Segmentation of RADARSAT-2 Dual-Polarization Sea Ice Imagery

by

Peter Yu

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Systems Design Engineering

Waterloo, Ontario, Canada, 2009

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

The mapping of sea ice is an important task for understanding global climate and for safe shipping. Currently, sea ice maps are created by human analysts with the help of remote sensing imagery, including synthetic aperture radar (SAR) imagery. While the maps are generally correct, they can be somewhat subjective and do not have pixel-level resolution due to the time consuming nature of manual segmentation. Therefore, automated sea ice mapping algorithms such as the multivariate iterative region growing with semantics (MIRGS) sea ice image segmentation algorithm are needed.

MIRGS was designed to work with one-channel single-polarization SAR imagery from the RADARSAT-1 satellite. The launch of RADARSAT-2 has made available two-channel dual-polarization SAR imagery for the purposes of sea ice mapping. Dual-polarization imagery provides more information for distinguishing ice types, and one of the channels is less sensitive to changes in the backscatter caused by the SAR incidence angle parameter. In the past, this change in backscatter due to the incidence angle was a key limitation that prevented automatic segmentation of full SAR scenes.

This thesis investigates techniques to make use of the dual-polarization data in MIRGS. An evaluation of MIRGS with RADARSAT-2 data was performed and showed that some detail was lost and that the incidence angle caused errors in segmentation. Several data fusion schemes were investigated to determine if they can improve performance. Gradient generation methods designed to take advantage of dual-polarization data, feature space fusion using linear and non-linear transforms as well as image fusion methods based on wavelet combination rules were implemented and tested. Tuning of the MIRGS parameters was performed to find the best set of parameters for segmentation of dual-polarization data. Results show that the standard MIRGS algorithm with default parameters provides the highest accuracy, so no changes are necessary for dual-polarization data. A hierarchical segmentation scheme that segments the dual-polarization channels separately was implemented to overcome the incidence angle errors. The technique is effective but requires more user input than the standard MIRGS algorithm.

Acknowledgements

Contents

B Test Images **94**

List of Tables

List of Figures

List of Symbols

β spatial context weighting parameter in MIRGS that determines the influence of the spatial context model on the segmentation process.

c number of channels in the multivariate image y.

C_1 first user defined MIRGS parameter that determines the value of β.

C_2 second user defined MIRGS parameter that determines the value of β.

∇_s normalized image gradient at site s.

$\nabla_s^{(pq)}$ normalized image gradient at site s from channel pq, such as $\nabla_s^{(HH)}$ or $\nabla_s^{(HV)}$.

∇_s^{VFG} normalized image gradient at site s calculated with the vector field gradient approach.

D_F wavelet coefficients for the fused image.

$D(r, c, k, l)$ wavelet coefficient at row r, column c in the coefficient image l at decomposition level k.

J minimum Fisher criterion computed between all pairs of classes in image.

κ Kappa coefficient, a measure of accuracy.

μ_i mean feature vector for class i.

M number of columns in an image.

N	number of rows in an image.
n	number of classes into which the image is segmented.
n_{train}	number of training feature vectors given to an NLDR method for learning the manifold coordinates.
Ω_i	set of all sites belonging to class i in the segmentation result.
pq	SAR polarization channel, where p is the transmit polarization and q is the receive polarization.
$\mathbf{\Psi}_s$	non-linear manifold coordinates of the feature vector at site s.
s	site in an image represented by a pixel location.
σ°_{pq}	backscatter coefficient for channel pq, such as σ°_{HH} or σ°_{HV}.
σ°	backscatter coefficient.
$\mathbf{\Sigma}_i$	covariance matrix of class i.
x	segmentation result image where each pixel is assigned a class number.
x_s	class number assigned to site s in the segmentation result x.
y	multivariate image which is input to MIRGS, where each pixel contains a feature vector.
$\mathbf{y}_s$	multivariate feature vector at site s in image y.
$y_s^{(pq)}$	element of feature vector $\mathbf{y}_s$ for channel pq, such as $y_s^{(HH)}$ or $y_s^{(HV)}$.

List of Acronyms

ADWA   absolute difference weighted average.
Auto-SITH   automatic generation of SITH hierarchies.

CG   Canny gradient combination rule.
CIS   Canadian Ice Service.

DB   Davies-Bouldin index.
DN   digital number.
DWT   discrete wavelet transform.

ENVISAT ASAR   ENVISAT advanced SAR.

HH   horizontal transmit, horizontal receive.
HV   horizontal transmit, vertical receive.

LEIGS   Laplacian eigenmaps [5].
LLE   locally linear embedding [47].
LTSA   local tangent space alignment [67].

MAGIC   Map Guided Ice Classification [12].
MAX   maximum gradient combination rule.
MIRGS   Multivariate Iterative Region Growing with Semantics [43].
MRF   Markov random field.
MSD   multiscale decomposition.

NLDR   non-linear dimensionality reduction.

PAL	parabolic arc-length projection.
PCA	principal components analysis.
R1	RADARSAT-1.
R2	RADARSAT-2.
RAG	region adjacency graph.
RGB	red-green-blue.
SAR	synthetic aperture radar.
SITH	sea ice type hierarchy.
SWT	stationary wavelet transform.
VFG	vector field gradient [27].
WA	weighted average.
WMO	World Meteorological Organization.

Chapter 1

Introduction

Sea ice mapping is an important application of remote sensing systems. It is essential for understanding the Arctic climate system [25] and for safe navigation of ships in waters where sea ice can form [59]. The primary task of sea ice mapping is to create maps that indicate the geographic distribution of different types of sea ice, with type being denoted by stage of development and other properties. Image data from satellite-based synthetic aperture radar (SAR), such as data from RADARSAT-1 (R1), are an important source of information for sea ice mapping [19]. RADARSAT-2 (R2), launched in 2007, is a Canadian SAR satellite that offers several technical enhancements over R1, including higher spatial resolution and additional imaging modes that are expected to improve discrimination of water from ice [44] and to better distinguish between different types of ice [50]. These enhancements are important because under certain but common circumstances, interpreting the various types of ice and water in the image can be difficult.

Since R2 is a relatively new satellite, little work has been done on evaluating the actual usefulness of these expected enhancements. Additionally, current operational sea ice maps are produced by human analysts with visual inspection of the image data [19]. This process is somewhat subjective, as different ice analysts can produce different results given the same data set. It is also extremely difficult for humans to produce a highly detailed, pixel-level accurate ice map in an operational setting due to the workload involved. Ice analysts currently provide only broad, regional maps that are outlines of areas with a certain composition of ice types. This has been one of the motivating reasons for interest in automated sea ice mapping algorithms for organizations such the Canadian Ice Service (CIS), which produces operational sea ice maps in Canada.

Automated algorithms are also useful for scientific research studies by reducing the workload and improving the objectivity of sea ice image analyses. For example, in order to generate information about the amount of ice present in an area, Belchanskya and Douglas had to perform manual inspection and thresholding of SAR images to distinguish between water and ice [4]. Worbya and Comiso visually compared ice information from passive microwave with SAR imagery [62]. While they found a good match, pixel-level accurate ice maps would help to improve the objectivity of their comparison. As with the operational case, such an ice map is only feasible with an automated algorithm.

As explained above, automated sea ice mapping algorithms are desirable for a number of reasons. Therefore, this thesis investigates the benefits of the enhanced information available in R2 data for use in automated sea ice mapping. Specifically, the problem to be investigated is how to make the best use of R2 data obtained with the new *dual-polarization* mode in the iterative region growing with semantics (IRGS) algorithm [65]. This algorithm is part of a larger system called MAGIC (Map-Guided Ice Classification) [12], which aims to provide pixel level accurate ice maps given a manually created ice map.

IRGS provides an unsupervised segmentation of the image to be analyzed, dividing the image into disjoint regions. In image segmentation, each region groups together image pixels that are similar in gray level or some other feature. In the case of sea ice mapping, the different regions ideally correspond to certain ice types, under the assumption that the ice types can be distinguished by the features available. The term *feature* generally refers to properties of objects that are either direct measurements of the object or can be extracted by some operation on the measurements [18]. Image segmentation produces maps that are *unlabeled*, meaning that there is no assignment of ice type to any of the regions produced. In the MAGIC system, the labeling process is performed after the segmentation stage by assigning an ice type to each of the regions.

The dual-polarization mode of R2 provides additional features unavailable from R1 to distinguish the different ice types. It should therefore be able to improve the image segmentation results obtained with IRGS. Thus, the problem being considered in this thesis is to determine whether this is the case and find the methods to use these additional features which produce the best segmentation result. The results of this work can then be used to extend the MAGIC system to make use of R2 dual-polarization data.

Chapter 2 gives an overview of sea ice mapping and SAR imaging and present observations about the appearance of sea ice in R2 images. Image segmentation and the existing MIRGS algorithm are also explained to give context for the rest of the thesis. Chapter 2 also presents an initial evaluation of MIRGS for segmenting R2 data to establish a set of goals that the methods presented in the thesis will address. Chapter 3 presents the proposed methods. Chapter 4 evaluates the performance of the proposed methods through a number of experiments. Finally, Chapter 5 concludes the thesis by presenting major findings and recommendations for future work.

Chapter 2

Background

2.1 Sea ice mapping

Sea ice refers to any ice that is found floating at sea and can be categorized into many different *stages of development* (ice types), each of which has very different properties [10]. The stages of development refer to how thick the ice is, which roughly corresponds to the age of the ice. For example, new ice that has just formed is thin and very different from ice that has had time to freeze and consolidate into a substantial navigation hazard.

An example of a partial sea ice chart (an operational sea ice map) created by the CIS is shown in Figure 2.1a. The ice chart polygon regions are overlaid on top of the SAR image that corresponds to the date and location of the charts. Each polygon region is coded with an *egg code* that contains the ice analyst's estimate of the composition of the ice types present in the polygon region. The ice in each polygon is not homogeneous and the egg code can contain more than one ice type. Since the polygons are created manually, a more detailed break down of the ice types is not often feasible. For the purposes of this thesis, an ideal sea ice mapping process should assign a unique ice type to each pixel in the image.

Egg codes for four of the polygons are shown on the right side of Figure 2.1a. The egg code follows standards set by the World Meteorological Organization (WMO) and the interpretation of the egg code is explained in Figure 2.1b [10]. The egg code values of interest for this thesis are $S_{\{a,b,c,d\}}$, which represent the multiple ice types (stages of development) that are present within the polygon. The value C_t represents the concentration of sea ice in tenths of the polygon's area, with the
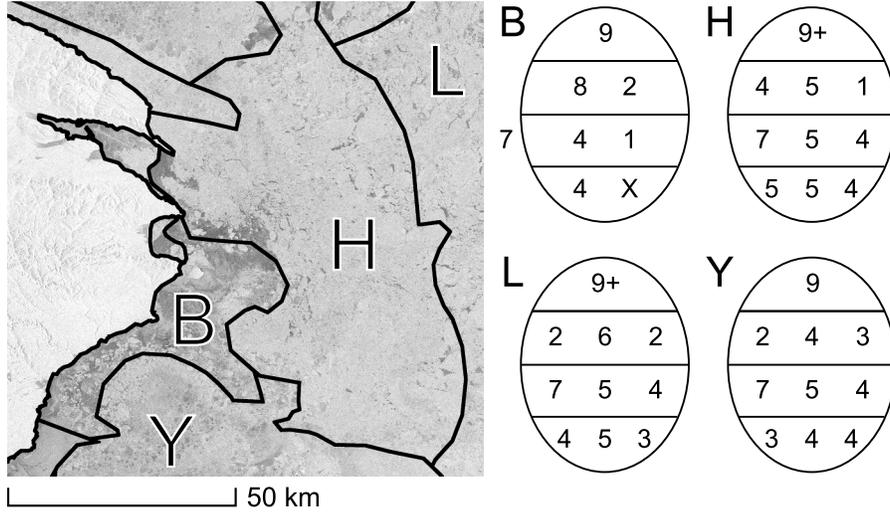
(a) Ice chart example.
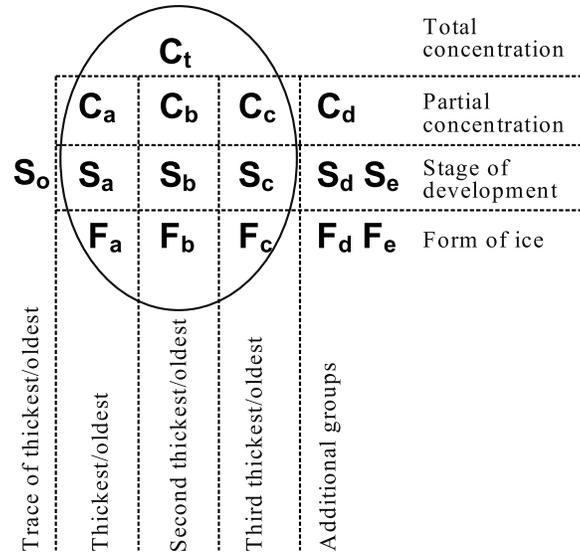
(b) Egg code meanings [10].

Figure 2.1: (a) An example of ice chart polygons overlaid on top of a SAR image, with the associated egg code information for each polygon. (b) Each egg code consists of fields that indicate the total concentration of ice C_t (in tenths, with 9+ indicating greater than nine-tenths concentration) and the partial concentrations ($C_{\{a,b,c,d\}}$) and forms ($F_{\{a,b,c,d\}}$) that correspond to each stage of development ($S_{\{a,b,c,d\}}$), respectively. Trace amounts of one additional stage of development can be indicated by the presence of S_o. S_e and F_e provide for an additional stage of development in Canada but this is rarely used.

5

Table 2.1: Listing of sea ice stages of development, along with thickness and code used to denote each type, from [10].

Name	Thickness (cm)	Code
New Ice	< 10	1
Nilas, Ice Rind	< 10	2
Young Ice	10 - 30	3
Gray Ice	10 - 15	4
Gray White Ice	15 - 30	5
First Year Ice (FYI)	≥ 30	6
Thin FYI	30 - 70	7
First Stage Thin FYI	30 - 50	8
Second Stage Thin FYI	50 - 70	9
Medium FYI	70 - 120	1·
Thick FYI	> 120	4·
Brash		-
Old Ice		7·
Second Year Ice		8·
Multi-Year Ice		9·
Ice of Land Origin		▲·
Undetermined or Unknown		X

remainder being water. $C_{\{a,b,c,d\}}$ represent the concentrations of each of the stages of development present. $F_{\{a,b,c,d\}}$ lists the form of ice corresponding to each of the stages of development. Form refers to the typical size within the polygon region of the ice floes for the corresponding type of ice. S_o, S_e and F_e allow trace amounts of additional stages of development to be specified but are rarely used.

Table 2.1 shows all sea ice stages of development defined by the WMO, their thickness and their corresponding codes [10]. This table lists the different types of ice that are the ice types of interest in ice mapping. First year and thinner ice types are formed during the most recent freezing season from sea water. Brash refers to ice formed from fragments of various types of ice after collisions of ice structures. Old ice refers to ice that has experienced at least one melting season after formation. The salinity content and physical structure of each type of ice is distinct and gives each ice type a different appearance in SAR imagery. Not all of the ice types can be distinguished based on gray tone in SAR images and ice analysts frequently have to use shape information and ancillary data such as weather conditions to make a determination. In addition to mapping the location of each of the aforementioned ice types, ice maps must also indicate where open water occurs.
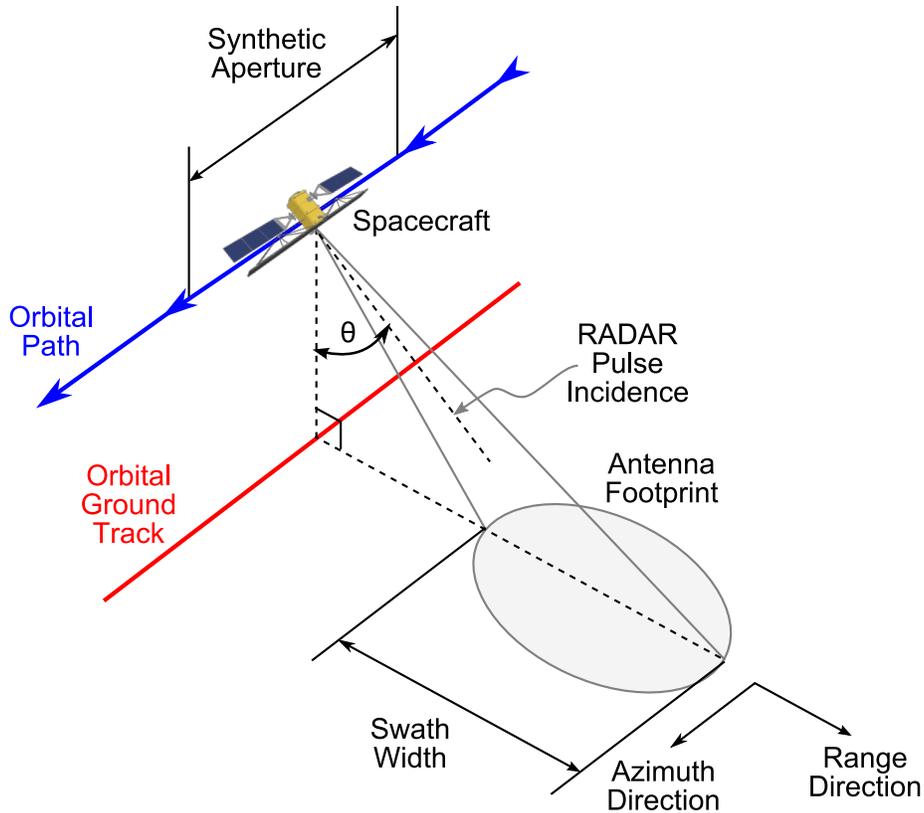
Figure 2.2: A diagram explaining the operation of SAR, adapted from [45]. The spacecraft sends a pulse of RADAR energy at the Earth's surface, which is reflected back and received as backscatter. The motion of the spacecraft creates a synthetic antenna aperture that is larger than the size of the antenna. The incidence angle θ defines the angle to any ground range, which is the distance from the orbital ground track.

2.2 Overview of SAR imaging

SAR imagery is one of the main sources of information for sea ice mapping. Therefore, an understanding of SAR and the new SAR imaging capabilities of R2 will be presented in this section.

