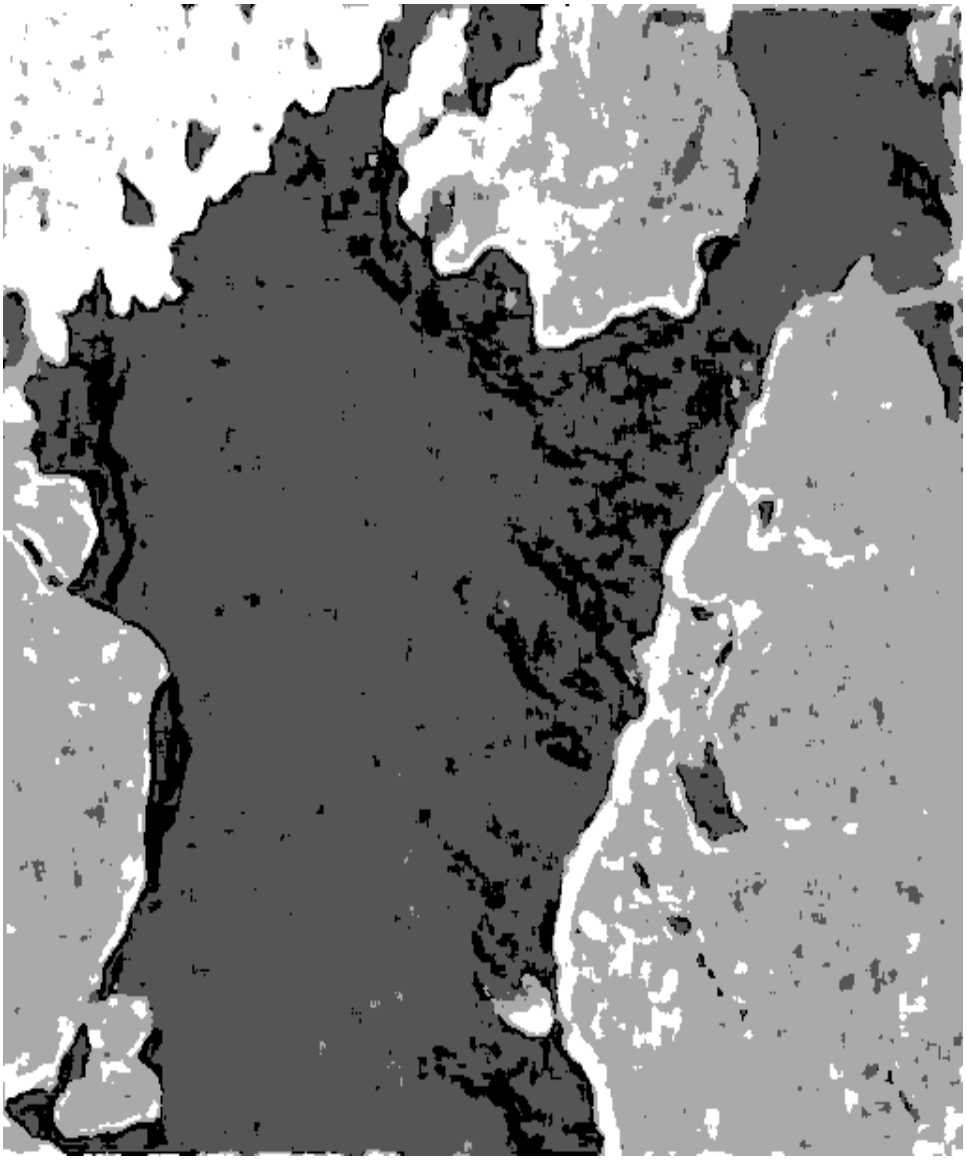


Figure 6.20: Segmented Figure 6.17 based on cooccurrence features.

	8x8	16x16	32x32
w/o updating	138	633	3601
w/ updating	74	219	555

Table 6.2: Completion times (seconds) for cooccurrence feature set collected for Figure 6.5a.

6.2.3 Algorithm Speeds

Image segmentation speeds are compared using an IBM RISC System/6000 Model 43P (64Mb RAM, 100MHz, SPECint92 - 128.1, SPECint92 - 120.2). In Section 5.1.3, speeds are compared using a single class of high frequency data. Here, a complete 4-class Brodatz image (Figure 6.5a) is used to compare algorithm completion times.

Completion times for generating the cooccurrence texture features without and with updating are presented in Table 6.2. Although only the features for the 16x16 window are used for segmentation, features are calculated for window sizes 8x8 and 32x32 just to compare the completion times. As in Section 5.1.3, using the updating approach is quite advantageous and becomes more important with increasing window size.

Clustering times for mixture analysis are presented in Table 6.3. Since the Gabor feature set has a higher dimensionality (24-d versus 12-d), the clustering takes longer. The dimensionality is especially important for the iFLD algorithm, where inverses of the S_w matrix (dimensioned to the feature space size) must be calculated (Section 4.2.2). Gabor features are better able to quickly find the centroids of the clusters (15 iterations instead of 38). A frequency band or two could be removed from the Gabor feature set to segment this image and obtain similar results. This would reduce the completion times to better approximate those obtained for the

		Time (seconds)	Iterations
Gabor			
	Features	127	-
	K-means	72	15
	iFLD	141	2
	Total	340	-
Cooccurrence			
	Features	219	-
	K-means	93	38
	iFLD	41	2
	Total	353	-

Table 6.3: Completion times for unsupervised segmentation using Gabor and co-occurrence feature sets collected from Figure 6.5a.

cooccurrence features. Since it is desirable to utilize a consistent Gabor feature set that can be applied to a wide variety of image types, results for the full 24-d feature space are presented. The unsupervised segmentations require more computing time: 859s for the Gabor features and 462s for the cooccurrence features. Not only must time be spent finding the appropriate clusters, but a considerable amount of time (one-quarter for Gabor, one-sixth for cooccurrence) is spent determining whether or not the final clusters are truly single classes. This takes considerable time because the iFLD will try to divide one class into two clusters and will often reach the maximum number of iterations (5) before stopping.

Since the amount of time required for clustering is a function of the distinctiveness and number of the classes in the image, a theoretical prediction of the computational order is not appropriate.

6.3 Discussion of Segmentation Results

6.3.1 Comparison to Other Texture Segmentation Comparative Studies

Research publications comparing different texture segmentation approaches are few. Many publications discuss novel individual segmentation schemes, but there is a noticeable lack of comparative studies. Reed and du Buf review recent texture segmentation approaches in a summary paper [92]. (The popular paper by Haralick [51] describes general texture interpretation without an emphasis on segmentation.) Given that Gabor and cooccurrence techniques are two popular texture analysis methods, there has been very little research published to compare their ability for image segmentation. Three such papers have been noted for discussion here.

Strand and Taxt [101] compare Gabor and cooccurrence texture segmentation with their new method (one that essentially calculates a local frequency measurement). The Gabor and cooccurrence techniques are implemented ineffectively. For example, the Gabor approach uses some centre frequencies that are clearly too low to be helpful for discrimination given the textures used in their study ($\sqrt{2}$cpi and $2\sqrt{2}$cpi for 256x256 images). Also, the two highest octave frequency bands ($5\sqrt{2}$cpi and $6\sqrt{2}$cpi) are not included. These would have definitely assisted the segmentation since the cotton and herringbone Brodatz textures are better recognized using these higher centre frequencies. For both the cooccurrence and Gabor approaches, the feature sets are restricted to only 0 and 90 degrees orientation, even though the Brodatz herringbone texture (D17) (which has strong 45 and 135 degree features) is used in the test image. The number of grey levels is reduced to eight for the

cooccurrence approach, possibly destroying important textural information.

Du [29] compares texture segmentation of remotely sensed imagery using Gabor filters and cooccurrence measures. His results are similar to those obtained in Section 6.2.2: both Gabor and cooccurrence features are able to identify the pertinent aspects of the texture. Gabor results tend to blur the boundaries while cooccurrence features preserve finer detail. Regions are better approximated by Gabor features than coocurrence features. Du's approach is highly supervised since not only are samples trained on, but the Gabor filter parameters are based on Fourier analysis of the training sets.

Dubuf *et al.* [29] compare seven different approaches for their ability to perform texture segmentation. Cooccurrence measures represent one of the methods. Laws masks [70] are another method used. Some of the Laws masks have shapes that are similar to Gabor filters, however, Laws masks do not have the design flexibility that the Gabor filters offer. Spann and Wilson's [99] unsupervised segmentation algorithm is used. This method identifies the boundaries between textures, but does not seem to simultaneously group similar textural regions together. For the SAR sea ice segmentation problem, additional processing would be required to group regions of similar textural characteristics. Dubuf *et al.* describe quantitative measures for comparing different segmentations. Their methods demonstrated a tendency for cooccurrence and Gabor-like features to be better than the other methods. It is difficult to compare their results to the current study since they only used bipartite images for their testing. Visual observations are used in this thesis to compare different image segmentation tests.

6.3.2 Summary of Segmentation Findings

The KIF method to cluster feature vectors was developed and implemented. Results of applying the technique to Gabor texture features, both in mixture resolving and unsupervised clustering modes have been, in general, successful. If the clusters are well separated in feature space, then the K-means algorithm is able to find the class centroids. Since clusters do not typically have equal, isotropic covariance matrices, the Fisher linear discriminant is used to correct for errors generated by the K-means algorithm. Iteratively applying the Fisher linear discriminant further improves the segmentation. Other algorithms using the Gabor filters have not always accounted for the disparity in covariance matrices from cluster to cluster and this is demonstrated to be an important aspect for proper segmentation.

The unsupervised approach requires setting the parameter τ. This is not unreasonable because the computer must be told what humans consider to be similar or dissimilar. Setting τ to a value of 10 is a reasonable estimate for unsupervised segmentation of the Brodatz imagery. Different imagery may require different values of τ. For this reason, plus the fact that expert information must be used when segmenting SAR imagery, the hierarchical assisted segmentation scheme is proposed. This method is successful for the operational segmentation of SAR sea ice imagery.

Both Gabor and cooccurrence features have the ability to identify texture, however, they have different aptitudes. These differences will be summarized in three ways: the ability for the Gabor features to interpret multiresolutional textures, the clusterability of the Gabor features, and the ability for cooccurrence measures to capture detailed information.

Gabor features are better for solving the general texture segmentation problem since they are able to distinguish patterns with different textural resolutions. This

is clearly represented by the results in Figures 6.7 and 6.9. Cooccurrence features are not appropriate for segmenting images with multiple textural resolutions.

Cooccurrence measures require a fixed window size. This is not the reason why they are not capable of multi-resolutional analysis. For example, one could select multiple window sizes. However, when the window size is too small for a particular texture in the image, erroneous feature measurements are made which impedes clustering. When the window size is too large, features at boundaries are not accurate. This not only affects the boundary estimation, but contributes to confusion when clustering.

Consider two distinct textures (A and B) in an image segmentation context. Assume Gabor filter (F, θ) is tuned to texture A, but not to texture B. Thus, a strong response is invoked using (F, θ) while within texture A. As the texture boundary is approached, the filter (F, θ) begins to decrease its response as its spatial extent begins to overlap the texture boundary. The filter response of (F, θ) will be relatively low when the spatial extent is fully within texture B. Thus, ideally, a well defined step edge appears between textures A and B. The cooccurrence data behaves in a different manner. For a given orientation and statistic (S, θ) texture A has an expected response and texture B has another expected response. The magnitude of the response is not necessarily dependent on the texture - it is only hoped that the texture response is consistent for texture A and texture B. As the window moves across the textural boundary, it is uncertain how the given (S, θ) will respond due to the mixture of textures. Although the response is expected to be consistent when the window is completely in A or completely in B, the response when crossing the boundary is unknown. This unexpected response may represent the response of another texture in the image and this will confuse any unsupervised texture segmentation scheme.

In the case of supervised classification using cooccurrence measures, the discriminant can be trained to emphasize the informative features. The confusion rising from the uncertain features will not play a role in the classification. Thus, cooccurrence features are strong when classifying the data that the discriminant is trained on, but these types of analyses have no practical operational purpose (why bother classifying data that already have class assignments?).

These observations suggest that, in general, Gabor features are more clusterable than the cooccurrence features. The magnitude of each Gabor filter output is relative to the other filter outputs. The cooccurrence features do not have these characteristics which impedes clustering.