The basic operating configuration of a spaceborne SAR is shown in Figure 2.2. The system consists of a spacecraft with the SAR equipment moving along an orbital track, which traces out the orbital ground track along the Earth's surface [45]. The azimuth direction is the direction of motion of the spacecraft, while the range direction corresponds to the distance from the orbital ground track. The incidence angle θ is the angle to points on the ground along the range direction. The SAR system emits microwave pulses at the Earth's surface. The antenna footprint is

7

(a) Horizontal (H) Polarization   (b) Vertical (V) Polarization
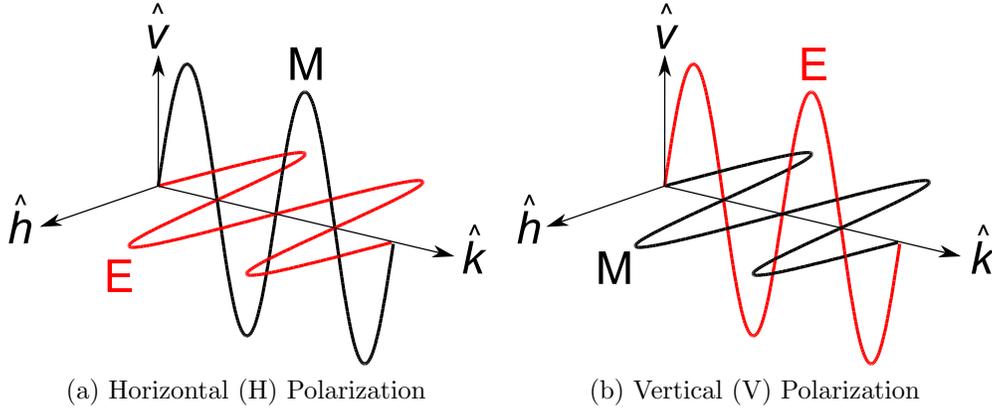
Figure 2.3: Waves of electromagnetic radiation consist of orthogonal electric (E) and magnetic fields (M) traveling in direction $\hat{\mathbf{k}}$. Linearly polarized waves, such as that used in SAR, have fixed directions for the fields. (a) Horizontally polarized waves have electric fields in the $\hat{\mathbf{h}}$ direction. (b) Vertically polarized waves have electric fields in the $\hat{\mathbf{v}}$ direction.

the area illuminated by the pulse. The swath width determines the width of the strip that the system can image. Once a pulse is emitted, the receiver on the SAR system waits for a return pulse of energy that has been *backscattered* toward it by the surface.

The antenna lengths required to create high resolution images are too large to launch into orbit for traditional RADAR. SAR solves this by using signal processing techniques and the motion of the spacecraft to create the effect of a larger antenna, giving rise to the *synthetic aperture* [37], [34]. The motion of the spacecraft also allows the field of view to be advanced in the direction of motion so that two dimensional SAR images of the surface can be generated.

The values of pixels in SAR images are the *backscatter coefficients* (σ°) on a decibel scale (dB) at each ground location corresponding to the pixels. The backscatter coefficient σ° is a dimensionless value that indicates the amount of backscatter from the surface.

SAR systems can transmit and receive EM energy in different polarizations. Polarization refers to the orientation of the electric field component of the electromagnetic wave. Figure 2.3 shows the two possible polarizations for the EM radiation that can be transmitted or received from SAR systems: horizontal (H) and vertical (V). The polarizations differ in that the electric field of one is orthogonal to the electric field of the other.

Since there are two possible polarizations for transmit and two possible polarizations for receive, four possible *polarization channels* can be measured, corresponding to the four possible backscatter coefficients σ_{pq}°, where p (transmit polarization) and q (receive polarization) can each be either H or V. When $p = q$, the σ_{pp}° channel is called the *co-polarization* (co-pol) channel. When $p \neq q$, the σ_{pq}° channel is called the *cross-polarization* (cross-pol) channel.

SAR systems are not necessarily designed to measure all four polarization channels. A *single-polarization* SAR system can only measure one of the four possible σ_{pq}°. A *dual-polarization* SAR can measure σ_{HH}° and σ_{HV}° *or* (σ_{VH}° and σ_{VV}°). Only a *quad-polarization* (or *fully-polarimetric*) SAR can measure all four possible σ_{pq}°. R1 is a single-polarization SAR while R2 has single-, dual- and quad-polarization imaging modes.

Different ice types have different backscatter characteristics that can be measured by SAR, since each ice type differs in factors such as surface roughness, volumetric structure and salinity [38]. Each of these factors affects the backscatter level in each of the polarization channels, making it possible to distinguish between the different types of ice with these backscatter "signatures". Very smooth salinated ice reflects EM radiation away from the sensor and appears very dark in all channels. Ice that only scatters EM radiation incident on its surface tends to scatter in the same polarization as the incoming radiation so that the co-polarization channel is brighter than the cross-polarization. Ice that scatters radiation after it has penetrated the surface tends to have closer values in the co- and cross-polarization channels.

Because of the large number of factors involved in determining the backscatter characteristics of ice, there is considerable variability even within one ice type. Smooth first year ice will have a lower backscatter than rough first year ice. Although both types belong to the same stage of development, they look quite different in SAR images. Ice at different stages of development can also look the same due to intra-type overlap of backscatter. Dual-polarization and quad-polarization SAR provide more information to help disambiguate different ice types than single-polarization SAR. R2 improves upon R1 by adding an operationally useful dual-polarization mode.

2.3 RADARSAT-2 capabilities and data

The most important improvement provided by R2 for operational sea ice mapping is the dual-polarization ScanSAR Wide mode [44]. This mode combines the same 500 km swath width as R1's single polarization ScanSAR mode (needed to provide adequate spatial coverage), with the additional information provided by dual-polarization imaging. The European Space Agency's ENVISAT Advanced SAR (ASAR) also provides dual-polarization data but only has a swath width of 100km [48]. ScanSAR Wide has a pixel resolution of 100 m × 100 m, with a pixel spacing of 50 m × 50 m (the pixels actually represent the backscatter of overlapping RADAR footprints). A full SAR scene is therefore approximately 10000 × 10000 pixels. This thesis will focus on investigating the use of R2 dual-polarization data from its ScanSAR Wide mode since it will be the mode used for sea ice monitoring. The Canadian Ice Service (CIS) expects to use the co-polarization (σ_{HH}°) and the cross-polarization (σ_{HV}°) channels for their operations and has provided several real R2 scenes for testing in this thesis. These scenes were recorded over the Gulf of St. Lawrence area on February 25 and March 4, 2009. CIS has also provided operational ice charts for Gulf of St. Lawrence on these dates, although they were created from R1 data using the HH channel since CIS had not yet integrated R2 imagery into their operational pipeline.

The HH channel provides the same data as R1, while the addition of the HV channel is expected to improve the discrimination of ice and water, particularly water that has been wind roughened, which looks very similar to some types of ice in the HH channel at small incidence angles. Experiments carried out by Manore et al. [31] and Scheuchl et al. [50] on airborne and ENVISAT ASAR dual-polarization data confirmed that the HV channel improves the discrimination of ice and water under these circumstances. R2's dual-polarization mode should be similar.

2.4 Appearance of sea ice in R2 imagery

The previous sections have discussed the theoretical aspects of sea ice imaging. This section describes the appearance of sea ice in R2 imagery by investigating the available R2 data. Table 2.2 shows the appearance of several different ice types in the HH and the HV channels, along with a brief description of the ice characteristics. The first column of the table lists the name of the ice type, with the stage of

10

development code in parenthesis. The ice types were identified by consulting CIS ice charts and finding locations where the ice type could be identified manually (e.g. ice chart polygons that are predominantly one type of ice or parts of the ice chart polygon where there is a clear distinction between the different ice types present).

The HV channel always appears darker than the HH channel because the backscatter is always lower in HV. This is because the SAR transmission is H-polarization, so that there is a tendency to backscatter H-polarization (seen in HH) rather than V-polarization (seen in HV). First-year ice is bright in both HH and HV while the two gray ice types are darker. New ice and open water are very similar, appearing bright in HH and dark in HV. However, open water tends to be smoother with no visible structures. Not shown in the table are smooth versions of gray ice, new ice and water, which are all dark in both HH and HV bands. In the first year ice images, some form of ice that is bright in HH and dark in HV is visible between the first year ice floes (which are very well defined in the HV image). This is wind roughened water or new ice, which are difficult to distinguish from the first year ice in the HH band alone at near range (small) incidence angles. The HV band, however, provides information on the location of the first year ice since water and new ice are dark.

Table 2.2 shows images that each span a very small incidence angle range of no more than 1.5°, while ScanSAR Wide images span 30° from 20° to 50°. The backscatter values of wind-roughened open water and new ice, and to a lesser extent the other ice types, are dependent on incidence angle [17], [57]. The incidence angle dependence is significant and can cause the appearance of open water and new ice to vary dramatically from near range (small incidence angles) to far range (large incidence angles). This is shown in Figure 2.4(a) for open water in the HH channel. The data scaling in the figure has not been adjusted to exaggerate the effect; the dynamic range of all ice types is between -35 dB to -5 dB, as shown in the figure. Thus, open water can take on the backscatter of a variety of other ice types depending on the incidence angle and can cause difficulties in separating it and other ice types. The HV channel is much less sensitive to the incidence angle, as shown in Figure 2.4(b) but there is some systematic banding in the HV band. Experiments with ENVISAT ASAR data showed a similar trend and may be related to the signal-to-noise ratio varying across the image due to insufficient transmitter power [49]. It is generally not possible to apply a correction to the scene to eliminate the incidence angle effect because the correction factor is different for each ice class and the ice class is not known beforehand [26]. Therefore, some way to harness the

11

Table 2.2: Appearance of different ice types in RADARSAT-2 images

Ice Type (Code)	Appearance	HH	HV
First Year Ice (7)	The ice appears bright in both HH and HV. Ice floes are well defined, have round corners and appear to be fairly homogeneous.		
Gray White Ice (5)	Moderately bright in HH but dark in HV. Dark in both HH and HV when smooth. Floes take on a cracked appearance.		
Gray Ice (4)	The ice appears moderately bright in both HH and HV and is similar to first year ice except it is not as bright.		
New Ice (1)	Similar to open water (bright to dark from near range to far range incidence angle) in HH but with visible structure. Dark areas in HH represent very smooth new ice. No backscatter in HV and appears very dark.		
Open Water	HH brightness varies with incidence angle and weather (rough water and near range are bright, calm water and far range are dark). No backscatter in HV. No visible structures.		
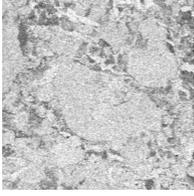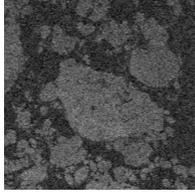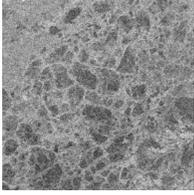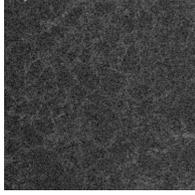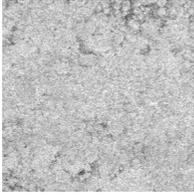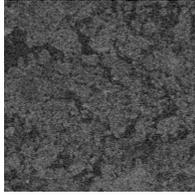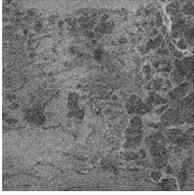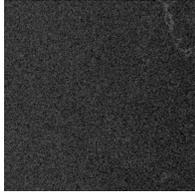
−5 dB

−35 dB

49   45   40   35   30   26

Incidence Angle (Degrees)

(a) HH

−5 dB

−35 dB

49   45   40   35   30   26

Incidence Angle (Degrees)

(b) HV
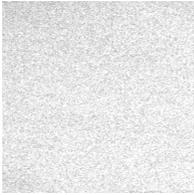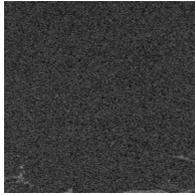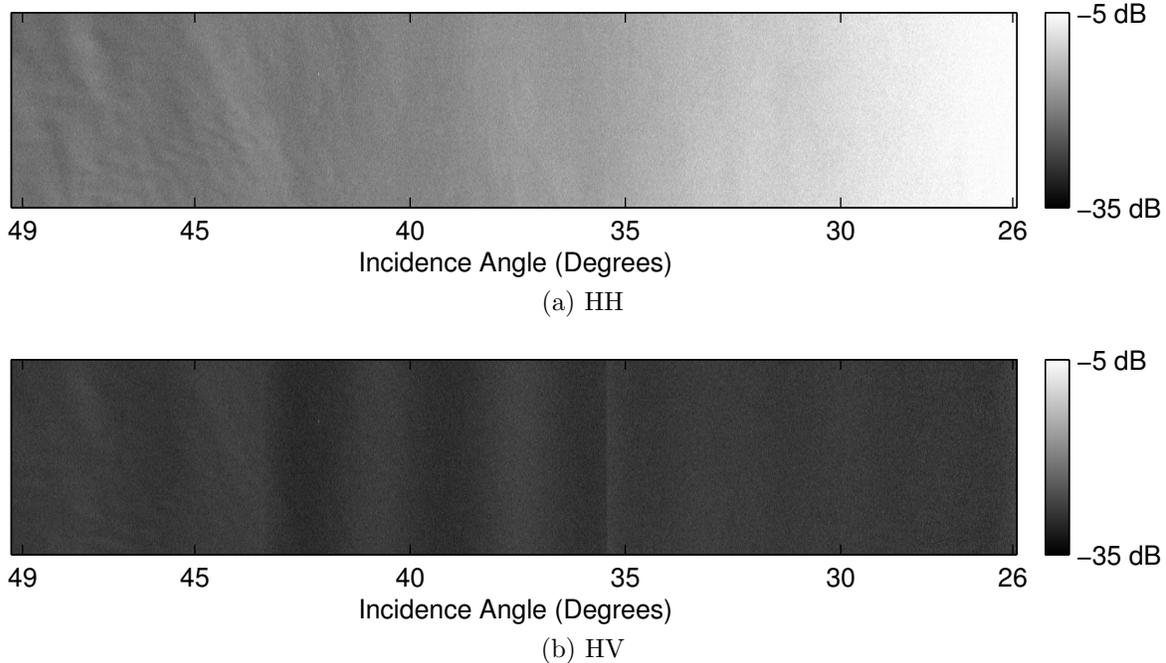
Figure 2.4: Appearance of open water as a function of incidence angle. (a) Open water exhibits an incidence angle dependent variation in backscatter in the HH band, with the near range (small incidence angle) being much brighter than the far range (larger incidence angle). New ice has similar characteristics. (b) The same scene in the HV band does not show as much incidence angle dependence, although there is systematic banding, in which the backscatter is not constant across incidence angles.

information in the incidence angle insensitive HV band is needed.

The proposed methods will take into account the above observations in order to improve sea ice segmentation in SAR scenes.

2.5 Image segmentation

As this thesis focuses on improving the image segmentation portion of the MAGIC system, the image segmentation problem must first be defined. The description used here is summarized from [65]. Let there be n classes into which the image is segmented. Let S be a set of sites on a lattice and $s \in S$ be a site in the lattice. Let $\mathbf{X} = \{X_s | s \in S\}$ be a set of discrete random variables forming a random field on S, with each X_s taking a value from $\{1, \ldots, n\}$. X_s indicates the class that is assigned to site s. The class numbers can be subsequently converted into meaningful ice class names through a classification and labeling process but this is not part of the image

segmentation process.

Let $\mathbf{Y} = \{Y_s | s \in S\}$ be the random field on S that is realized by the observed image. Each Y_s takes on a feature vector, with each element representing the tonal value from each of the available image channels. The tonal values could be the backscatter in dB from the HH and HV channels, for example, but in this thesis they could also take on other meanings depending on what transforms are applied to the image data.

Let $x = \{x_s | s \in S\}$ and $y = \{\mathbf{y}_s | s \in S\}$ be realizations of $\mathbf{X}$ and $\mathbf{Y}$, respectively. Based on the information contained in the image y, the image segmentation algorithm must generate the segmentation image x. There will be n classes in the segmentation image, denoted by disjoint regions $\Omega_1, \ldots, \Omega_n$, each of which contains all the pixels assigned to one class. Note that $y_s^{(pq)}$ represents the value of channel pq at site s. For example, a dual-polarization feature vector consists of $y_s = (y_s^{(HH)}, y_s^{(HV)})$.

The definition of the "best" segmentation for the purposes of sea ice mapping is to ensure that each Ω_i corresponds to a unique ice type. Different segmentation methods have different methods for estimating the best segmentation.

2.6 Multivariate IRGS algorithm

The IRGS algorithm [65] and its multivariate extension, MIRGS [43], is the image segmentation algorithm that is to be adapted to dual-polarization R2 data since it is fully incorporated into the MAGIC system and its results when applied to R1 data have been evaluated by CIS experts and found to be reasonable [12], [64]. This section summarizes [65] and [43] in its description of the algorithm, which will henceforth be referred to as MIRGS.

Figure 2.5 shows the major steps of the MIRGS algorithm. The following sections detail each of the steps of the MIRGS algorithm, based on the latest implementation [43].