The cooccurrence features do excel at capturing fine details, as indicated by the results of the SAR imagery. The Gabor technique requires smoothing of the feature maps, blurring textural boundaries belonging to small regions. The cooccurrence features do not require smoothing and capture fine boundary details when the textures are distinct with the same textural resolution. Also, the cooccurrence probabilities treat every pixel in the window with the same weight. For a small window size, this may be appropriate.

Chapter 7

Summary of Results and Their Implications

7.1 Summary of Results and Research Contributions

This thesis has generated a variety of results. These have been summarized in detail at the end of each of the two testing chapters, Chapters 5 and 6. Here, a summary of the major results and their contribution to the research literature are presented.

- Design and implementation of a fast procedure for capturing cooccurrence texture features.

  The cooccurrence texture features have been in use for over two decades. A severe restriction of the method for operational applications is the computational demands of the matrix (GLCM) approach for calculating the cooccurrence probabilities and features. Although cooccurrence features are popular

in the remote sensing community, no commercial software package is known that calculates the cooccurrence texture features in a computationally efficient manner. The linked list implementation in this thesis has exceptional computational advantages over the traditional matrix approach and is thus important to the remote sensing community.

- Investigation of parameters for three texture analysis methods. Generation of preferred parameters for the two most successful approaches.

  The linked list approach provides the opportunity for rapid indepth testing of the cooccurrence features. For example, testing of statistics across different quantization levels generated insights into their abilities. There does not seem to be any other research that has performed this particular type of analysis. Also, features that are determined to be highly correlated and thus redundant are removed from the feature set. The end result is a more robust feature set requiring significantly less time to calculate.

  Power spectrum features generated texture features that had poorer classification ability compared to the cooccurrence measures.

  Different Gabor filter parameters and manipulations of the filter responses were compared analytically for their classification ability. This helps to determine which manipulations is more appropriate for the task at hand. Here, a successful methodology has been identified (Section 5.4.1) for implementing Gabor filters for texture identification.

- Identification of cooccurrence and Gabor measures that agree with psychovisual characteristics of the HVS.

  Rao and Lohse [90] identified repetition, directionality, and complexity as important psychovisual characteristics necessary for texture identification.

These characteristics describe the most critical texture features generated by the cooccurrence and Gabor features. In fact, the cooccurrence and Gabor features tuned to a texture's dominant orientation and frequency will be strongly correlated. This indicates that the important information captured by these two methods are quite similar, even though both use completely different approaches. Both, in essence, generate a local estimate of the directional frequency.

- Design and implementation of an unsupervised texture segmentation scheme (KIF).

  K-means generates a reasonable texture segmentation for distinct textures. Textures that are not so distinct require additional information to segment properly. Some textures have variations in their dominant frequency and orientation textural content, making them difficult to identify. These complex textures require special handling since the covariances of their feature dimensions must be taken into account. This is performed using a class-pairwise iterative Fisher linear discriminant. From an unsupervised perspective, determining the covariances is quite difficult. Here, the clusters generated by the K-means algorithm are successfully used as a starting point for the iterative Fisher linear discriminant.

- Detailed comparison of cooccurrence and Gabor texture features for unsupervised segmentation.

  Gabor filters and cooccurrence probabilities generate similar information using different techniques, however, their segmentation capabilities are different. Gabor filters are able to capture information in a multi-resolutional manner. If an image contains patterns with different textural resolutions, the Gabor

filter can easily capture this type of information while maintaining spatial localization. The cooccurrence features are unable to capture features in a multi-resolutional fashion.

The cooccurrence features are able to capture detailed information, which is important in the context of natural imagery. The cooccurrence measures are perhaps better suited for supervised classification than the Gabor filters. The cooccurrence measures have a finite extent and treat every pixel in the same window (which is already assumed to contain only one texture only) in the same manner. The Gabor filters weight the contribution of the pixels as a function of distance from the centre pixel, a better approach for image segmentation.

- Design and implementation of an operational approach for the segmentation of SAR sea ice imagery.

  The unsupervised approach to image segmentation is not necessarily appropriate for all SAR sea ice imagery under all circumstances. In addition to texture, various other aspects of the imagery must be taken into account. The divisive hierarchical model is appropriate as a tool for assisting the operational segmentation of SAR imagery.

7.2 Future Research Directions

7.2.1 Improve Accuracy of Texture Features for Segmentation

Cooccurrence Features

Cooccurrence features become confused at textural boundaries. To try to improve segmentation near boundaries, the cooccurring probabilities can be weighted by a Gaussian centred on the window instead of uniformly weighting all the cooccurring grey level probabilities. The probability contributed by grey level pairs that are closer to the centre of the window would be emphasized and those near the edge of the window would be deemphasized. This might allow for improved classification near boundaries.

The cooccurrence data has always considered only the probability of two occurring pixels. What type of information can be derived from n-occurring pixels? This would create opportunities to generate diverse texture features based on multidimensional space as opposed to a 2-dimensional space. This suggestion is not feasible for the matrix approach because of the excessive computational demands, but it is extendable using linked lists with a moderate increase in the computational costs. If a more robust texture feature set will be created is uncertain. Since it is more unlikely that n-occurring pairs would repeat within the same window, the homogeneity features will probably be ineffective. Moment features would have to be determined with respect to the n-dimensional diagonal. The reliability of these estimates is unknown.

Gabor Features

The Gabor feature set determined by evenly covering the spatial-frequency domain creates a high dimensional feature space that requires substantial computing time for unsupervised clustering. However, only a few distinct filters are required for robustly identifying any particular texture. Devising a scheme that finds an optimal set of filter responses for a local spatial region is appealing. If an accurate local estimate of the dominant Gabor filters can be made, then adjacent local regions can be determined to be either the same or different from the current region. This would allow the texture to "grow", in much the same manner that the HVS would track and identify a consistent texture within an image. I have spent significant time attempting this type of approach. Many problems are associated with this design for unsupervised segmentation: determining the number of distinct filters required, determining whether the spatial region is within a constant texture region or near a texural boundary, and estimating the statistics of the filter parameters. Promising results were obtained for bipartite images with distinct textures, however, generating a robust solution for diverse imagery proved to be elusive.

Linear Gaussian smoothing may not be the optimal choice for smoothing the Gabor feature maps since textural boundaries are smoothed which can affect the segmentation (see Figure 6.8c, for example). An adaptive smoothing of the Gabor feature maps may be more appropriate. Smoothing the constant regions and avoiding edge smoothing should prevent modifying the feature measurements near the textural boundaries. There is also a significant increase in the computational time required for feature extraction if adaptive smoothing is employed.