2.6.1 Step 1: Image gradient and watershed generation

The algorithm starts by accepting as input an image with at least one image channel. The image is first segmented with a watershed algorithm [58] that divides the image into many small regions with relatively uniform backscatter in each. Each region v
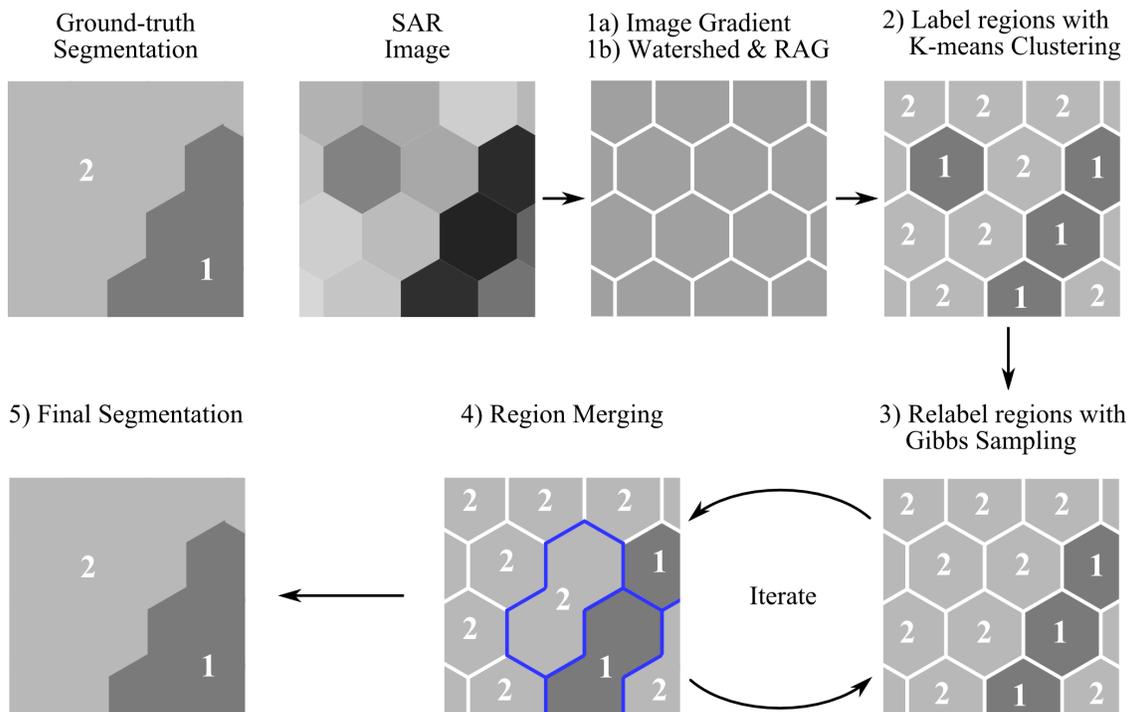
Figure 2.5: Major steps of the MIRGS algorithm. To initialize the system, MIRGS computes the 1a) image gradient, 1b) generates the watershed regions, the region adjacency graph (RAG) and 2) the initial segmentation with K-means. The 3) relabeling and 4) merging processes follows and are repeated for a user-specified number of iterations until the 5) final segmentation is produced.

consists of a set of sites S_v that belongs to it. By grouping sites into regions, the effect of speckle-noise is reduced since the feature vectors of the individual sites $\{\mathbf{y}_s | s \in S_v\}$ can be averaged into one feature vector $\mathbf{y_v}$ for the entire region. The image is represented by a data structure called a region adjacency graph (RAG) $\mathcal{G}$ whose nodes consist of the set of regions $\mathcal{V}$ and whose edges represent boundary sites between each pair of adjacent regions. The segmentation definitions are modified from that presented in Section 2.5. Each site no longer has a separate label in the segmentation image x. Rather, all sites within a region are assigned a single label. Let x^r be this region-based segmentation image and $\mathbf{X^r}$ be the set of all possible x^r.

In order to generate the watershed segmentation, the image gradient must be computed. Since there can be multiple channels in an image, MIRGS uses a vector field gradient (VFG) approach [27] to calculate a joint image gradient from all channels. The calculated gradient is always normalized by dividing by the largest gradient in the scene so that the largest gradient has a value of 1.0. This allows MIRGS to evaluate relative edge strength rather than absolute edge strength so that the algorithm scales properly for scenes with different dynamic ranges. The watershed is then calculated based on the joint, normalized image gradient.

2.6.2   Step 2: Region-based K-means initialization

As part of the initialization process and as a technique to push the algorithm to find a globally acceptable solution, each of the obtained watershed regions v is assigned a label x_v^r via a region-based K-means algorithm. For each region v, the K-means algorithm chooses the label that best satisfies:

$$x_v^r = \arg\min_i \sum_{s \in S_v} (\mathbf{y}_s - \mu_i)^{\mathrm{T}} (\mathbf{y}_s - \mu_i) \qquad (2.1)$$

where $\mathbf{y}_s$ is the feature vector at site s whose elements are the values of the image channels and where $^{\mathrm{T}}$ is the transpose operator. This process is iterative and begins with random means μ_i that are updated on each iteration.

2.6.3 Step 3: Labeling with Gibbs Sampling

Once the initial K-means segmentation is generated, MIRGS enters the iterative portion of the algorithm. The goal is to find the optimal labeling of each watershed region. This is done by finding a configuration of labels that globally minimizes a cost function. MIRGS iteratively performs labeling of the watershed regions, followed by region-merging. At each iteration, an intermediate segmentation result is generated. Region-merging reduces the number of nodes in the RAG by combining adjacent regions. This makes the labeling process in the subsequent iterations more efficient as fewer nodes have to be considered. Additionally, the solution process is not as likely to be trapped in a local minimum when regions are merged together [65]. The number of iterations for MIRGS is set by the user and 100 iterations is the number typically used.

The cost function that MIRGS minimizes in order to produce the optimal segmentation x^{r*} is the following:

$$x^{r*} = \arg \min_{x^r \in \mathbf{X^r}} E_f + E_s$$

(2.2)

where E_f is the feature model energy and E_s is the spatial context model energy. E_f is defined as follows:

$$E_f = \frac{1}{2c} \sum_{i=1}^{n} \sum_{S_v \subseteq \Omega_i} \sum_{s \in S_v} \log(|\mathbf{\Sigma}_i|) + (\mathbf{y}_s - \mu_i)^{\mathrm{T}} \mathbf{\Sigma}_i^{-1} (\mathbf{y}_s - \mu_i)$$

(2.3)

where c is the number of channels in the image, n is the number of classes, S_v are the sites in each watershed region v that are part of the region Ω_i that is assigned to class i, $\mathbf{\Sigma}_i$ is the class covariance matrix of class i and μ_i is the mean value of class i. E_f is the energy associated with the assumption that the image values of the watershed regions in each class follow a multivariate Gaussian distribution, which gives reasonable results [65], [64].

The spatial context model energy E_s is defined as follows:

17

$$E_s = \beta \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \sum_{s \in \partial\Omega_i \cap \partial\Omega_j} g(\nabla_s) \tag{2.4}$$

where β is a positive weighting value that determines the strength of the spatial context model in the cost function, and $\partial\Omega_i$ and $\partial\Omega_j$ are the sets of sites along the border of class i and j, respectively. $g(\nabla_s)$ is the edge penalty and is defined as:

$$g(\nabla_s) = \exp\left[-\left(\frac{\nabla_s}{K}\right)^2 \right] \tag{2.5}$$

where ∇_s is the normalized image gradient at site s (which is used to represent edge strength and is the same image gradient as that calculated for Step 1) and K is a parameter that determines the strength of the image gradient's effect. The setting of K is described thoroughly in [65] and [43] and is not repeated here.

E_s adds a penalty value to the total energy whenever the segmentation result has assigned two different classes to adjacent regions. This is similar to the Markov random field (MRF) based multi-level logistic (MLL) segmentation model, which only differs in that the edge penalty $g(\nabla_s)$ is always unity. MLL operates under the intuitive assumption that regions adjacent to each other are more likely to be from the same class than they are from different classes. The addition of the edge penalty $g(\nabla_s)$ in MIRGS refines this assumption. The penalty added will be higher when the edge strength between two regions is weak, and small when the edge strength is high. The MIRGS model agrees even more with intuition: humans would expect that if there is a strong edge between two regions, they are more likely to be from different classes than when there is no edge.

The entire cost function produces a segmentation that is a balance between labeling each region based on feature space similarity and spatial context. The cost function considers a segmentation more likely to be "true" when the regions in each class are similar to each other in feature space and when adjacent regions belong to the same class if the edge between them is weak. The parameter β controls the influence of the spatial context model. Large β will make the spatial context model stronger and will create very smooth segmentation results, ideal for simple images. For complex images, β should be small so the segmentation result is more detailed.

MIRGS incorporates a method to derive β from the data and is implemented as follows:

$$\beta = C_1 \frac{J/C_2}{1 + J/C_2} \beta_0 \qquad (2.6)$$

where J is the minimum Fisher criterion [18] between any two classes in the image according to the current segmentation result (i.e., it is the Fisher criterion between the two classes that are least separable from each other), C_1 and C_2 are user-defined constants and β_0 is an intermediate parameter. β_0 is calculated by considering the boundary length of the previous intermediate segmentation. It uses a Monte Carlo method to obtain a maximum likelihood estimation that will keep the boundary length the same on subsequent segmentations [65], which will preserve the level of complexity in the segmentation result.

The Fisher criterion J is a measure of class separability in feature space. Large J implies greater separability. The use of J to adjust β ensures that the effect of both feature space and spatial context models are balanced in the cost function. When J is large (as it is on the initial iterations), the previous segmentation result was strongly influenced by the feature space model, so β is made larger to compensate. As the effect of the spatial context model gets larger on subsequent iterations, J decreases. This causes β to decrease, preventing excessive influence by the spatial context model. The user is able to set C_1 and C_2 to control the relative strength of β. Large C_1 and small C_2 emphasize the spatial context model.

The actual optimization for Equation 2.2 is accomplished with Gibbs sampling [21] which chooses a label for each node in the RAG.

2.6.4   Step 4: Region Merging

After the labeling process is completed for all nodes, region merging is performed in a greedy fashion. The process only considers all pairs of adjacent regions which have the same class. Let ∂E be the total change of the energy in Equation 2.2 if a pair of such regions is merged. The algorithm will merge the pair of regions that has the smallest negative ∂E and update the RAG. This continues until the smallest ∂E is non-negative.

When region-merging is complete, the algorithm will go back to Step 3 until the desired number of iterations is reached, at which point the final segmentation will be produced.

2.7   Evaluation of MIRGS with R-2 data

In this section, MIRGS is evaluated using real R2 data to identify areas for improvement. A subimage was extracted from the February 25, 2008 Gulf of St. Lawrence scene and calibrated according to the steps in Appendix A. Based on CIS provided ice charts, a pixel-level manual segmentation was created of the scene, consisting of four classes: water, smooth ice of indeterminate type, first year ice and gray ice / gray white ice (types 4 and 5 jointly). In this case, it was extremely difficult to separate the scene into the exact stages of development listed in Table 2.1, so the ice was grouped by appearance in a red-green-blue (RGB) composite of the HH and HV channels. The manual segmentation is used as ground-truth to evaluate the performance of MIRGS on the test scene. Additional details about this image can be found in Section B.1.

Figures 2.6 (a)-(c) shows the HH and HV channels and the manually segmented ground-truth. Figures 2.6 (d)-(f) shows four class segmentation results using only the HH channel, only the HV channel and both HH and HV channels as features. The overall accuracy of each segmentation is calculated and listed in the figure captions. The HH channel, which contains the same data as R1, only achieves 61.7% accuracy. The HV channel only achieves 46.3% accuracy. Only the HH & HV feature set produces an accuracy of higher than 80%, which according to CIS requirements [20] is the target value that any ice mapping processes must achieve consistently.

Several observations can be made from these results. The HH channel contains significant feature space ambiguity between water and first year ice, causing the first year ice to frequently be grouped with open water. The HV channel does not contain sufficient information to segment four classes. Visually, the image only shows two classes. Therefore, the segmentation result is quite poor. The combined feature set appears to overcome all of these problems, producing a segmentation that appears closest to the ground-truth. Even so, the level of detail in the HH & HV result is somewhat lacking. Fine structures, particularly those of water, appear to be merged with neighbouring classes.

(a) HH   (b) HV   (c) Ground-truth

(d) HH result, Accuracy = 61.7%   (e) HV result, Accuracy = 46.3%   (f) HH & HV result, Accuracy = 84.7%

■ Smooth ice   ■ Gray Ice   ■ First Year Ice   ■ Open Water
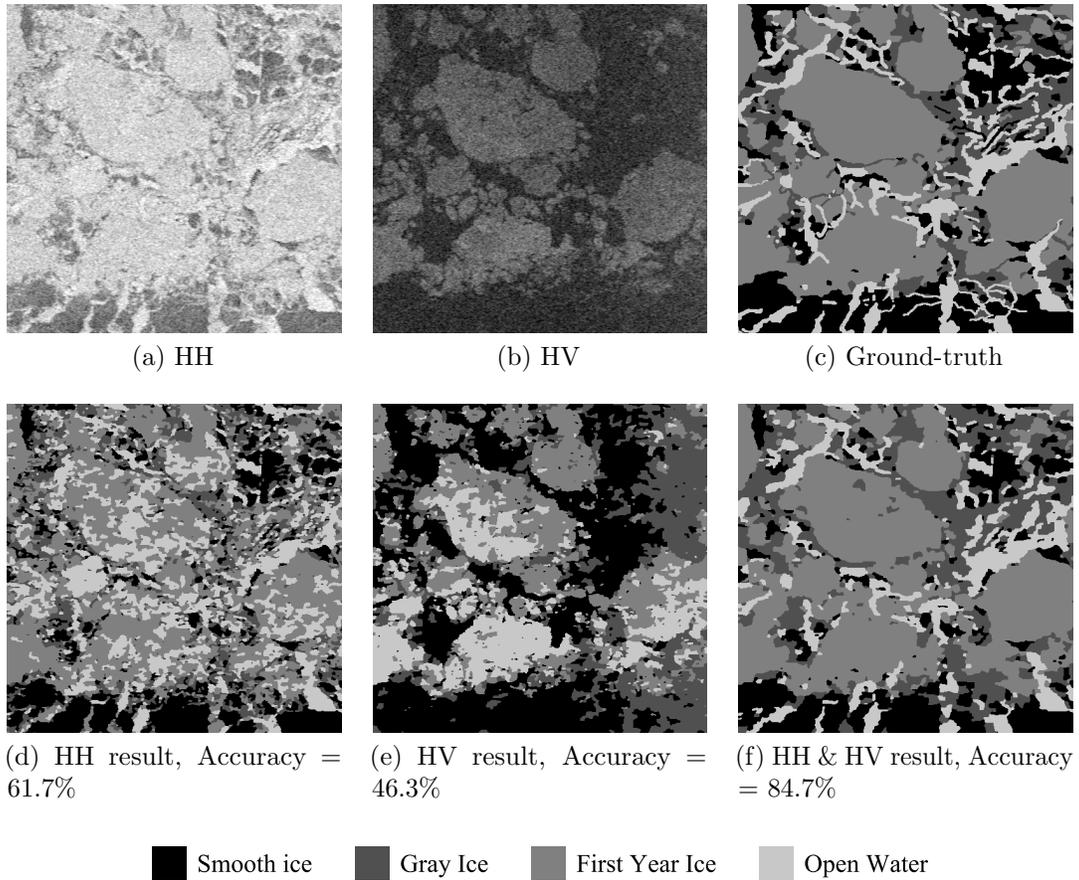
Figure 2.6: An evaluation of MIRGS with R2 data. A subimage from a R2 scene was selected for evaluating MIRGS using default parameters ($C1 = 3, C2 = 0.5$) by comparison with a manually segmented ground-truth. The HH result shows ambiguities in the segmentation and the HV result is incorrect. Only the HH & HV result is accurate but it appears to lose details due to overmerging.
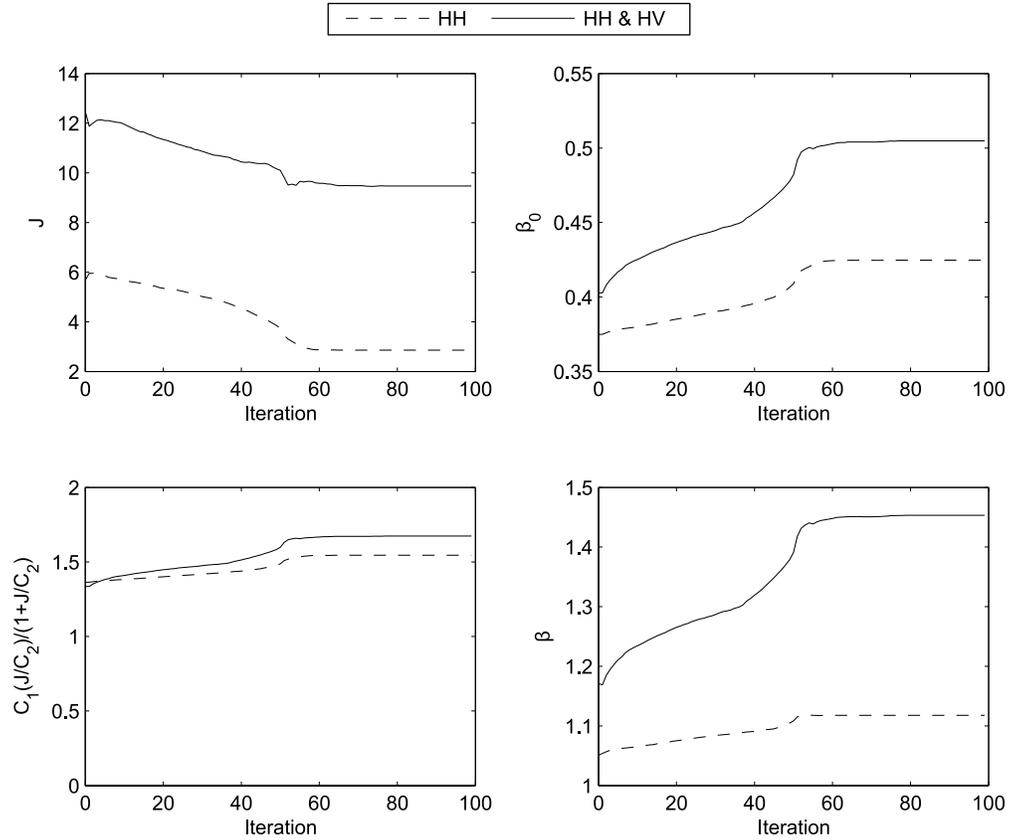
Figure 2.7: Comparison of the spatial context parameter β as a function of iteration between univariate and multivariate feature sets. The parameter β is affected by β_0 and the Fisher criterion J for the two least separable classes. The dual-pol data inherently causes more merging by causing larger β values for the same set of C_1 and C_2. This is because the feature space separability is higher which causes larger J and β_0 values. These in turn cause β to be larger, in accordance with Equation 2.6.

Qualitatively, the HH & HV segmentation result appears much smoother than the HH result, suggesting that the spatial context model is more dominant in the multivariate case. Figure 2.7 confirms this quantitatively. Plots of the minimum Fisher criterion J, β_0, the adjustment factor in Equation 2.6 and the β used at each iteration during the segmentation process show that the β is always larger for the multivariate HH & HV case compared to the HH-only case. This arises because of the better feature space separability between classes when using HH & HV. This means the initial segmentation of HH & HV is already smoother, which tends to make β_0 larger. Additionally, with J being larger, the adjustment factor encourages β to be larger as well. The end result is that the spatial context model is weighted more in the multivariate case, smoothing out the segmentation result.

The current combined gradient method may also contribute to loss of detail. Figure 2.8 illustrates this for a synthetic dual-polarization example. In Figure 2.8(a), a horizontal slice through the step edges in the image is shown. In the HH gradient (Figure 2.8(b)), the edges with the strongest relative strength with respect to other edges in the HH channel are the left-most and right-most edges. In the HV gradient (Figure 2.8(c)), the strongest edges are the left-most and middle edges. In the combined gradient (Figure 2.8(d)), the strongest edge is the left-most edge because it has a high edge strength in both of the channels. On the other hand, the middle edge is considered a weaker edge in the combined gradient even though it was one of the strongest edges in the HV channel. The combined gradient method [27] was designed to assign the highest strength only to edges that appear in both channels but for sea ice segmentation, strong edges in either of the dual-polarization channels should be considered meaningful. An example is open water and smooth ice having no contrast in HV but high contrast in HH: the edge that appears in HH definitely separates the two classes.

Figures 2.9(a)-(b) show the HH and HV channels of a full R-2 scene across the full range of incidence angles. Only the left side of the scene has ice; the smooth gradient that transitions from light to dark in the HH image is open water. This scene is a typical example of the incidence angle variation of open water. Additional details about this image can be found in Section B.3.

Figures 2.9(c)-(e) show three class segmentation results using HH, HV and HH & HV, respectively. All three segmentation results break up the open water portion into a number of regions, reflecting the difficulty that the algorithm has in dealing with the incidence angle effect. This scene has no associated manual segmentation

(a) Horizontal image slice (b) HH gradient

(c) HV gradient (d) Combined HH & HV gradient
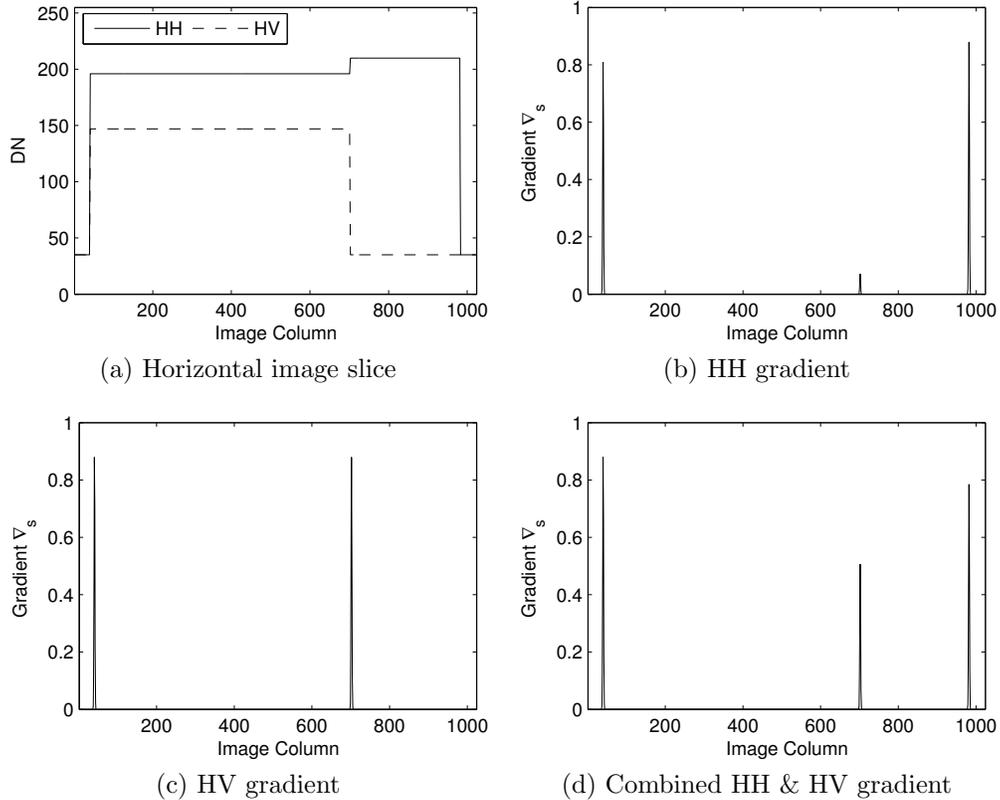
Figure 2.8: An example showing the dual-polarization performance of the joint VFG gradient method. (a) A plot of the intensity in digital numbers (DN) of a horizontal slice through a synthetic dual-polarization image with step edges shows that some edges are strong in HH and some are strong in HV. (b) Gradient through the same slice derived from the HH channel does not pick up the middle edge which is low contrast in HH. (c) Gradient from the HV channel does not pick up the right edge but assigns high relative strength to the left and middle edges. (d) The combined gradient picks up all edges, but the relative strength is strongest for the edge that is in both images even though all edges are expected to be equally meaningful.

(a) HH image   (b) HV image

(c) HH segmentation result   (d) HV segmentation result

(e) HH & HV segmentation esult   (f) Expected segmentation

■ Land   ■ Gray Ice   ■ First Year Ice   ■ Open Water
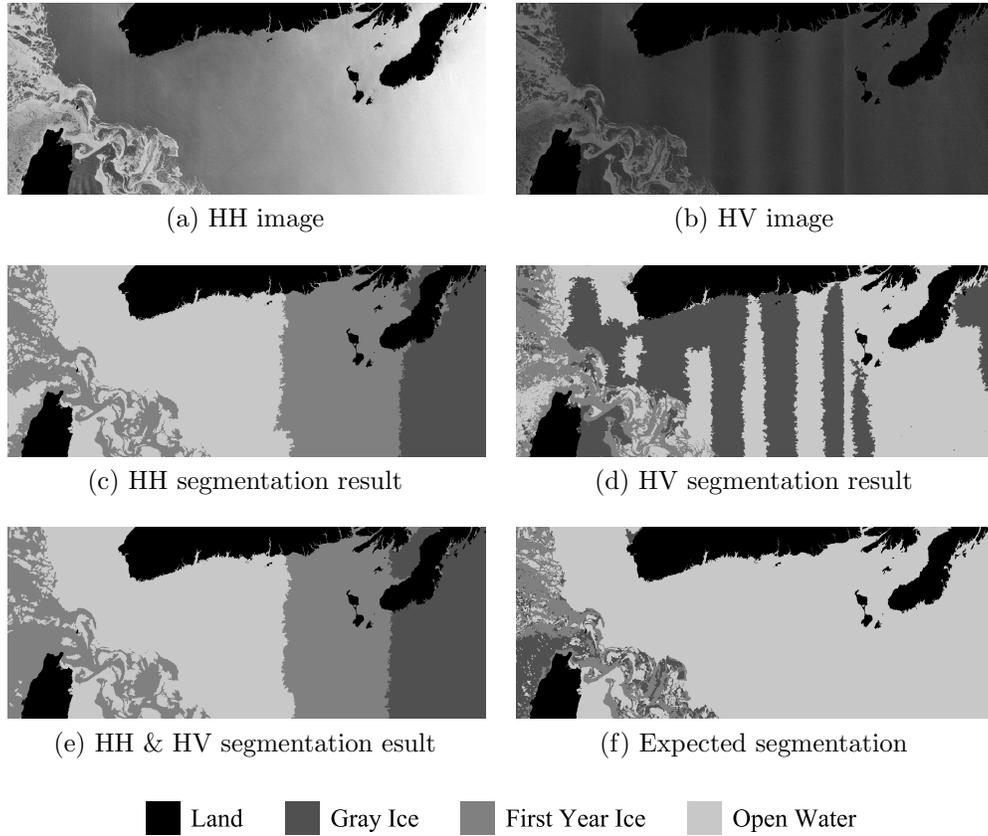
Figure 2.9: Segmentation results using MIRGS applied to a full scene R2 image. The incidence angle variation of open water causes both IRGS and MIRGS to produce incorrect segmentation results. The open water portion should not be split into multiple sections.

for accuracy comparisons as the size of the scene makes it impossible to attempt one but the segmentation results are clearly wrong. Although the appearance of open water in HV is relatively insensitive to incidence angle, the HV band alone cannot distinguish three classes.

Currently, MIRGS is unaffected by the incidence angle effect when the image to be segmented is small (as in Figure 2.6) or when a manually produced CIS ice chart is available to exclude open water regions from the segmentation process. This limits its flexibility for fully automated full scene segmentation.

In summary, although the use of dual-polarization data improves segmentation accuracy, possible areas of improvement that will be addressed by this thesis include increasing level of detail in the segmentation result and dealing with the effect of incidence angle on the segmentation. These objectives are detailed in Section 2.9.

25

2.8 Related work

Previous studies have examined the utility of multipolarimetric information for a variety of remote sensing applications. Much of the research literature has focused on using either fully-polarimetric SAR or systems where both the HH and VV channels are available. Images from fully-polarimetric airborne SAR systems have been assessed qualitatively for improving sea ice identification [31]. The authors identified that using either the cross-polarization channels independently or fusing the co-polarization channels by taking the ratio HH/VV can improve ice-water discrimination. The same ratio has also been used to estimate sea ice thickness [35]. In another work [51], fully-polarimetric backscatter data from the space-shuttle-based SIR-C SAR system was transformed into several different measures which can separate ice and water using simple thresholds. Fully polarimetric or HH & VV data has also been used for other applications, such as land cover classification [41], ship detection [29] and crop monitoring [8].

Although the results from these previous studies were promising, the techniques used are not directly applicable to operational R2 data, since they require channels not available in the dual-polarization mode. However, dual-polarization data consisting of HH and HV channels has also been studied in the literature. Dual-polarization ENVISAT ASAR data, which is similar to RADARSAT-2 data, was used with an unsupervised segmentation algorithm and tested for its ability to distinguish sea ice types [49]. The segmentation algorithm uses transformations on complex-valued dual-polarization data and a minimum-distance, Bayesian framework to classify each image pixel. As with previous studies, the separation of water and ice was improved by the dual-polarization data. RGB composite images of HH and HV data (R = HV, G = HV, B = HH) were found to significantly improve visual discrimination of open water and ice [1]. These findings are also supported in [44] and [50].

Another work tested dual-polarization data from a Ku-band SAR (a different frequency than the C-band R2) with a multivariate Gaussian maximum likelihood classifier and two neural-network classifiers [39]. All three classifiers performed similarly and the classification accuracies of first year ice, multi-year ice and icebergs were improved. Fusion of the co-polarized and cross-polarized channels was also performed using principal component analysis (PCA). This was found to improve the visual distinction of the different ice types but did not improve classification accuracy using the tested classifiers.

Most of the work mentioned has focused on making use of the multidimensional feature space provided by multipolarimetric data. Data fusion methods such as PCA in [39] transform the multidimensional data into a single image, relying on the transform to preserve separability of classes. PCA has been frequently tested for reducing the dimensionality of multisensor, multifrequency and/or multipolarization data sets, such as passive microwave and SAR data [14], passive microwave and scatterometer data[63] and Landsat visible and near-infrared imagery [42]. Dimensionality reduction with PCA is a linear projection of the multidimensional feature space onto a reduced dimension feature space that preserves maximum amount of variance [52]. Non-linear dimensionality reduction (NLDR) techniques can also be used to fuse multiple channels into a single image. Multidimensional data points in hyperspectral satellite imagery were projected onto reduced dimension manifold coordinates in [3], producing reduced dimensionality images that exceeded the quality of those produced by PCA. The improvement was achieved because the structure of the underlying data was non-linear.

Image fusion methods such as the HH/VV band ratio in [31] are also able to fuse multiple images together into a single image. Further examples include the HV/HH band ratio for fusing dual-polarization SAR [50], [14] and wavelet image fusion methods which decompose the images into wavelet coefficients at different scales and combine the coefficients from two or more images according to defined rules [40]. Wavelet image fusion has been used for multipolarization SAR image fusion [54], [23] and attempt to create fused images that combine details from all of the source images.

2.9 Objectives

The answers to the following research questions are the objectives of this thesis. These are formulated based on the observations in Section 2.7. Listed under each question are hypotheses that will be tested to answer the questions.

1. What methods can be used to increase the accuracy in the dual-polarization segmentation results?

   (a) An image gradient calculation method designed for dual-polarization data may be more effective at improving accuracy and capturing more detail. This method must be designed such that high contrast boundaries

that appear in only one of the polarization channels are treated as being as strong as an edge that appears in both images.

(b) Data and image fusion techniques can create a combined single channel image that preserves the feature space separability of the two original channels. The single channel image will contain high contrast boundaries between all classes separable with the dual-polarization data. A combined edge map that contains all the edges in both original channels can then be generated from this fused image.

(c) An appropriate choice of user-defined parameters for the MIRGS spatial context model parameter β can improve the accuracy and level of detail by compensating for the extra merging caused by the multivariate feature space.

2. Can the incidence angle effect that causes confusion between open water and ice be eliminated by using dual-polarization data?

Relative to the HH channel, the HV channel is insensitive to the incidence angle effect and can be used to separate open water from other ice types. Since the HV channel does not have sufficient information to distinguish between all possible ice types, segmentation will have to be performed hierarchically, where HV can be segmented into as many classes as it can distinguish (e.g., ice and open water). Each of the classes produced can be further subdivided by using the HH channel to segment all possible classes.

Chapter 3

Methods

This section describes the proposed methods. Sections 3.1 to 3.4 focus on improving accuracy for smaller scenes without the incidence angle effect. These methods seek to improve the current MAGIC application of segmenting manually created ice polygons down to the pixel level [12]. The manually created polygons exclude open water with large incidence angle related backscatter changes, so that the incidence angle effect can usually be safely ignored. Section 3.5 focuses on a hierarchical segmentation scheme that is designed to deal with the incidence angle effect for full R2 scenes. This can potentially lead to fully automatic full scene segmentation of sea ice.

Before being input into any of the methods in this section, the R2 data were first calibrated using the procedure described in Appendix A.

3.1 Gradient and edge map computation

As discussed in Section 2.7, the multivariate VFG gradient calculation method used by MIRGS, which computes the gradient using the two channels jointly, assigns a strong edge strength only to edges that appear in both the HH and the HV channels. However, strong edges that appear in only one channel but not in the other are equally meaningful because they denote an ice type boundary and should have a high edge strength to reflect this.

Given that both the initial watershed segmentation and the MIRGS algorithm rely on a proper accounting of meaningful edges, three different gradient generation rules were tested that were designed to treat strong edges that appear in any channel

as being as meaningful as strong edges that appear in both channels. In this section, let $\nabla_s^{(HH)}$ be the normalized gradient at site s from the HH image, $\nabla_s^{(HV)}$ be the normalized gradient for the HV image and ∇_s^{VFG} be the normalized vector field gradient [27] that MIRGS currently uses.

The simplest way of combining strong edges in any of the channels is to take the maximum normalized gradient (MAX):

$$\nabla_s^{MAX} = \max\left\{\nabla_s^{(HH)}, \nabla_s^{(HV)}, \nabla_s^{VFG}\right\} \tag{3.1}$$

In the MAX gradient rule, the gradient value given to MIRGS is the maximum relative strength that occurs at site s from any of the two polarization channels individually or from the dual-polarization feature set. Inclusion of ∇_s^{VFG} covers cases where a site has a weak relative gradient magnitude in each individual channel but a strong relative magnitude when both channels are considered jointly. The MAX ensures that the ∇_s always gives maximum normalized relative strength available at any site.

In an ideal example with no noise, the MAX gradient rule gives the desired results: any edges that are strong in one of the images is guaranteed to have a high edge strength in the combined gradient. However, in the case of sites that have no edge but a non-zero gradient due to noise, the MAX rule will always choose the largest gradient value, which may amplify noise. To alleviate this, two alternative gradient combination rules are proposed.

The first is an Absolute Difference Weighted Average (ADWA) gradient rule. This rule is defined as follows:

$$\nabla_s^{ADWA} = w_s^{ADWA}\nabla_s^{MAX} + (1 - w_s^{ADWA})\nabla_s^{VFG} \tag{3.2}$$

where:

$$w_s^{ADWA} = \frac{|\nabla_s^{(HH)} - \nabla_s^{(HV)}|}{\max_s\left\{|\nabla_s^{(HH)} - \nabla_s^{(HV)}|\right\}} \tag{3.3}$$

In the ADWA rule, the combined gradient weights heavily toward ∇_s^{MAX} if the difference in gradient magnitude between the HH and HV band is large and weights heavily toward ∇_s^{VFG} if the difference is small. Because the gradient magnitude due to noise tends to be small, the difference in gradient magnitude at noisy sites will be small. In this case, the VFG gradient is appropriate since it was designed to minimize noise [27]. If the difference in gradient magnitude is large between the HH and HV channels, then a case where one of the channels has an strong edge and the other does not has been encountered. In this case, the MAX gradient value is appropriate and the ADWA gradient weights toward the MAX gradient. Finally, if the difference in gradient magnitude in HH and HV is small but the site is a true edge, then the edge appears in both images and weighting toward the VFG gradient will be appropriate, which is reflected in the ADWA rule.

The second alternative to the MAX rule is the Canny gradient combination rule (CG). This rule is similar to ADWA but the weight is different:

$$\nabla_s^{CG} = w_s^{CG}\nabla_s^{MAX} + (1 - w_s^{CG})\nabla_s^{VFG} \qquad (3.4)$$

where:

$$w_s^{CG} = \begin{cases} 1 & \text{if site } s \text{ is a local maximum} \\ 0 & \text{otherwise} \end{cases} \qquad (3.5)$$

The local maximum is defined as in the Canny edge detection algorithm [11]: a site s is a local maximum if the gradient magnitude is larger than that of its immediate neighbours in the gradient direction. In the CG rule, maxima are assigned the gradient value of ∇_s^{MAX} since they are more likely to be real edges. Non-maxima are assigned the gradient value of ∇_s^{VFG}.