Other Texture Features

The research literature lacks comparative analyses between different texture segmentation approaches. This thesis compares two statistical based methods for their segmentation performance. Model based approaches have also been demonstrated to be successful for texture segmentation. Comparison of a model based approach with the Gabor and cooccurrence methods would be enlightening.

The Gabor and cooccurrence features may be quite effective when combined. Gabor filters are important when multi-resolutional textures are involved. However, the cooccurrence probabilities are more effective at capturing high frequency features from isotropic imagery. If the Gabor feature set is augmented with a small window isotropic cooccurrence feature (one that averages all four directions into a single measure), then this combined feature set may be more effective for solving the general texture segmentation problem. Another way to handle this problem is to average the Gabor responses of all the orientations of the highest frequency band and use this as an additional texture feature.

7.2.2 Improve Computations

The linked list approach for generating cooccurrence texture features is a tremendous improvement over the traditional matrix approach. The linked list is a burden for searching since a linear search of the list is required to determine whether or not a grey level pair exists. The linked list approach may be further improved by using a tree structure [96] which would reduce the amount of time required for searching. Additional memory would also be used.

Clustering feature vectors derived from large images is unwieldly. The Equitz algorithm [39] can reduce the number of feature vectors by clustering similar feature

vectors. The data set could be reduced by a certain percentage of the original. Each of the features in the reduced feature set would also be represented by the number of samples assigned to it based on the Equitz algorithm. Running K-means and iterative Fisher linear discriminant on the reduced data set would accelerate the execution times.

7.2.3 Other Applications

Further potential of the segmentation scheme can be exploited by segmenting other sources of SAR imagery ie. forestry, agricultural crops, etc. As long as there is some noticeable textural distinction between classes, Gabor and cooccurrence measures can generate useful discriminating information. The clustering algorithm can also be attempted to solve other clustering problems. As long as the class centroids can be identified by K-means, this clustering algorithm should be successful.

7.2.4 Create Standard Texture Analysis Image Library

Texture analysis is at a stage of development whereby a universal set of test images is important. Computer implementations for practical applications started well over twenty years [52] ago. Methods for seriously considering the fully unsupervised approach initiated about one decade ago [91]. Current research indicates an abundance of texture analysis publications, however, few scientists test methodologies on images previously used for texture segmentation analysis. Whether a method is useful for generic texture identification or whether it is only useful for certain images is difficult to ascertain. Creating a texture image library that scientists can all access and use will enable these types of comparisons.

To enable a test bed of appropriate images, different application areas would first be identified. Remotely sensed, biomedical, theoretical, and surface inspection imagery are suggested applications areas. Then each method would have subareas of interest. For example, remotely sensed imagery could contain images dealing with different sensors (aerial, SAR, visible, satellite, infrared, etc.) and different terrain (sea ice, forestry, agriculture, etc.). Theoretical images would include Brodatz, Markov modelled, and fractal collages. Contributions from scientists in different research fields would assist generation of a robust set of images that can be used to consistently compare various texture segmentation algorithms.

Bibliography

[1] M.F. Augusteijn, Clemens L.E., and K.A. Shaw. Performance evaluation of texture measures for ground cover identification in satellite images by means of a neural network classifier. *IEEE Trans. Geosci. Remote Sensing*, 33(3):616–626, May 1995.

[2] A. Baraldi and F. Parmiggiani. An investigation of the textural characteristic associated with gray level cooccurrence matrix statistical parameters. *IEEE Trans. Geosci. Remote Sensing*, 33(2):293–304, Mar 1995.

[3] D.G. Barber, D.D. Johnson, and E.F. LeDrew. Measuring climatic state variables from SAR images of sea ice: The SIMMS SAR validation site in Lancaster Sound. *Arctic*, 4(1):108–121, 1991.

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

[5] D.G. Barber and E.F. LeDrew. The Seasonal Sea Ice Monitoring and Modelling Site (SIMMS) '91 Science Plan. Technical Report ISTS-EOL-SIMS-TR91-001, Earth Observation Laboratory, Institute for Space and Terrestrial

Science, and the Department of Geography, University of Waterloo, Waterloo, Ontario, Canada, N2L 3G1, 1991.

[6] D.G. Barber and E.F. LeDrew. The Seasonal Sea Ice Monitoring (SIMMS) '93 Science Plan. Technical Report ISTS-EOL-SIMS-TR93-004, Earth Observation Laboratory, Institute for Space and Terrestrial Science, and the Department of Geography, University of Waterloo, Waterloo, Ontario, Canada, N2L 3G1, 1993.

[7] D.G. Barber, M.J. Manore, T A. Agnew, H. Welch, E.D. Soulis, and E.F. LeDrew. Science issues relating to marine aspects of the cryosphere: Implications for remote sensing. *Can. J. of Remote Sensing*, 18(1):46–55, Jan 1992.

[8] D.G. Barber, M.E. Shokr, R.A. Fernandes, E.D. Soulis, and D.G. Flett. A comparison of second-order classifiers for SAR sea ice discrimination. *Photogrammetric Engineering and Remote Sensing*, 59(9):1397–1408, Sept 1993.

[9] M.G. Bello. A combined Markov random field and wave-packet transform-based approach for image segmentation. *IEEE Trans. Image Processing*, 3(6):834–846, Nov 1994.

[10] J. Bigun and J.M. du Buf. N-folded symmetries by complex moments in Gabor space and their application to unsupervised texture segmentation. *IEEE Trans. Pattern Anal. Machine Intell.*, 16(1):80–87, Jan 1994.

[11] C. Bouman and B. Liu. Multiple resolution segmentation of textures images. *IEEE Trans. Pattern Anal. Machine Intell.*, 13(2):99–113, Feb 1991.

[12] A.C. Bovik. Analysis of multichannel narrow-band filters for image texture segmentation. *IEEE Trans. Signal Processing*, 39(9):2025–2043, Sept 1991.

[13] A.C. Bovik, M. Clark, and W.S. Geisler. Multichannel texture analysis using localized spatial filters. *IEEE Trans. Pattern Anal. Machine Intell.*, 12(1):55–73, Jan 1990.

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

[15] F.W. Campbell and J.J. Kulikowski. Orientational selectivity of the human visual system. *J. Physiol.*, 187:437–445, 1966.

[16] F.W. Campbell and J.G. Robson. Application of Fourier analysis to the visibility of gratings. *J. Physiol.*, 197:551–565, 1968.

[17] F. Carsey. Review and status of remote sensing of sea ice. *IEEE J. Oceanic Eng.*, 14(2):127–138, Apr 1989.

[18] T. Chang and C.-C.J. Kuo. Texture analysis and classification with tree-structured wavelet transform. *IEEE Trans. Image Proc.*, 2(4):429–441, Oct 1993.

[19] R. Chellappa and S. Chatterjee. Classification of textures using Gaussian Markov random fields. *IEEE Trans. Acoustics, Speech, and Signal Processing*, 33(4):959–963, Aug 1985.

[20] J-L. Chen and A. Kundu. Rotation and gray scale transform invariant texture identifiation using wavelet decomposition and hidden Markov model. *IEEE Trans. Pattern Anal. Machine Intell.*, 16(2):208–214, Feb 1994.

[21] L.L. Chen and A. Kundu. Unsupervised texture segmentation using multichannel decomposition and hidden Markov models. *IEEE Trans. Image Processing*, 4(5):603–619, May 1995.

[22] D.A. Clausi and M.E. Jernigan. A fast method to determine cooccurrence texture features using a linked list implementation. In *26th Symposium on Remote Sensing of Environment and 18th Annual Symposium of the Canadian Remote Sensing Society*, Mar 1995.

[23] J.M. Coggins and A.K. Jain. A spatial filtering approach to texture analysis. *Pattern Recognition Letters*, 3(3):195–203, May 1985.

[24] F.S. Cohen, Fan. Z., and M.A. Patel. Classification of rotated and scaled texture images using Gaussian Markov random field models. *IEEE Trans. Pattern Anal. Machine Intell.*, 13(2):192–203, Feb 1991.

[25] R.W. Conners and C.A. Harlow. A theoretical comparison of texture algorithms. *IEEE Trans. Pattern Anal. Machine Intell.*, 2(3):204–222, May 1980.

[26] F. D'Astous and M.E. Jernigan. Texture discrimination based on detailed measures of the power spectrum. In *Proc. IEEE Comput. Society Conf. on Pattern Recognition and Image Processing*, pages 83–86, 1984.

[27] F.T. D'Astous. *Textural Feature Extraction in the Spatial Frequency Domain.* PhD thesis, Department of Systems Design Engineering, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1, 1983.

[28] J.G. Daugman. 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, July 1985.

[29] Li-Jen Du. Texture segmentation of SAR images using localized spatial filtering. In *IGARSS*, pages 1983–1986, 1990.

[30] J.M.H. du Buf. Gabor phase in texture discrimination. *Signal Processing*, 21:221–240, 1990.

[31] J.M.H. du Buf and P. Heitkamper. Texture features based on Gabor phase. *Signal Processing*, 23:227–244, 1991.

[32] J.M.H. du Buf, M. Kardan, and M. Spann. Texture feature performance for image segmentation. *Pattern Recognition*, 23(3/4):291–309, 1990.

[33] R.O. Duda and P.E. Hart. *Pattern Classification and Scene Analysis*. John Wiley and Sons, Toronto, 1st edition, 1973.

[34] I.S. Duff, A.M. Erisman, and J.K. Reid. *Direct Methods for Sparse Matrices*. Clarendon Press, London, 1st edition, 1986.

[35] D. Dunn and W.E. Higgins. Optimal Gabor filter for texture segmentation. *IEEE Trans. Image Processing*, 4(7):947–964, July 1995.

[36] D. Dunn, W.E. Higgins, and J. Wakeley. Texture segmentation using 2-d Gabor elementary functions. *IEEE Trans. Pattern Anal. Machine Intell.*, 16(2):130–149, Feb 1994.

[37] C. Elachi. *Introduction to the Physics and Techniques of Remote Sensing*. John Wiley and Sons, Toronto, 1st edition, 1987.

[38] D.T. Eppler and L.D. Farmer. Texture analysis of radiometric signatures of new sea ice forming in arctic leads. *IEEE Trans. Geosci. Remote Sensing*, 29(2):233–241, Mar 1991.

[39] W.H. Equitz. A new vector quantization clustering algorithm. *IEEE Trans. Acoustics, Speech, and Signal Processing*, 37(10):1568–1575, Oct 1989.

[40] F. Farrokhnia. *Multi-Channel Filtering Techniques for Texture Segmentation and Surface Quality Inspection.* PhD thesis, Michigan State University, Department of Electrical Engineering, East Lansing, MI, 1990.

[41] R.A. Fernandes and M.E. Jernigan. Unsupervised multi-level segmentation of multispectral images. In *Neural Networks for Signal Processing*, volume II, pages 363–372. 1992 IEEE-SP Workshop, 1992.

[42] M. Fily and A. Rothrock. Extracting sea ice data from satellite SAR imagery. *IEEE Trans. Geosci. Remote Sensing*, GE-24(6):849–854, Nov 1986.

[43] R.A. Fisher. The use of multiple measurements in taxonomic problems. In *Contributions to Mathematical Sciences.* John Wiley and Sons, New York, 1950.

[44] J.M. Francos, A.Z. Meiri, and B. Porat. A unified texture model based on a 2-d Wold-like decomposition. *IEEE Trans. Signal Processing*, 41(8):2665–2678, Aug 1993.

[45] J.M. Francos, A. Narasimhan, and J.W. Woods. Maximum likelihood parameter estimation of textures using a Wold-like decomposition based model. *IEEE Trans. Image Processing*, 4(12):1655–1666, Dec 1995.

[46] S.E. Franklin and D.R. Peddle. Texture analysis of digital image data using spatial cooccurrence. *Computers and Geosciences*, 13(3):293–311, 1987.

[47] K. Fukunaga. *Introduction to Statistical Pattern Recognition.* Computer Science and Scientific Computing. Academic Press, Inc., Toronto, 2nd edition, 1990.

[48] R.C. Gonzalez and R.E. Woods. *Digital Image Processing.* Addison-Wesley, Toronto, 1993.

[49] C.F. Hall and E.L. Hall. A nonlinear model for the spatial characteristics of the human visual system. *IEEE Trans. Syst., Man, Cybern.*, SMC-7(3):161–170, Mar 1977.