All three combined gradient rules assign high gradient strength to all three edges for the synthetic example shown previously in Figure 2.8. Figure 3.1 shows the results of using VFG, MAX, ADWA and CG on the same synthetic image with multiplicative Gaussian (mean 1.0, variance 0.1) noise (in SAR imagery, noise is considered multiplicative [37]). The mean of the gradient value is shown for non-edge pixels and edge pixels using each method. Non-edge pixels should have small gradient values because they are caused by noise while edge pixels should

have high gradient values. In the figure, VFG has the lowest sensitivity to noise but also the lowest sensitivity to edges. The other three methods have comparatively higher noise sensitivity along with edge sensitivity but it is not clear from this test which is the best method to use for MIRGS segmentation of R2 imagery.

The experiments in Chapter 4 tested each of the four available gradient combination methods (VFG, MAX, ADWA and CG) for the purposes of combining the gradients of HH & HV dual-polarization data. The combined gradients were then used in MIRGS with the HH & HV dual-polarization feature set in order to determine which method provides the best segmentation accuracy. These results are presented in Chapter 4.

3.2   Image fusion

The literature review in Section 2.8 identified two main methods for image-based fusion: band ratios and wavelet image fusion. These ideas are described and adapted to the dual-polarization R2 data in this section.

3.2.1   HV / HH band ratio

The cross-polarization ratio HV/HH is a relatively simple image fusion technique and was suggested as a useful fusion of information from the HH and HV channels for human visualization [50]. This ratio was also found to be one of the principal variables amongst a multisensor, multifrequency and multipolarization data set [14]. This means that the HV/HH ratio accounts for a significant amount of the variance in the multidimensional data set. Therefore, the HV/HH ratio should be investigated as an image fusion technique.

Calculation of the ratio is straightforward. The backscatter coefficient in HV is simply divided by the backscatter coefficient in HH. Since the HH and HV images are digital numbers on a dB scale, the fusion at each image site s is accomplished simply by:

$$y_s^{HV/HH} = y_s^{(HV)} - y_s^{(HH)} \tag{3.6}$$

The resulting single channel image still has units of dB.
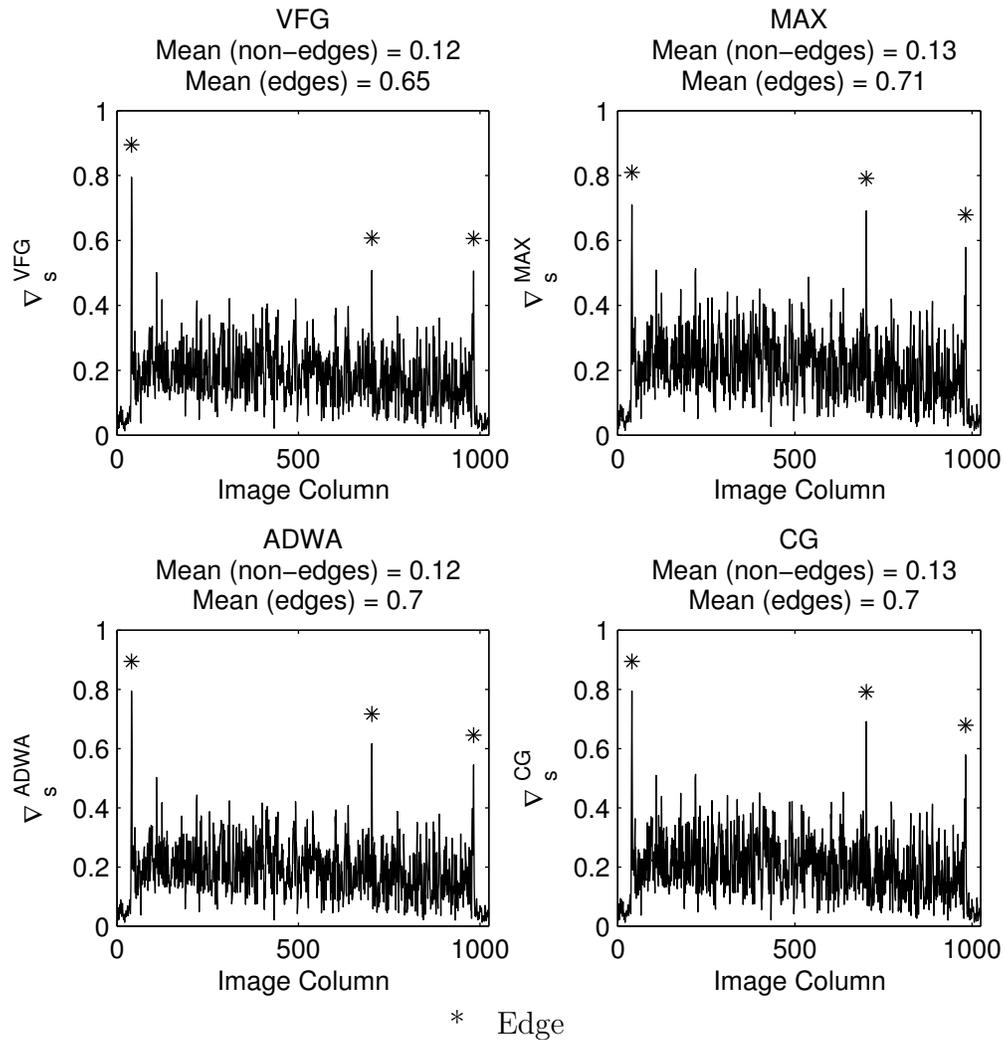
Figure 3.1: VFG, MAX, ADWA and CG gradient combination methods for a noisy synthetic two-channel image with step edges. Mean (non-edges) indicates the mean gradient value over all non-edge pixels with lower being better. Mean (edge) indicates the mean gradient value over all edges pixels, with higher being better.

3.2.2 Wavelet image fusion

As explained in Section 2.8, wavelet-based image fusion has been attempted for a variety of remote sensing applications. Conceptually, wavelet image fusion decomposes the images to be fused into their wavelet coefficient representations and uses various rules to fuse the coefficients into a final set of fused wavelet coefficients, which is then transformed back into a fused image [40]. Wavelet-based image fusion is able to preserve fine details of the original images in the fused image [54]. Here, wavelet image fusion is briefly described, along with the coefficient fusion rules that will be tested.

Wavelets are functions whose energy is limited in time and possess wave-like or periodic characteristics [55]. Signals can be expressed as a summation of scaled and translated wavelets. The wavelet transform finds the summation coefficients for each scaled and translated wavelet in a manner similar to how the Fourier transform finds the summation coefficients for expressing the signal in terms of sinusoidal functions. Since the wavelets are limited in time, unlike sinusoidal functions, a wavelet transform is able to provide both time and frequency information. Due to this, wavelet transforms are well suited for analyzing non-stationary signals like images [2].

The wavelet image fusion approach begins with the multiscale decomposition (MSD) of the images to be fused. MSD decomposes images into a collection of coefficient images at different spatial scales (decomposition level) that can be used to fully reconstruct the original image [66]. Mallat [30] showed that the wavelet transform can be implemented as a MSD with the use of high-pass and low-pass filters to create the wavelet coefficient images. The filters are derived from the wavelet basis that will decompose the images. The MSD can be constructed in stages: one level of decomposition consists of a smoothed approximation coefficient image and three detail coefficient images showing details in the horizontal, vertical and diagonal directions. The approximation image can, in turn, be decomposed into four coefficient images, which represents the next decomposition level.

Let $D(r, c, k, l)$ represent one MSD coefficient image at decomposition level k. Indices r and c refer to the row and column locations in each image and $l = 0 \ldots 3$. When $l = 0$, the coefficient image is referred to as the approximation coefficient image. Higher l correspond to horizontal (1), vertical (2) and diagonal (3) detail coefficient images. The original image can be considered $D(r, c, k, 0)$ with $k = 0$.
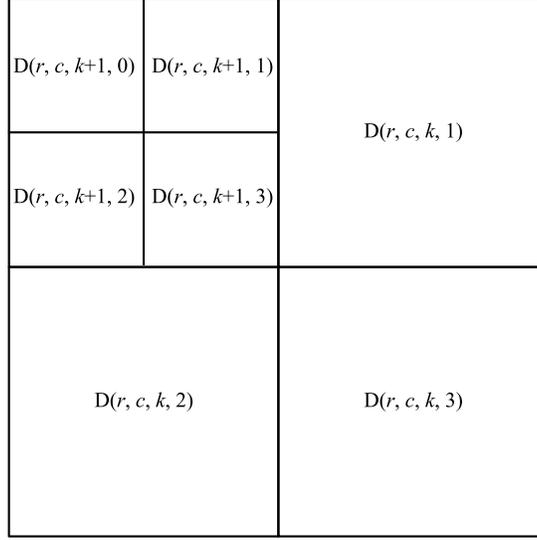
Figure 3.2: A two level wavelet multi-resolution decomposition. Level k contains three detail coefficient images $D(r, c, k, 1 \ldots 3)$. Level $k + 1$ is the highest stage of this decomposition and contains an approximation coefficient image $D(r, c, k + 1, 0)$ and three detail coefficient images $D(r, c, k + 1, 1 \ldots 3)$

A two-level wavelet MSD is shown in Figure 3.2. $D(r, c, k, 1 \ldots 3)$ are detail coefficients corresponding to level k. The original image can be considered $D(r, c, k, 0)$ with $k = 0$. The smoothed approximation at level k was decomposed into the next decomposition level $k + 1$. In this example, the decomposition stopped at level $k + 1$ and so the smoothed approximation $D(r, c, k + 1, 0)$ is kept. The sizes of the coefficient images shown in the figure reflect their relative pixel dimensions. Detail coefficients $D(r, c, k, 1 \ldots 3)$ have relatively high spatial resolution and contains high frequency details. $D(r, c, k + 1, 1 \ldots 3)$ has lower spatial resolution and contains lower frequency details. $D(r, c, k, 0)$ contains the lowest frequency details.

Because it separates the image into high frequency and low frequency coefficient images that retain spatial localization, the wavelet MSD allows the fusing of the coefficients at different scales and locations to be controlled independently. The coefficient fusion process is shown in Figure 3.3. The coefficients in the same location in each corresponding image are combined according to a fusion rule which can make different decisions for each location and scale.

Previous work on using wavelets to fuse SAR imagery used the discrete wavelet transform (DWT) to create the MSD [54]. In this thesis, the stationary wavelet transform (SWT) [36] is used because it was found to produce better image fusion results in both other work [46] and in initial tests with the R2 SAR images, where
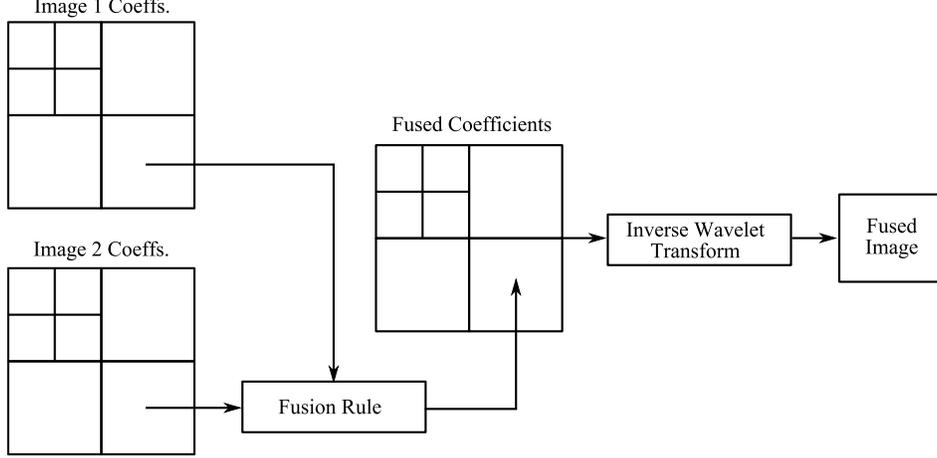
Figure 3.3: Wavelet image fusion is accomplished by decomposing the individual images using the wavelet transform, fusing the corresponding coefficient images into a single fused decomposition and then performing the inverse wavelet transform to obtain the fused image.

the SWT fused images had fewer artifacts. The fusion rules and procedures for the SWT case are fully interchangeable with those of the DWT case.

Let D_1 and D_2 represent the wavelet decomposition of the two SAR channels (HH and HV) to be fused. Let D_F be the decomposition of the fused image. For conciseness, let $p = (r, c, k, l)$ so that $D(r, c, k, l) = D(p)$. The detail coefficients $D_1(r, c, k, 1 \ldots 3)$ and $D_2(r, c, k, 1 \ldots 3)$ can be fused by the choose-max rule:

$$D_F(p) = \max \{D_1(p), D_2(p)\} \quad \text{for} \quad l = 1, \ldots, 3 \qquad (3.7)$$

This rule was found to be the best for detail coefficients in [40]. By choosing the strongest detail coefficients, the wavelet fusion method should create an image that incorporates all the strong edge information present in the HH and HV images, in a way that is similar to the MAX gradient calculation in Section 3.1.

Two fusion rules for the approximation coefficients $D_1(r, c, k, 0)$ and $D_2(r, c, k, 0)$ are tested: the weighted-average (WA) rule, first introduced in [9], and an Absolute Difference Weighted Average (ADWA) rule that has been formulated to take advantage of differences in the HH and HV channels, in a similar manner as the ADWA gradient rule. The WA rule was found to give the best results in [40]. It is formulated as follows [9]:

$$D_F(p) = w_1(p)D_1(p) + w_2(p)D_2(p) \quad \text{for} \quad l = 0 \tag{3.8}$$

where w_1 and w_2 are weights that are assigned based on how well the coefficients at each location match each other. If there is good matching (i.e., the coefficients appear similar), then the assigned weights will average the two coefficients. If matching is poor, then the assigned weights will choose the more salient feature. Salience at a particular row r and column c in the coefficient image is defined as:

$$S(p) = \sum_{r' \in R', c' \in C'} D(r', c', k, l)^2 \tag{3.9}$$

where R' and C' are sets of row and column indices centered around row r and column c, respectively. Salience is computed by considering a small window around (r, c) defined by R' and C'. A 3×3 window was used in this thesis after promising results in initial tests. The match measure between coefficients in D_1 and D_2 is defined as:

$$M_{12}(p) = \frac{2 \sum\limits_{r' \in R', c' \in C'} D_1(r', c', k, l) D_2(r', c', k, l)}{S_1(r', c', k, l) + S_2(r', c', k, l)} \tag{3.10}$$

The match measure is similar to a normalized correlation between the values within the small neighbourhood defined by R' and C'. The weights are assigned as follows. Let $w_{min}(p)$ and $w_{max}(p)$ be two intermediate weights. If $M_{12}(p) \leq \alpha$ then,

$$w_{min}(p) = 0 \tag{3.11}$$
$$w_{max}(p) = 1 \tag{3.12}$$

else if $M_{12}(p) > \alpha$ then:

$$w_{min}(p) = \frac{1}{2} - \frac{1}{2}\left(\frac{1 - M_{12}(p)}{1 - \alpha}\right) \tag{3.13}$$

$$w_{max}(p) = 1 - w_{min}(p) \tag{3.14}$$

Finally, if $S_1(p) > S_2(p)$, then $w_1(p) = w_{max}(p)$ and $w_2(p) = w_{min}(p)$, else $w_1(p) = w_{min}(p)$ and $w_2(p) = w_{max}(p)$. The parameter α is set empirically; the value chosen was 0.5 after initial tests. The WA rule has previously been used for fusing SAR images for visual analysis [54].

The ADWA rule is a simpler fusion rule. It can be defined as follows:

$$w_1(p) = \frac{|D_1(p) - D_2(p)|}{\max_{r,c}\{|D_1(p) - D_2(p)|\}} \tag{3.15}$$

$$w_2(p) = 1 - w_1(p) \tag{3.16}$$

$$D_F(p) = w_1(p)D_1(p) + w_2(p)D_2(p) \tag{3.17}$$

where the maximum value of $|D_1(p) - D_2(p)|$ is taken over all rows and columns at the particular decomposition level and D_1 corresponds to the HH band and D_2 corresponds to the HV band. The ADWA rule emphasizes the coefficients in the HH band when the difference between HH and HV is large. When the difference between the HH and HV band is small, it emphasizes the HV band. For first year ice, the backscatter difference between the HH and HV band is small. This makes the first year ice darker in the fused image because the darker HV band is emphasized. For open water at near range incidence angles, the HH band is much brighter than the HV band. The fused image will therefore retain the HH level of brightness in open water regions. This fusion technique increases contrast between first year ice and open water in the fused image. As seen in Section 2.4 and 2.7, first year ice and open water are frequently difficult to distinguish in the HH band at near range incidence angles. The ADWA rule should create a single image that retains the improved dual-polarization separability between these ice classes.