[50] T.E. Hall and G.B. Giannakis. Bispectral analysis and model validation of texture images. *IEEE Trans. Image Processing*, 4(7):996–1009, July 1995.

[51] R.M. Haralick. Statistical and structural approaches to texture. *Proc. IEEE*, 67(5):786–804, May 1979.

[52] R.M. Haralick, K. Shanmugan, and I. Dinstein. Textural features for image classification. *IEEE Trans. Syst., Man, Cybern.*, SMC-3(6):610–621, Nov 1973.

[53] D. He, L. Wang, and J. Guibert. Texture discrimination based on an optimal utilization of texture features. *Pattern Recognition*, 21(2):141–146, 1988.

[54] T. Hirose, T. Heacock, and R. Duncan. A report on the evaluation of ice classification algorithms. Technical report, Prepared by Noetix Research Inc. for Environment Canada Ice Branch and Ice Applications Group - Canada Centre for Remote Sensing, October 1993.

[55] 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*, GE-22(2):113–120, Mar 1984.

[56] D.H. Hubel and T.N. Wiesel. Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat. *J. Neurophysiol.*, 28:229–289, 1965.

[57] Radarsat International Inc. Value-added opportunities in support of the Radarsat program: Market feasibility study. Technical Report 670DIP-1-1465/01-SV, Industry Science and Technology Canada, June 1993.

[58] A.K. Jain and R.C. Dubes. *Algorithms for Clustering Data.* Prentice Hall, Englewood Cliffs, New Jersey, 1st edition, 1988.

[59] A.K. Jain and F. Farrokhnia. Unsupervised texture segmentation using Gabor filters. *Pattern Recognition*, 24(12):1167–1186, 1991.

[60] J.R. Jensen. *Introductory Digital Image Processing: A Remote Sensing Perspective.* Prentice-Hall, Englewood Cliffs, New Jersey, 1st edition, 1986.

[61] M.E. Jernigan and F. D'Astous. Entropy-based texture analysis in the spatial frequency domain. *IEEE Trans. Pattern Anal. Machine Intell.*, 6(2):236–243, Mar 1984.

[62] B. Julesz. Visual pattern discrimination. *IRE Trans. on Information Theory*, IT-8:84–92, Feb 1962.

[63] B. Julesz. A theory of preattentive texture discrimination based on first-order statistics of textons. *Biol. Cybern.*, 41:131–138, 1981.

[64] B. Julesz, E.N. Gilbert, and J.D. Victor. Visual discrimination of textures with identical third-order statistics. *Biological Cybernetics*, 31:137–140, 1978.

[65] C. Kervrann and F. Heitz. A Markov random field model-based approach to unsupervised texture segmentation using local and global spatial statistics. *IEEE Trans. Image Processing*, 4(6):856–862, June 1995.

[66] S. Krishnamachari and R. Chellappa. GMRF models and wavelet decomposition for texture segmentation. In *Proceedings IEEE International Conference on Image Processing*, pages Vol.3, 568–571, 1995.

[67] A. Kundu and J. Chen. Texture classification using QMF bank-Based subband decomposition. *CVGIP: Graphical Models and Image Processing*, 54(5):369–384, Sept 1992.

[68] R. Kwok, J.C. Curlander, McConnell, and S.S. Pang. An ice-motion tracking system at the Alaska SAR facility. *IEEE J. Oceanic Eng.*, 15(1):44–54, Jan 1990.

[69] R. Kwok, E. Rignot, and B. Holt. Identification of sea ice types in spaceborne synthetic aperture radar data. *J. Geophys. Res.*, 97(C2):2391–2402, Feb 1992.

[70] K.L. Laws. *Textured Image Segmentation.* PhD thesis, Univ. Southern Calif., Los Angeles, CA, 1980.

[71] E.F. LeDrew. Polar regions, influence on climate variability and change. *Encyclopedia of Earth System Science*, 3:647–659, 1992.

[72] T.M. Lillesand and R.W. Kiefer. *Remote Sensing and Image Interpretation.* John Wiley and Sons, Toronto, 2nd edition, 1987.

[73] J.S. Lim. *2-d Signal and Image Processing.* Prentice-Hall, Englewood Cliffs, 1st edition, 1990.

[74] S. Liu and M.E. Jernigan. Texture analysis and discrimination in additive noise. *CVGIP*, 49:52–67, 1990.

[75] J.D. Lyden, B.A. Burns, and A.L. Maffett. Characterization of sea ice types using synthetic aperture radar. *IEEE Trans. Geosci. Remote Sensing*, GE-22(5):431–439, Sept 1984.

[76] J. Malik and P. Perona. Preattentive texture discrimination with early vision mechanisms. *J. Opt. Soc. Am. A*, 7(5):923–932, May 1990.

[77] S.G. Mallat. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Trans. Pattern Anal. Machine Intell.*, 11(7):674–693, July 1989.

[78] B.S. Manjunath and R. Chellappa. Unsupervised texture segmentation using Markov random field models. *IEEE Trans. Pattern Anal. Machine Intell.*, 13(5):478–482, May 1991.

[79] B.S. Manjunath and R. Chellappa. A unified approach to boundary perception: Edges, textures, and illusory contours. *IEEE Trans. Neural Networks*, 4(1):96–107, Jan 1993.

[80] B.S. Manjunath, T. Simchony, and R. Chellappa. Stochastic and deterministic networks for texture segmentation. *IEEE Trans. Acoustics, Speech, and Signal Processing*, 38(6):1039–1049, June 1990.

[81] R. McKillop. Arctic sea ice classification in SAR imagery using texture. Master's thesis, Department of Civil Engineering, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1, 1990.

[82] G.F. McLean. Vector quantization for texture classification. *IEEE Trans. Syst., Man, Cybern.*, SMC-23(3):637–649, May/June 1993.

[83] J.A. Nystuen and F.W. Garcia, Jr. Sea ice classification using SAR backscatter statistics. *IEEE Trans. Geosci. Remote Sensing*, 30(3):502–509, May 1992.

[84] P.P. Ohanian and R.C. Dubes. Performance evaluation for four classes of textural features. *Pattern Recognition*, 25(8):819–833, 1992.

[85] J.S. Paterson, B. Brisco, S. Argus, and G. Jones. In situ measurements of micro-scale surface roughness of sea ice. *Arctic*, 44(1):140–146, 1991.

[86] A.P. Pentland. Fractal-based description of natural scenes. *IEEE Trans. Pattern Anal. Machine Intell.*, 6(6):661–674, 1984.

[87] D.A. Pollen and S.F. Ronner. Visual cortical neurons as localized spatial frequency filters. *IEEE Trans. Syst., Man, Cybern.*, SMC-13(5):907–916, Sep-Oct 1983.

[88] M. Porat and Y.Y. Zeevi. Localized texture processing in vision: Analysis and synthesis in the Gaborian space. *IEEE Trans. Biomedical Engineering*, 36(1):115–129, Jan 1989.

[89] T. Randen and J.H. Husϕy. Optimal texture filtering. In *Proceedings IEEE International Conference on Image Processing*, pages Vol.1, 374–377, 1995.

[90] A.R. Rao and G.L. Lohse. Identifying high level features of texture perception. *CVGIP: Graphical Models and Image Processing*, 55(3):218–233, May 1993.

[91] T. Reed and H. Wechsler. Tracking of nonstationarities for texture fields. *Signal Processing*, 14:95–102, 1988.

[92] T.R. Reed and Du Buf J.M.H. A review of recent texture segmentation and feature extraction techniques. *CVGIP: Image Understanding*, 57(3):359–372, May 1993.

[93] T.R. Reed, H. Wechsler, and M. Werman. Texture segmentation using a diffusion region growing technique. *Pattern Recognition*, 23(9):953–960, 1990.

[94] O. Rioul and M. Vetterli. Wavelets and signal processing. *IEEE SP Magazine*, pages 14–38, Oct 1991.

[95] R. Schalkoff. *Pattern Recognition: Statistical, Structural, and Neural Approaches*. John Wiley and Sons, Inc., Toronto, 1st edition, 1992.

[96] R. Sedgewick. *Algorithms in C*. Addison-Wesley Publishing Company, Don Mills, Ontario, 1990.

[97] K.S. Shanmugan, V. Narayanan, V.S. Frost, J.A. Stiles, and J.C. Holtzman. Textural features for radar image analysis. *IEEE Trans. Geosci. Remote Sensing*, GE-19(3):153–156, July 1981.

[98] M.E. Shokr. Evaluation of second-order texture parameters for sea ice classification from radar images. *J. Geophys. Res.*, 96(C6):10625–10640, June 1991.

[99] M. Spann and R. Wilson. A quad-tree approach to image segmentation which combines statistical and spatial inforamtion. *Pattern Recognition*, 18(3/4):257–269, 1985.

[100] G. Strang. Wavelets. *American Scientist*, 82:250–255, May-June 1994.

[101] J. Strans and T. Taxt. Local frequency features for texture classification. *Pattern Recognition*, 27(10):1397–1406, 1994.

[102] K. Sun Fu. *Syntactic Pattern Recognition and Applications.* Prentice-Hall, Toronto, 1st edition, 1982.

[103] F.T. Ulaby, F. Kouyate, B. Brisco, and T.H.L. Williams. Textural information in SAR images. *IEEE Trans. Geosci. Remote Sensing*, GE-24(2):235–245, Mar 1986.

[104] M. Unser. Sum and difference histograms for texture classification. *IEEE Trans. Pattern Anal. Machine Intell.*, 8(1):118–125, Jan 1986.

[105] M. Unser. Texture classification and segmentation using wavelet frames. *IEEE Trans. Image Processing*, 4(11):1549–1560, Nov 1995.

[106] J.F. Vesecky, R. Samadani, M.P. Smith, J.M. Daida, and R.N. Bracewell. Observation of sea-ice dynamics using synthetic aperture radar images: Automated analysis. *IEEE Trans. Geosci. Remote Sensing*, 26(1):38–48, Jan 1988.

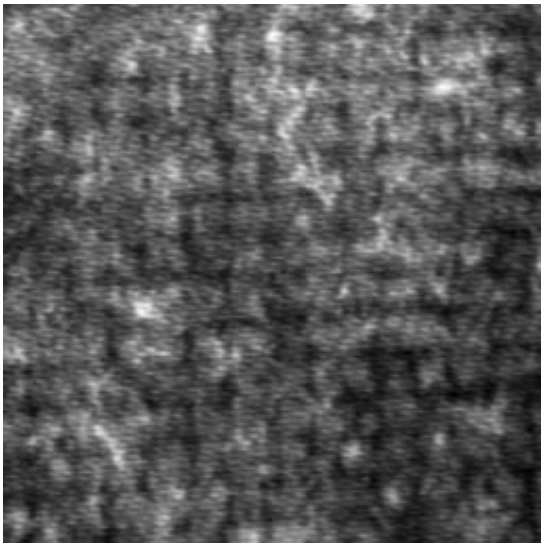
[107] J.S. Weszka, C.R. Dyer, and A. Rosenfeld. A comparative study of texture measures for terrain classification. *IEEE Trans. Syst., Man, Cybern.*, SMC-6(4):269–285, Apr 1976.

[108] G.A. Wright. Feature selection for texture coding. Master's thesis, Department of Systems Design Engineering, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1, 1986.

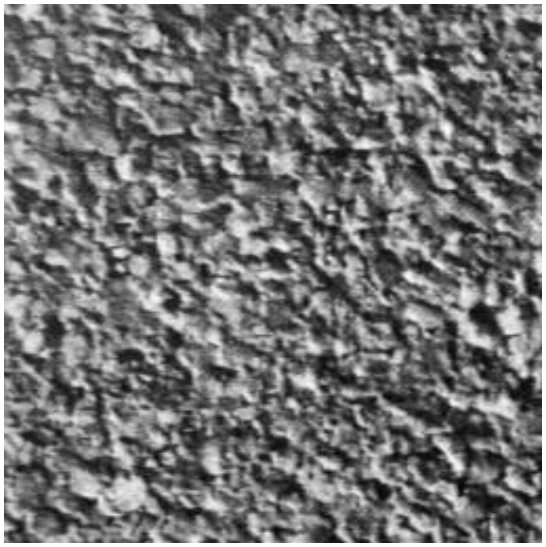
[109] R.K. Young. *Wavelet Theory and Its Applications.* Kluwer Academic Publishers, Boston, 1st edition, 1993.