The wavelet basis chosen for the MSD was the Daubechies wavelet with eight coefficients. Although many other wavelet bases can be used, the tests here are not meant to be an exhaustive investigation of the optimal wavelet basis. In initial tests, the choice of wavelet made very little difference in terms of image appearance or
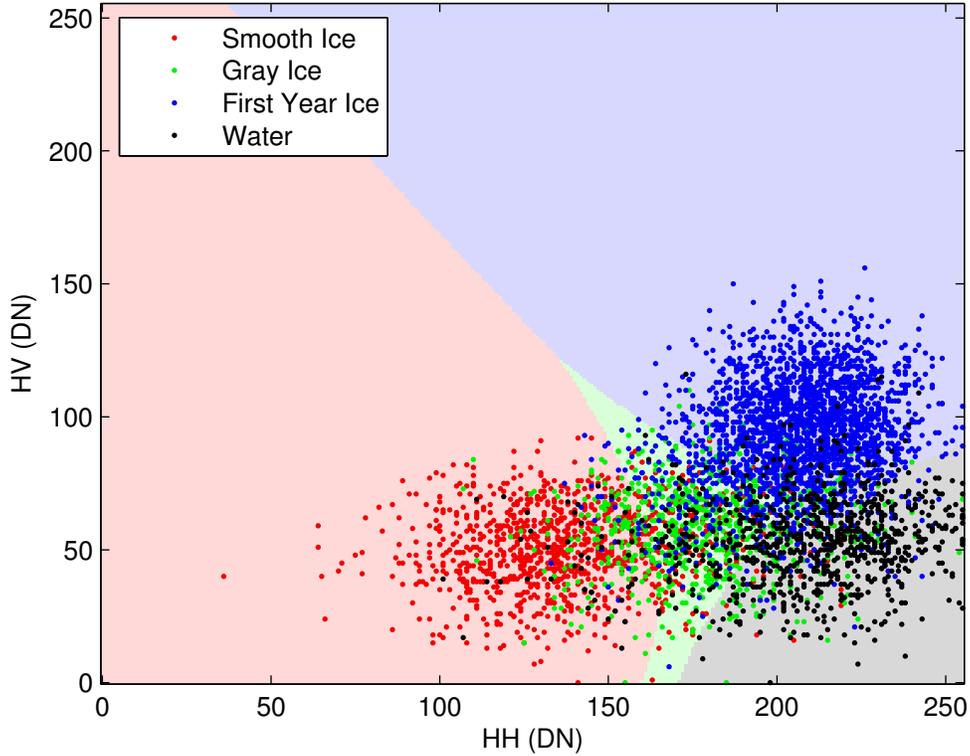
Figure 3.4: The HH & HV feature space plot of the real image (Figure B.1). Light shades in the background indicate the Gaussian maximum likelihood decision boundaries.

class separability. For all tests, four levels of decomposition were used for the MSD as it gave the best results during initial testing.

3.3   Feature space fusion

Figure 3.4 shows a plot of the feature space of the real image (Figure B.1) with the dual-polarization HH & HV feature set. All feature values are integer digital numbers (DN) from $[0, 255]$, which are a direct mapping from backscatter values of $[-35, -5]$ dB. For clarity, only 5000 data points, selected from a regular image grid, are shown. Lighter shades in the background indicate the decision boundaries for a Gaussian maximum likelihood (ML) classifier. Feature space fusion methods attempt to transform the feature space shown in Figure 3.4 into a single dimensional feature space. The fused single image should then theoretically contain edges between all the classes for the spatial context model and retain class separability for the feature space model.
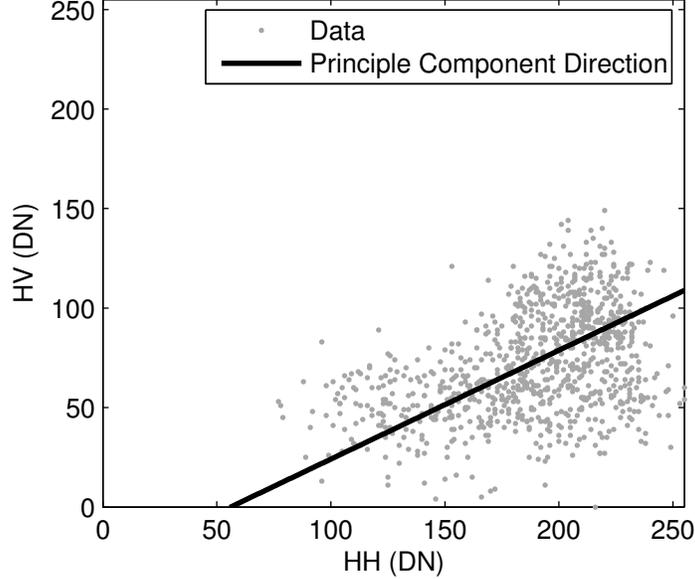
Figure 3.5: The line represents the principal component direction onto which the data points are projected during PCA. The data points shown here correspond to the real image (Figure B.1).

3.3.1 Principal component analysis

As mentioned in Section 2.8, principal component analysis (PCA) can be used as an image fusion technique for multichannel image data. In order to fuse the HH and HV channels, PCA can be regarded as a projection of the 2-D $\mathbf{y}_s$ feature vectors onto the axes in the direction of maximum variance of the data (the principal component direction) [52]. Figure 3.5 shows the feature space for the real image, with the principal component direction shown as a line. Under PCA, all data points (feature vectors) are projected onto the direction of the line. This creates a new, single channel image referred to as PC1 (principal component 1). If the classes are well separated along the PC direction, PCA can be an excellent way of reducing dimensionality while retaining class separability.

Mathematically, the fused PCA channel for each feature vector at site s can be calculated as follows [52]:

$$y_s^{PC1} = \mathbf{e^T}(\mathbf{y}_s - \bar{\mathbf{y}}) \tag{3.18}$$

where $\mathbf{y}_s$ is the feature vector at site s, $\bar{\mathbf{y}}$ is the mean of the feature vectors and $\mathbf{e}$ is the eigenvector corresponding to the largest eigenvalue of the feature vector
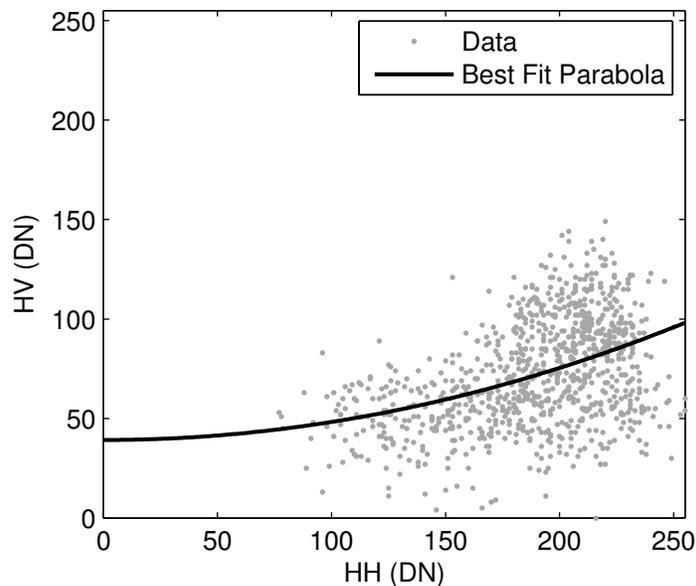
Figure 3.6: A least-squares fit of a parabola to the HH & HV data points in feature space. The data points shown here correspond to the real image (Figure B.1).

covariance matrix.

The locality preserving projection [22] (LPP) is similar to PCA but it explicitly attempts to find a linear projection that preserves distances between points in the transformed feature space. The LPP will keep points that were close in the original space close in the transformed space and likewise with points that were far apart in the original space. This can potentially improve results since it ensures that feature vectors that are dissimilar do not get mapped to similar values. However, in practice with the dual-polarization data, the LPP produced results very similar to PCA and for conciseness was not included in the testing presented here.

3.3.2 Parabolic arc-length projection

Feature space fusion is not limited to linear projections such as PCA. In fact, the feature space distribution of points in Figure 3.4 appears to be non-linear. The classes appear to be dark in both HH & HV, dark in HV but bright in HH, and bright in both HH & HV. There are no data points that are bright in HV and dark in HH. These characteristics produce the shape of the points seen in Figure 3.4. All other R2 scenes are similar.

The non-linear distribution of points suggests that projection of the points onto a coordinate system defined by a non-linear curve may be a useful way of fusing the

two channels. In Figure 3.6, a parabolic curve is shown. This curve was created by least-squares fitting of the coefficients a and b in the following equation:

$$y^{(HV)} = a(y^{(HH)})^2 + c \tag{3.19}$$

to the data points. Projection of each data point to the arc-length coordinate of this parabolic curve represents a non-linear transformation to a univariate feature space which should be able to "unwrap" the non-linear feature space into a single dimension. The projection is accomplished as follows. Let $\mathbf{y}'_s$ be the point on the parabola closest to the feature vector at site s $(\mathbf{y}_s)$:

$$\mathbf{y}'_s = \arg\min_{\mathbf{y}_p} ||\mathbf{y}_s - \mathbf{y}_p|| \tag{3.20}$$

where $\mathbf{y}_p$ is any point on the parabola and $||\mathbf{y}_s - \mathbf{y}_p||$ indicates Euclidean distance between $\mathbf{y}_s$ and $\mathbf{y}_p$. $\mathbf{y}'_s$ is found by minimizing the Euclidean distance equation analytically. When more than one root is found, the one which corresponds to the largest HH $(y^{(HH)})$ value is chosen. The parabolic arc-length (PAL) coordinate is calculated from the standard arc-length formula for the quadratic curve in Equation 3.19:

$$y_s^{PAL} = \int_0^{y'^{(HH)}_s} \sqrt{1 + (2ay_p^{(HH)})^2} dy_p^{(HH)} \tag{3.21}$$

where a is the least-squares fitted coefficient from Equation 3.19 and PAL_s is the Parabolic Arc-Length coordinate of site s, where the arc-length is measured from $y^{(HH)} = 0$. Equation 3.21 is solved analytically.

Other curves can also be used. For example, a general parabola $y^{(HV)} = a(y^{(HH)})^2 + by^{(HH)} + c$ can also be considered. However, the general parabola sometimes produces an inverted parabola when fitted to the data, which does not give the desired non-linear "unwrapping" of the feature space. Other families of curves can also be tested in future work; the parabolic curve is tested here as

42

a representative of non-linear projections using curve fitting because there are analytical solutions to the above equations.

3.3.3   Non-linear dimensionality reduction techniques

In addition to the PCA and PAL techniques, three non-linear dimensionality reduction (NLDR) techniques were also used to perform feature space fusion: locally linear embedding (LLE) [47], Laplacian eigenmaps (LEIGS) [5] and local tangent-space alignment (LTSA) [67]. In many ways, these techniques are similar to PCA or LPP but rather than projecting onto a linear axis, they attempt to project the data onto a non-linear "manifold" on which the data are assumed to lie. The PAL technique assumed that the data lie on a parabola due to the backscatter characteristics of dual-polarization data. The NLDR techniques learn the manifold coordinates from the data points themselves, with no assumption for a particular manifold shape.

Conceptually, NLDR methods assume that the image feature vector at site s arises from:

$$\mathbf{y}_s = f(\mathbf{\Psi}_s) + \epsilon \tag{3.22}$$

where f is some unknown non-linear function or process that forms a manifold in the feature space of $\mathbf{y}_s$, $\mathbf{\Psi}_s$ are the manifold coordinates of site s and ϵ is noise [67]. NLDR methods find $\mathbf{\Psi}_s$ without explicitly knowing f. The three NLDR methods considered here operate similarly: for every feature vector $\mathbf{y}_s$, the local geometry as defined by the k nearest Euclidean distance neighbours in feature space is determined and manifold coordinates are then found that in some way preserves the local geometry for all the original feature vectors. The three methods mainly differ in the nature of the local geometry that they preserve. In the following discussion, let n_{train} represent the number of feature vectors that are given to each NLDR algorithm to learn the manifold coordinates.

LLE [47] assumes that the manifold is locally linear. Then an $n_{\text{train}} \times k$ weight matrix $\mathbf{W}$ is computed by minimizing the reconstruction error $\mathcal{E}(\mathbf{W})$:

$$\mathcal{E}(\mathbf{W}) = \sum_s |\mathbf{y}_s - \sum_{j=1}^{k} w_{sj} \mathbf{y}_{s_j}|^2 \qquad (3.23)$$

where $\mathbf{y}_{s_j}$ refers to one of the k nearest neighbours of the feature vector at site s. Minimizing Equation 3.23 via least-squares fitting finds the set of weights that best linearly reconstructs each of the feature vectors from its k nearest neighbours. The weights are constrained such that $\sum_{j=1}^{k} w_{sj} = 1$ in order to make the weights invariant to translation and the form of Equation 3.23 makes them invariant to rotation and scaling [47]. LLE assumes that there is a linear mapping between the original feature space and the manifold coordinates on a local level that consists of a translation, rotation and scaling. Since the weights were designed to be invariant to these three transformations, the manifold coordinates $\mathbf{\Psi}_s$ of a data point can be written as a linear combination of the manifold coordinates of its k nearest neighbours $\mathbf{\Psi}_{s_j}$ with the exact same set of weights. Therefore, choosing $\mathbf{\Psi}$ (the set of all $\mathbf{\Psi}_s$ for every site s in the scene) to minimize the cost function $\Phi(\mathbf{\Psi})$:

$$\Phi(\mathbf{\Psi}) = \sum_s |\mathbf{\Psi}_s - \sum_{j=1}^{k} w_{sj} \mathbf{\Psi}_{s_j}|^2 \qquad (3.24)$$

will give the desired manifold coordinates. In Equation 3.24, the w_{sj} are fixed and the optimization is performed to find a set of $\mathbf{\Psi}_s$ that globally minimizes the expression. The above problem is converted into an eigenvalue problem and solved to obtain the manifold coordinates. LLE preserves the reconstruction weights of each of the k nearest neighbours of each data point. Only one parameter, k, needs to be chosen. This was set to $k = 8$ after tests from 4 to 16 neighbours showed little difference in results.

LEIGS [5] constructs an adjacency graph with feature vectors $\mathbf{y}_s$ as nodes. Any two nodes are connected by an edge if at least one of the nodes is among the k nearest neighbours of the other. A $n_{\text{train}} \times n_{\text{train}}$ weight matrix $\mathbf{W}$ is constructed where $w_{ij} = 1$ if nodes i and j are connected on the graph. LEIGS then finds the set of manifold coordinates $\mathbf{\Psi}$ minimizes the cost function $\Phi(\mathbf{\Psi})$:

$$\Phi(\mathbf{\Psi}) = \sum_{ij} (\mathbf{\Psi}_i - \mathbf{\Psi}_j) w_{ij} \tag{3.25}$$

where the summation is done over all pairs of nodes in the graph. The above minimization problem is converted algebraically to an eigenvalue problem subject to a scaling constraint in the manifold coordinates; the full details are in [5]. The idea behind LEIGS is that the manifold coordinates must map points that are close together in the original space (as indicated by $w_{ij} = 1$) to points that are close together on the manifold. The only parameter that needs to be chosen is k, the number of nearest neighbours. This value was again set to $k = 8$ after initial testing from 4 to 16 neighbours revealed little difference in the results.

LTSA [67] uses the k nearest neighbours of a feature vector to estimate the local manifold tangent space by finding the best fit n-dimensional hyperplane for the points, where n is the number of dimensions in the original data. Once the local tangent space is found, its k nearest neighbours are converted to local tangent space coordinates $\boldsymbol{\theta}_{s_j}$, where $j = 1 \ldots k$ to indicate the nearest neighbour points. LTSA assumes that there is an affine transformation L_s that approximately transforms the tangent space coordinates to manifold coordinates:

$$\mathbf{\Psi}_{s_j} = \bar{\mathbf{\Psi}}_{s_j} + L_s \boldsymbol{\theta}_{s_j} + \epsilon_{s_j} \tag{3.26}$$

where $\bar{\mathbf{\Psi}}_{s_j}$ is the mean of the manifold coordinates for the k nearest neighbours and ϵ_{s_j} is a reconstruction error. LTSA then finds the set of L_s and $\mathbf{\Psi}_{s_j}$ that minimizes the total reconstruction error over all of the original data points:

$$\sum_s \sum_{j=1}^{k} \epsilon_{s_j} = \sum_s \sum_{j=1}^{k} \mathbf{\Psi}_{s_j} - \bar{\mathbf{\Psi}}_{s_j} - \mathbf{L}_s \boldsymbol{\theta}_{s_j} \tag{3.27}$$

The above problem is algebraically converted to an eigenvalue problem and the optimal manifold coordinates L_s and $\mathbf{\Psi}_{s_j}$ are found. Then, all information to produce the manifold coordinates $\mathbf{\Psi}_s$ for each feature vector is available. As with LLE and LEIGS, the only parameter that needs to be set is k, the number of nearest neighbours and the same value of $k = 8$ was used.

All three NLDR techniques required finding the eigenvectors of matrices with $(MN)^2$ entries, where M is the image width in pixels and N is the image height. In order to reduce the computational requirements, the original images were resampled via bicubic interpolation to have a maximum of 4096 pixels prior to the NLDR process. This was chosen because it is the largest number of points that could be handled without exhausting available memory (2 GB) on the test system. The 4096 data points form the training set for learning the manifold. Once the NLDR technique is applied, the feature vector $\mathbf{y}_t$ at each site t in the resampled image will have been mapped to manifold coordinates $\mathbf{\Psi}_t$:

$$\mathbf{y}_t \rightarrow \mathbf{\Psi}_t \tag{3.28}$$

Assuming that the training feature vectors allowed the NLDR technique to learn a reasonably accurate manifold, a method is needed to map all of the original feature vectors $\mathbf{y}_s$ to coordinates on the learned manifold. The manifold, however, is defined only on the training points and cannot be extended to other points. A method to estimate the manifold coordinates of points outside of the training set has been suggested [28]. First, the mapping between each training feature vector and its manifold coordinates is approximated as an affine transform $\mathbf{L}_t$:

$$\mathbf{\Psi}_t - \bar{\mathbf{\Psi}}_t = \mathbf{L}_t(\mathbf{y}_t - \bar{\mathbf{y}}_t) \tag{3.29}$$

where $\bar{\mathbf{\Psi}}_t$ and $\bar{\mathbf{y}}_t$ are the mean of the transformed and original coordinates of the data point and its k nearest neighbours in the training set. This neighbourhood is the same as that used by the NLDR technique for learning the manifold. The transform $\mathbf{L}_t$ is then calculated [28]:

$$\mathbf{L}_t = (\mathbf{\Psi}_t - \bar{\mathbf{\Psi}}_t)(\mathbf{y}_t - \bar{\mathbf{y}}_t)^{\dagger} \tag{3.30}$$

where $(.)^{\dagger}$ is the pseudoinverse operator which operates on vectors by treating them as non-square matrices. Once $\mathbf{L}_t$ is computed for each training point, mapping of the original data points is accomplished by:

$$\mathbf{\Psi}_s = \bar{\mathbf{\Psi}}_{t'} + \mathbf{L}_{t'}(\mathbf{y}_s - \bar{\mathbf{y}}_{t'}) \tag{3.31}$$

where t' indicates the site of the training feature vector that is closest to $\mathbf{y}_s$ in feature space. Since $\mathbf{y}_{t'}$ and $\mathbf{y}_s$ are nearest neighbours, their transformed coordinates on the manifold should be similar. Thus, the affine transformation $\mathbf{L}_{t'}$ can be used to estimate the manifold coordinates of $\mathbf{y}_s$.