Appendix A

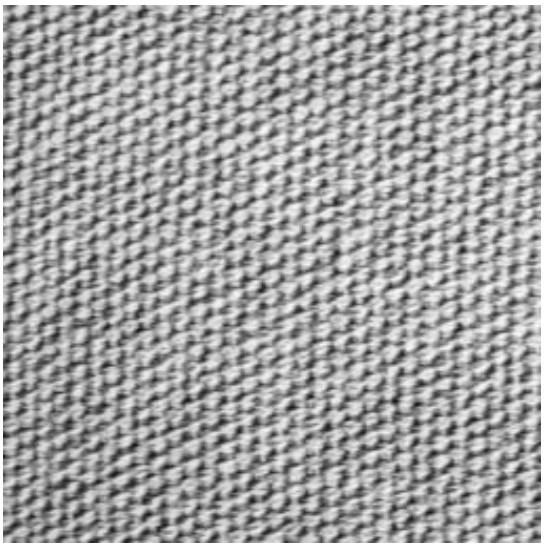
Brodatz Imagery



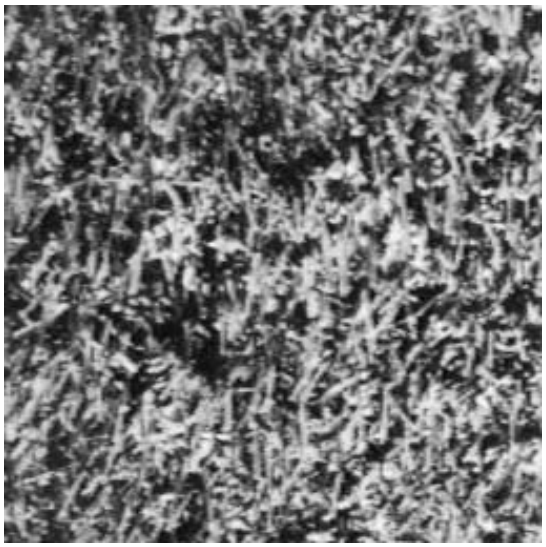
(a) Cloth (D19)



(b) Cork(D4)

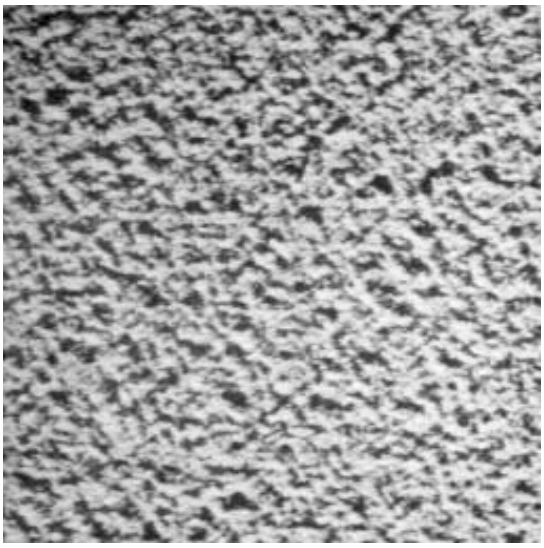


(c) Cotton (D77)

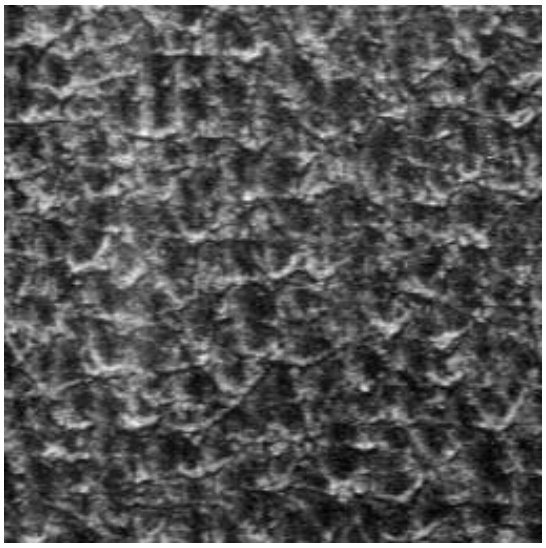


(d) Grass (D9)

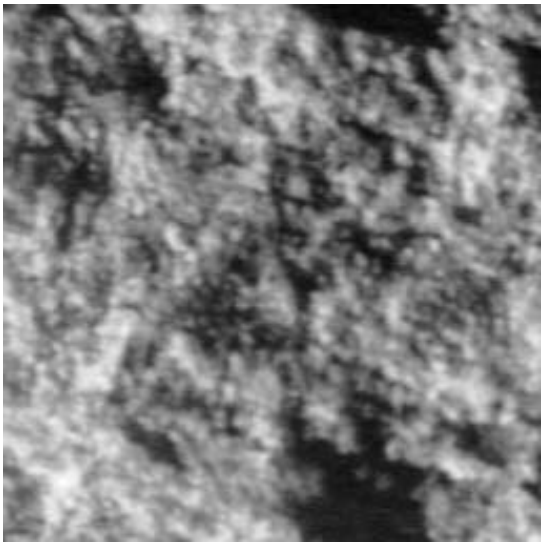
Figure A.1: Brodatz images used for classification testing.



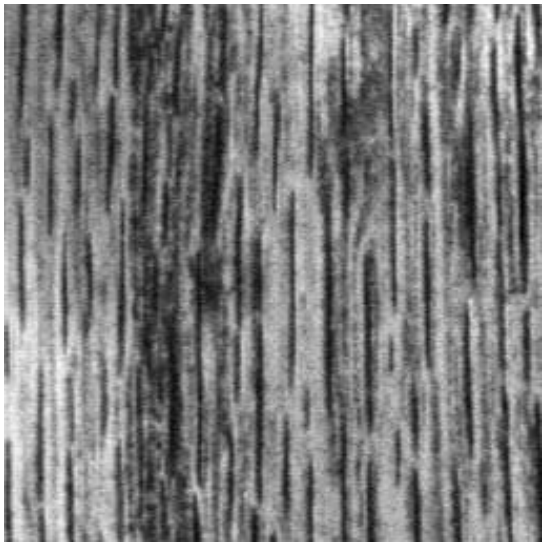
(e) Paper (D57)



(f) Pigskin(D92)



(g) Stone (D2)



(h) Wood (D68)

Figure A.1 Brodatz images used for classification testing (cont.).