All image feature vectors are mapped using this estimation method. The resulting NLDR feature vectors still have two dimensions and fusion to one image is accomplished by discarding one of the two dimensions. The most significant NLDR channel corresponds to the first channel and retains the most information. The least significant NLDR channel is the second channel and is usually discarded.

For efficient nearest neighbour searching to find training site t' in Equation 3.31 for each site s in the image, a KD-Tree [6] data structure was implemented. This is a space partitioning scheme for multidimensional data points that improves an exhaustive $O(MNn_{\text{train}})$ nearest neighbour search to an $O(MN \log n_{\text{train}})$ search on average. Tests showed that finding the nearest neighbour from a set of 4096 training points typically required searching fewer than 100 points, significantly improving the speed of the fusion process.

The implementation of each of the three NLDR methods was obtained from [60], which is a collection of code for NLDR techniques provided by the original authors. Another popular NLDR technique called ISOMAP [56] was also available but it was not tested because the provided code could not handle the transformation of 4096 data points. Using a much lower number of data points was tested but did not generate useful results, so these results, for conciseness, were not included here.

3.4 Parameter selection

Parameter selection of the C_1 and C_2 parameters in Equation 2.6 can affect results by changing the value of the spatial context model weighting β. In the parameter selection experiments, each of the methods explained in the previous sections are tested with MIRGS set at different values of the parameters. The range of C_1 tested was $[1.0, 7.0]$. Values below this range were unacceptably noisy and oversegmented and values above this range results were overmerged excessively. The range of C_2

47

tested was from $[0.1, 1]$. This was the range which has been found empirically to give reasonable results. In actual usage of MIRGS, C_1 is generally set first and then C_2 is used to "fine-tune" the segmentation results.

3.5   Hierarchical segmentation

The methods in the previous sections have focused on fusing the information from the dual-polarization bands before the segmentation process is executed. In this section, a hierarchical segmentation method is introduced which performs the segmentation in stages, where one segmentation result is used to guide the next one. Hierarchical segmentation generates image segments in a nested fashion [24]. The image is first partitioned into several segments, which are each partitioned into further segments. This partitioning can be continued until a desired number of total classes is reached, a stopping criteria is met or until each pixel is assigned its own class.

3.5.1   Overview

The proposed approach is to first segment the HV channel with MIRGS into as many classes as it can distinguish. Since the HV channel does not have sufficient information to distinguish all classes, as seen in Section 2.7, each of the obtained segments is then further partitioned using MIRGS but with information from the HH channel.

   The motivation for considering such a hierarchical approach for segmentation can be seen in Figure 3.7, which is a plot of the probabilities of a pixel having a certain value in the HH or HV channel given that it belongs to a certain class. These plots were generated with the ground-truth for the real R2 image shown in Figure B.1. In the plot for the HH channel, first year ice and water overlap and are not separable. However, these two classes are separable in the HV channel, although smooth ice, gray ice and water are not. If the proposed segmentation approach is used, then the HV channel will be segmented into two segments: one containing first year ice and another containing the three remaining classes. If this other segment is then input into MIRGS and segmented into three classes with the HH channel, then the result should correctly separate the three remaining classes.

   A schematic of the hierarchical segmentation process for this image is shown in Figure 3.8. When interpreting the hierarchies, the order of the branches from each

channel indicates the brightness of the class mean in the channel. The left most branch is the darkest and the rightmost branch is brightest.

This method is a way of using the enhanced class separability information given by two channels but using only one channel at a time. The reason for this is to help deal with the incidence angle effect which causes problems for MIRGS as described in Section 2.7 in Figure 2.9. Initially segmenting the HV channel into two classes allows the darker open water to be separated from first year ice since the appearance of water is insensitive in HV to incidence angle.

One question with this approach, however, arises when Figure 3.7(b) is considered again. It suggests that if the HV image is segmented into two segments, then one segment will consist of open water, gray ice and smooth ice that must be further segmented into three classes by segmenting the HH image, as in Figure 3.8. For an image without the incidence angle effect, such as the real image (Figure B.1), this does not pose a problem. However, the HH image may not be able to properly segment the open water / gray ice / smooth ice portion of the image if the water changes appearance due to the incidence angle effect.

Two possibilities exist to mitigate the aforementioned problem. First, Figure 3.7 only shows feature space separability. It is possible that gray ice will be separable from open water and grouped with first year ice due to edge strength and spatial context, which MIRGS inherently uses. This was the case for the full scene data set (Figure B.3) that was available for testing. Smooth ice is more problematic, but it is very dark in the HH band and should be separable in that channel from wind roughened open water regardless of incidence angle. There was no example of smooth ice occurring in an image with the incidence angle effect to test this assertion.

The second mitigating factor is that the HV channel could be segmented into more than two segments. If the HV image can be separated into open water / smooth ice, gray ice and first year ice, then the HH channel should be able to separate smooth ice from open water, as described above while gray ice and first year ice are already separate.

The test data provided by the CIS does not represent all the ice types listed in Table 2.1, so the observations in this section will need to be amended in future work as additional data becomes available.
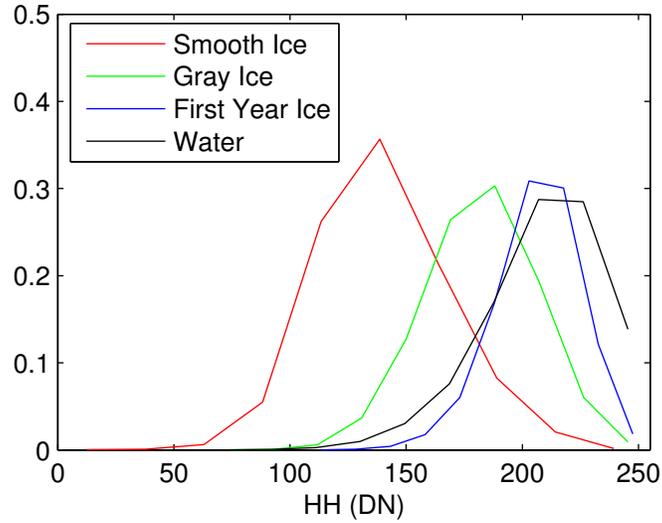
(a) $P(y^{(HH)}|x)$
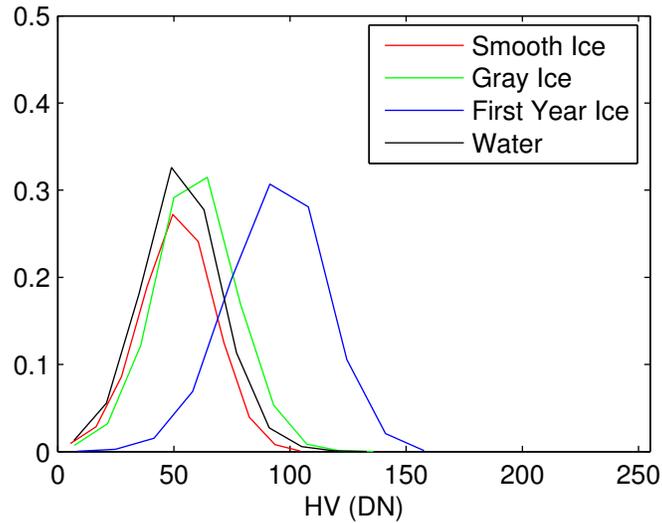
(b) $P(y^{(HV)}|x)$

Figure 3.7: Class conditional probability distribution functions for backscatter values for the real image (Figure B.1) from (a) HH and (b) HV. The separability of first year ice and water is poor in the HH channel but good in the HV channel. However, the separability of smooth ice, gray ice and water is poor in HV.
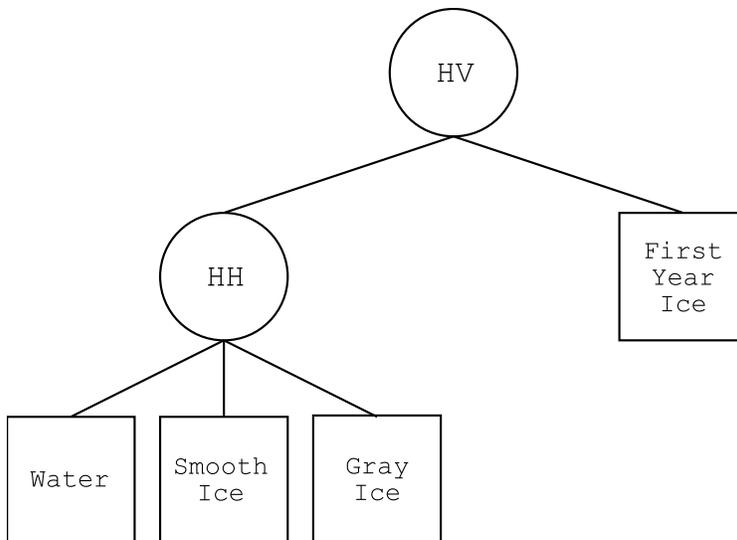
Figure 3.8: An example sea ice type hierarchy for hierarchical segmentation. The order of the branches from each channel from left to right indicates the brightness of each class in the channel from darkest to brightest.

3.5.2 Implementation

Two approaches for creating sea ice type hierarchies (SITH) like that shown in Figure 3.8 were tested. The first approach is simple: the hierarchies are user-defined based on domain knowledge. User-defined SITHs allow human operators to control precisely how many classes are segmented at each node in the hierarchy. This complicates the segmentation process because the user must define more than just the number of classes. For operational use, however, this should be acceptable as human experts should have some intuition to determine which hierarchy is appropriate for each scene.

The second approach for creating a SITH (Auto-SITH) is based on estimating the number of segments that can be distinguished in the HV channel. The HV channel is then segmented into precisely that number of segments. The number of subsegments in each segment that is distinguishable in the HH channel is then estimated and the segmentation is performed. To estimate the number of segments that can be distinguished, the Davies-Bouldin (DB) index [16] is used. The DB index is defined as:

$$DB(\eta) = \frac{1}{\eta} \sum_{i=1}^{\eta} R_i \tag{3.32}$$

51

where

$$R_i = \max_{j \neq i} \{R_{j,i}\} \tag{3.33}$$

$$R_{j,i} = \frac{e_j + e_i}{m_{j,i}} \tag{3.34}$$

$DB(\eta)$ evaluates a particular segmentation result with η classes, with e_j being the average Euclidean distance between the feature vectors assigned to class j and the centroid μ_j of class j and $m_{j,i} = ||\mu_j - \mu_i||$, the Euclidean distance between the two class centroids. The e_j can be considered the average error for class j in the current segmentation result. To estimate the number of classes in a scene, the DB index is evaluated for segmentation results using $\eta = 2, \ldots, n$ classes. The η with the smallest DB index is chosen as the number of classes present in the scene. The smallest DB index is achieved for segmentation results whose clusters in feature space are small (low e_j) and have large separation (large $m_{j,i}$).

The error terms e_j used in the DB index is defined in terms of Euclidean distance only, so the segmentation results used for estimating the number of classes were the initial region-based K-means segmentations rather than the MIRGS segmentation. This is because usage of the MIRGS segmentation results would require a new error term e_j to be derived, but it is currently unclear how to incorporate the MIRGS spatial context model into the error term.

In determining the number of classes for the HV channel, $\eta = 2, \ldots, n$, where n is the total number of classes in the scene. For determining the number of subsegments in each of the obtained HV segments, $\eta = 2, \ldots, (n - n^{HV} + 1)$. This is based on the fact that when the HV channel is divided in to n^{HV} segments, each of these segments can only have a maximum of $(n - n^{HV} + 1)$ classes if there are n classes in total. If $(n - n^{HV} + 1) \leq 2$, then the number of classes is set to $(n - n^{HV} + 1)$.

It should be noted that the DB index can only determine the number of classes for $\eta > 1$. An alternative method, outlined in [18], determines whether a single cluster should be split into two clusters but it does not determine the number of classes beyond that. This was not implemented because if a one class segment is oversegmented to two classes, a post-process merging (see below) can be used to eliminate the extra class. Other indices, such as the Modified Hubert (MH) index [24], can also be used instead of the DB index. However, the DB index is

attractive because it only requires the η at which the minimum value of the index occurs to be found rather than having to determine the η at which a "significant knee" in the index value curve occurs [24].

A method similar to K-means iterative Fisher (KIF) clustering [13] could also be used to automatically generate the proper hierarchies. At each stage, KIF always clusters segments into two subsegments; if the Fisher criterion between the two classes is below a certain threshold $\hat{\tau}$, then the clusters are merged since they likely represent only one class. Such a scheme is not implemented here because an appropriate $\hat{\tau}$ threshold has not yet been found for each of the dual-polarization channels. More image data will be necessary before this can be done in future work.

Two additional generic sea ice type hierarchies are also tested, shown in Figure 3.9. These are used to provide additional test cases for the test scenes.

When using both Auto-SITH and the generic SITH 2-3 or 3-2, the final number of classes may exceed the number of classes that the user desires for the scene. In the case of Auto-SITH, the DB index only determines the number of classes as being two or more. For the generic SITHs, the output is always six classes. To ensure that the final number of classes is equal to the number of classes that the user has chosen, the final segmentation result is post-processed to merge any excess classes. This is performed by repeated pairwise merging of the two closest classes according to the Fisher criterion [18] until the desired number of classes is reached. The experiments will test which feature set (HH, HV or HH & HV) should be used to calculate the Fisher criterion during merging. It is also possible to use a spatial context aware merging criterion such as that present in Step 4 (Section 2.6.4) of MIRGS but the classes that should be merged are not necessarily adjacent to each other. Therefore, only the feature space merging with the Fisher criterion is considered at this time.
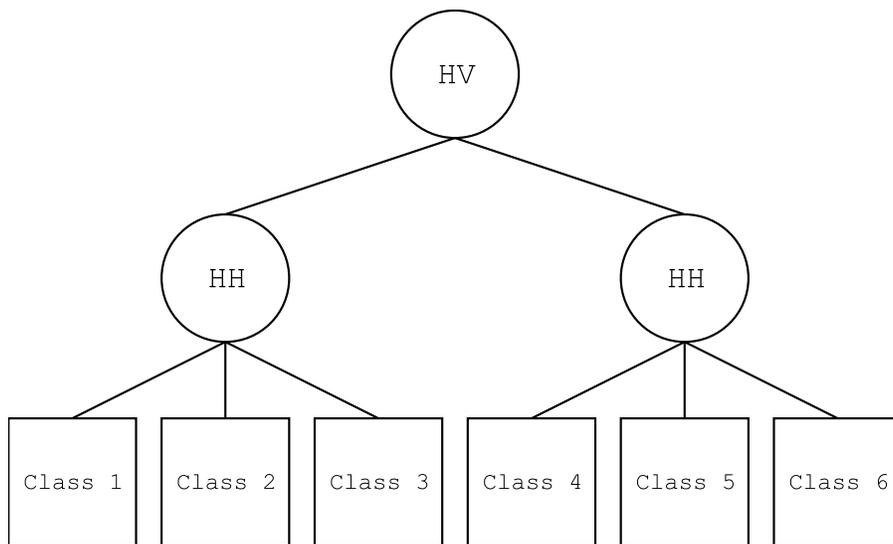
(a) SITH 2-3
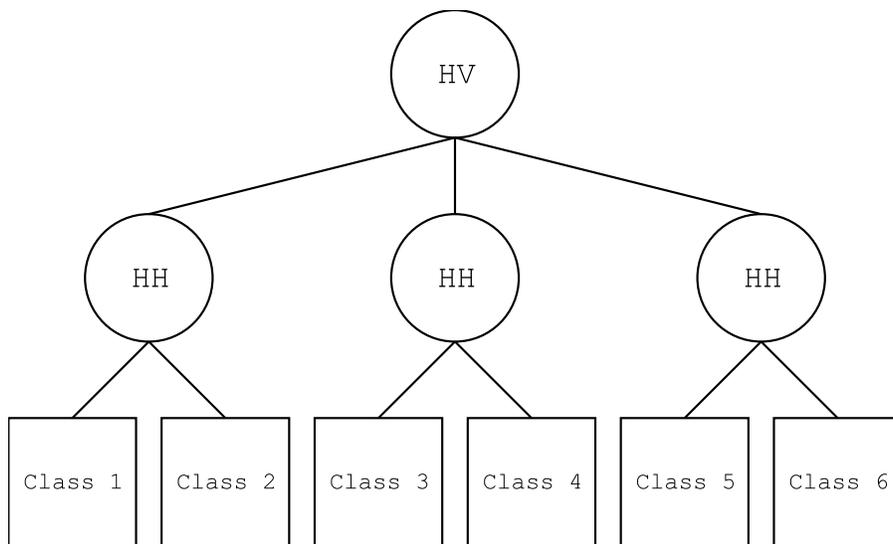
(b) SITH 3-2

Figure 3.9: Two generic sea ice type hierarchies that were also tested.

Chapter 4

Experimental Results and Discussion

The methods described in Chapter 3 were tested on three separate images. The first image is the real R2 image seen in Figure B.1. This image has a manual segmentation as ground-truth and represents ice types at near range incidence angles ($0°$ to $10°$). The second test image is a synthetic dual-polarization image created as described in Section B.2 and shown in Figure B.2. The ground-truth for the real image may contain manual segmentation errors whereas the ground-truth for the synthetic image is unambiguous. The gray levels in each of the classes in the synthetic image were derived from real R2 data and represent ice types at the $30°$ to $35°$ incidence angle range. There was a lack of unambiguous ice type samples at other incidence angle ranges to generate synthetic images for those ice types. The hierarchical segmentation techniques were tested with the real and synthetic images, as well as the third test image, which is the full scene image shown in Figure B.3. This image lacks ground-truth but Figure B.3(c) is a reasonable "expected" segmentation derived by manually combining several MIRGS segmentation results. Due to the lack of actual ground-truth, only qualitative evaluation for the full scene image was performed.

4.1   Data fusion for improved accuracy

This section presents the results of testing gradient generation, image fusion and feature space fusion techniques. All experiments performed in this section followed

the same procedure. Each technique is applied to the calibrated dual-polarization images and MIRGS is used to obtain a segmentation result. The segmentation result is then compared to the ground-truth image. Two measures were used for measuring segmentation accuracy: the overall accuracy, which is the percent of pixels correctly segmented and the κ coefficient [7],[15]. The Kappa coefficient is defined as follows [45]:

$$\kappa = \frac{P \sum_{k} x_{kk} - \sum_{k} x_{k+} x_{+k}}{P^2 - \sum_{k} x_{k+} x_{+k}} \tag{4.1}$$

where x_{ij} is the j-th entry on the i-th row of the segmentation error matrix, $x_{i+} = \sum_j x_{ij}$ and $x_{+j} = \sum_i x_{ij}$. P represents the total number of pixels in the image. κ is an accuracy assessment measure that ranges from $[-1, 1]$ and compares the segmentation result to random assignment. When $\kappa = 0$, the segmentation result is as good as random assignment. When $\kappa = 1$, the segmentation is perfect. Negative κ indicate results that are biased against the proper segmentation.

Overall accuracy and κ give two correlated but distinct measurements of accuracy. Prior to accuracy assessment, the segmented results are relabeled to match the ground-truth image labels to maximize overall accuracy. This process is necessary because the segmentation algorithm may assign a label value $(1, \ldots, n)$ to pixels of a certain ice type that is different from the label value $(1, \ldots, n)$ assigned to the same ice type in the ground-truth image or in another segmentation result.

Table 4.1 lists all the data fusion methods and feature sets that were tested. Both original bands are tested with MIRGS individually in order to provide a comparison with single-polarization data. The gradient combination rules are tested using HH & HV bands as the feature set given to MIRGS. The single images created by the image fusion techniques are also tested. LEIGS 1 & 2, LLE 1 & 2 and LTSA 1 & 2 use both NLDR channels as the multivariate feature set for MIRGS and were included to test whether the NLDR techniques transform the original two-dimensional feature space to another two-dimensional feature space with better separability. The NLDR transforms are very non-linear and the results are difficult to predict in advance. LEIGS 2, LLE 2 and LTSA 2, which are the second and least significant NLDR channels, are included for completeness.

Table 4.1: Data fusion methods and feature sets tested.

Name	Feature Set Used
Single Polarization	
HH	HH band only.
HV	HV band only.
Gradient Generation Methods for Dual-Polarization Data	
HH & HV, ADWA Gradient	HH & HV with gradient combination using the ADWA rule.
HH & HV, VFG Gradient	HH & HV with existing VFG gradient generation method.
HH & HV, MAX Gradient	HH & HV with gradient combination using the MAX rule.
HH & HV, CG Gradient	HH & HV with gradient combination using the CG rule.
Image Fusion Techniques	
ADWA Wavelet	Image fused with the wavelet ADWA rule.
WA Wavelet	Image fused with the wavelet WA rule.
HV / HH	Band ratio image.
Feature Space Fusion Techniques	
LEIGS 1	First LEIGS channel.
LEIGS 1 & 2	Both LEIGS channels (with VFG gradient combination).
LEIGS 2	Second LEIGS channel.
LLE 1	First LLE channel.
LLE 1 & 2	Both LLE channels (with VFG gradient combination).
LLE 2	Second LLE channel.
LTSA 1	First LTSA channel.
LTSA 1 & 2	Both LTSA channels (with VFG gradient combination).
LTSA 2	Second LTSA channel.
PAL	Fused PAL channel.
PCA	Fused PCA channel.

Table 4.2: Accuracy statistics for tested data fusion techniques on the real image (Figure B.1) at different values for the MIRGS merging parameter C_1. C_2 for each method was fixed at 0.4.

Method	$C_1 = 1$ Overall (%)	κ	$C_1 = 3$ Overall (%)	κ	$C_1 = 5$ Overall (%)	κ	$C_1 = 7$ Overall (%)	κ
ADWA Wavelet	62.78	0.48	74.14	0.61	*75.24*	*0.62*	74.59	0.60
HH	56.52	0.40	61.68	0.46	68.85	0.55	*72.14*	*0.57*
HH & HV, ADWA	80.44	0.72	*83.92*	*0.77*	78.40	0.69	75.20	0.64
HH & HV, VFG	**80.85**	**0.73**	**84.70**	**0.78**	77.48	0.68	71.17	0.59
HH & HV, MAX	80.77	0.73	*84.19*	*0.77*	78.64	0.69	72.71	0.61
HH & HV, CG	80.62	0.72	*83.92*	*0.77*	79.72	0.71	73.51	0.62
HV	40.86	0.21	46.27	0.28	*48.37*	*0.30*	47.46	0.29
LEIGS 1	65.59	0.54	67.14	0.56	72.00	0.62	*73.91*	*0.64*
LEIGS 1 & 2	65.92	0.54	70.25	0.59	73.78	0.64	*74.97*	*0.65*
LEIGS 2	51.28	0.35	54.69	0.39	59.11	0.44	*60.52*	*0.46*
LLE 1	64.26	0.52	72.28	0.62	*75.59*	*0.66*	72.70	0.62
LLE 1 & 2	79.55	0.71	*84.13*	*0.77*	**81.65**	**0.74**	**75.93**	**0.66**
LLE 2	49.53	0.30	56.29	0.37	*61.51*	*0.42*	58.27	0.35
LTSA 1	40.21	0.20	46.30	0.27	*49.30*	*0.31*	48.12	0.29
LTSA 1 & 2	61.35	0.48	62.40	0.50	*62.88*	*0.50*	57.09	0.42
LTSA 2	46.07	0.29	48.76	0.33	*50.38*	*0.29*	49.63	0.27
PAL	57.01	0.42	64.37	0.50	68.07	0.54	*68.78*	*0.55*
PCA	59.65	0.45	66.95	0.54	69.72	0.57	*70.55*	*0.58*
WA Wavelet	61.25	0.45	*72.52*	*0.59*	72.44	0.57	67.75	0.48
HV / HH	51.04	0.32	56.80	0.39	*57.51*	*0.41*	52.81	0.37

Bold Best result for given C_1 *Italic* Best result for given method   **Bold** Best result

4.1.1 Real image

Table 4.2 shows the accuracy assessment results for the real image (Figure B.1) for all data fusion methods. The output of each method was input into MIRGS, which was then run at four different settings of the C_1 parameter, with $C_2 = 0.4$. Cells highlighted in yellow represent the best C_1 value for a particular method, while those in light green represent the best data fusion method for a particular C_1 value. The best accuracies are obtained with multivariate feature sets (HH & HV, LLE 1 & 2) at $C_1 = 3$. All of these perform very similarly, but the best accuracy obtained in the table was provided by HH & HV, VFG Gradient. The proposed gradient combination rules did not improve accuracy for the real image. It appears that MIRGS is not very sensitive to the gradient combination method and all four gradient generation methods produce very similar results.

Many of the image fusion and feature space fusion techniques (ADWA Wavelet, WA Wavelet, LEIGS 1, LLE 1) produce accuracy results better than the best HH or HV channels alone. This is expected as they each attempt to fuse the dual-polarization information into a single band. These fused images are unable to provide the accuracy of the dual-polarization feature set HH & HV, indicating that

Table 4.3: Accuracy statistics for tested data fusion techniques on the real image (Figure B.1) at different values for the MIRGS merging parameter C_2. C_1 for each method was fixed at the best value for the method as indicated by Table 4.2.

Method	$C_2 = 0.1$ Overall (%)	κ	$C_2 = 0.4$ Overall (%)	κ	$C_2 = 0.7$ Overall (%)	κ	$C_2 = 1$ Overall (%)	κ
ADWA Wavelet	*75.67*	*0.63*	75.24	0.62	75.04	0.62	75.46	0.63
HH	70.69	0.54	*72.14*	*0.57*	67.33	0.53	67.36	0.53
HH & HV, ADWA	83.81	0.77	83.92	0.77	84.06	0.77	*84.15*	*0.77*
HH & HV, VFG	84.21	0.77	84.70	0.78	84.71	0.78	**84.79**	**0.78**
HH & HV, MAX	84.07	0.77	84.19	0.77	84.11	0.77	*84.28*	*0.77*
HH & HV, CG	83.89	0.77	83.92	0.77	84.15	0.77	*84.22*	*0.77*
HV	46.46	0.28	*48.37*	*0.30*	47.84	0.30	47.75	0.29
LEIGS 1	73.87	0.64	73.91	0.64	*74.02*	*0.64*	73.49	0.63
LEIGS 1 & 2	75.24	0.65	74.97	0.65	*75.48*	*0.66*	75.30	0.65
LEIGS 2	57.86	0.43	60.52	0.46	*60.50*	*0.46*	59.93	0.45
LLE 1	75.35	0.65	75.59	0.66	75.13	0.65	*75.62*	*0.66*
LLE 1 & 2	84.10	0.77	*84.13*	*0.77*	84.00	0.77	84.12	0.77
LLE 2	60.31	0.39	*61.51*	*0.42*	60.38	0.41	60.11	0.41
LTSA 1	49.07	0.30	*49.30*	*0.31*	49.01	0.30	48.42	0.30
LTSA 1 & 2	58.82	0.45	*62.88*	*0.50*	62.18	0.49	62.22	0.49
LTSA 2	49.78	0.27	*50.38*	*0.29*	50.04	0.29	49.57	0.28
PAL	67.02	0.52	68.78	0.55	*69.12*	*0.55*	68.11	0.54
PCA	68.93	0.55	*70.55*	*0.58*	69.97	0.58	70.44	0.58
WA Wavelet	*73.12*	*0.59*	72.52	0.59	72.38	0.59	71.68	0.58
HV / HH	52.81	0.37	57.51	0.41	*59.18*	*0.43*	57.95	0.41

Bold Best result for given C_2 *Italic* Best result for given method **Bold** Best result

some information has been lost. The best results obtained with LTSA 1, PAL and PCA were unable to improve upon the best HH only results. The second channel of all NLDR techniques perform poorly, which is expected since they carry the least information.

Figure 4.1 shows the effect of increasing C_1 on the segmentation result. As C_1 gets larger, more and more merging occurs due to β being larger and the segmentation result becomes less noisy and smoother. The increased merging causes the number of regions remaining in the MIRGS RAG to decrease as C_1 increases. With larger C_1, the spatial context model is weighted more by a larger β and there is less reliance on the feature model. This explains why univariate feature sets such as HH in Table 4.2 have increasing performance at larger C_1: the feature space model has poor separability for the univariate feature sets, so increased reliance on spatial context is necessary to obtain accurate performance. At the same time, the spatial context is insufficient and so the univariate feature sets do not provide as much accuracy as the multivariate feature sets.

Table 4.3 shows segmentation accuracy statistics for the real image for each of the data fusion methods at four different values of C_2. The C_1 for each method was
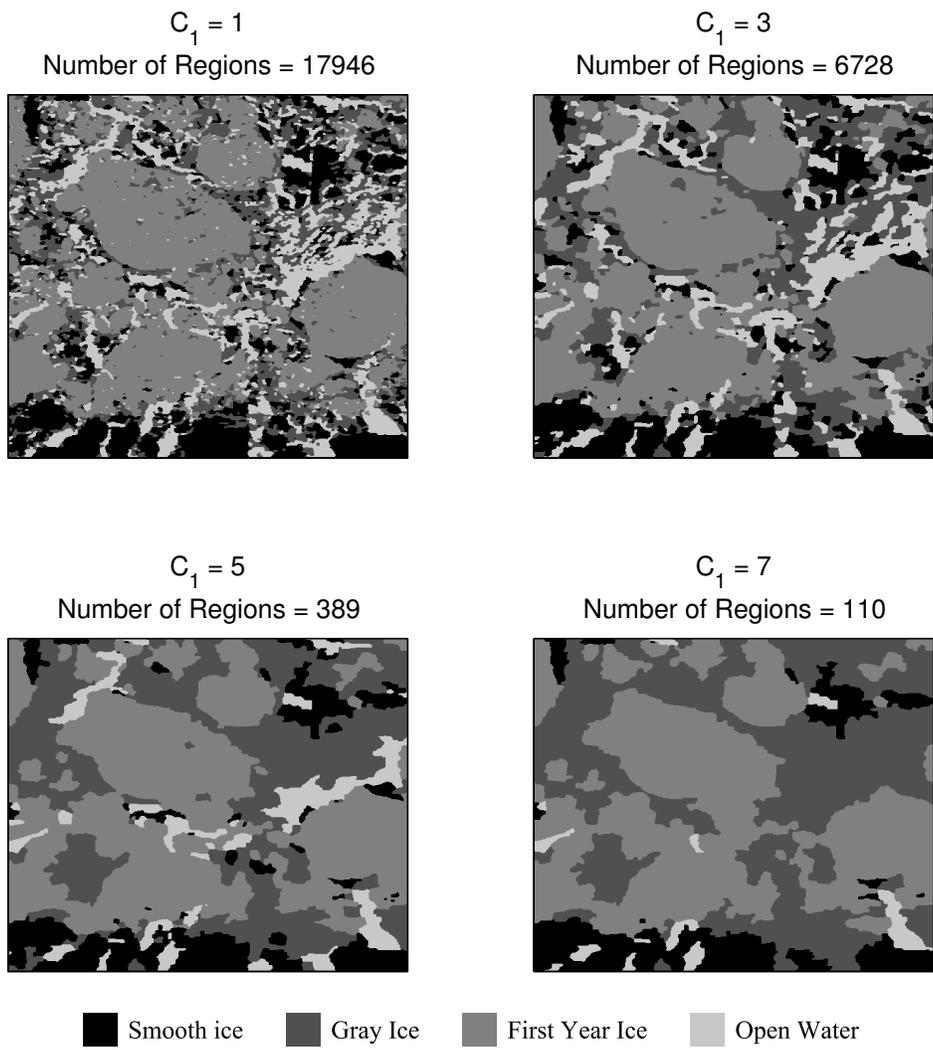
Figure 4.1: Segmentation results obtained for the real image (Figure B.1) using HH & HV with VFG gradient (the best feature set according to Table 4.2) at different values of C_1. The effect of larger C_1 is to increase the amount of merging. C_2 was fixed at 0.4.

C_2 = 0.1
Number of Regions = 5786

C_2 = 0.4
Number of Regions = 6728

C_2 = 0.7
Number of Regions = 7767

C_2 = 1
Number of Regions = 8646

■ Smooth ice ■ Gray Ice ■ First Year Ice ▢ Open Water
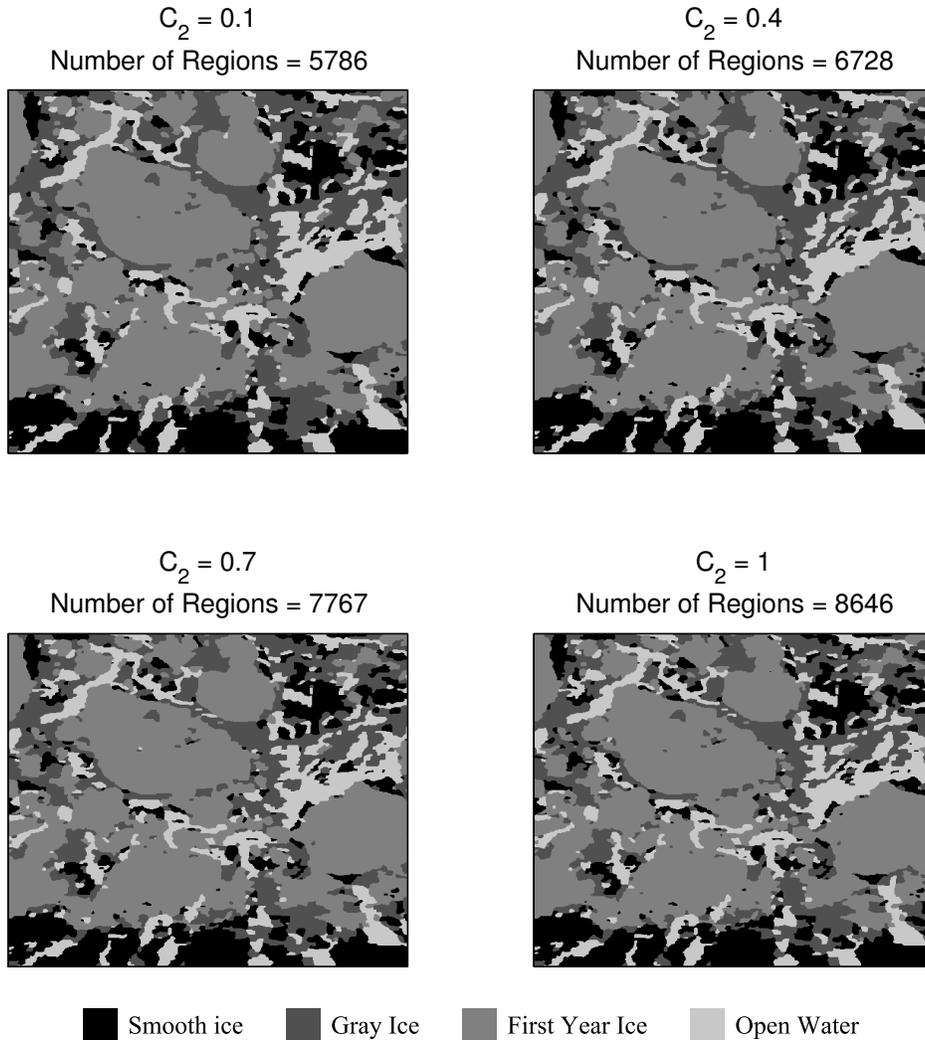
Figure 4.2: Segmentation results obtained for the real image using HH & HV with VFG gradient (the best feature set according to Table 4.3) at different values of C_2. The effect of larger C_2 is to decrease the amount of merging. C_1 was fixed at 3.

fixed at the best value obtained in Table 4.2. The effect of changing the value of C_2 within this range is much smaller than the effect of C_1. C_2 values beyond this range did not improve performance. None of the data fusion methods change in their relative performance; HH & HV with VFG gradient is still the best performing feature set. However, $C_2 = 1$ has improved accuracy over $C_2 = 0.4$. Figure 4.2 shows the effect of varying C_2 on the segmentation result. Larger C_2 leads to less merging (as indicated by the increasing number of regions remaining in the RAG) but this is not apparent visually. The best result remains nearly identical to the result obtained with default MIRGS parameters on HH & HV, VFG gradient, previously shown in Figure 2.6(f).
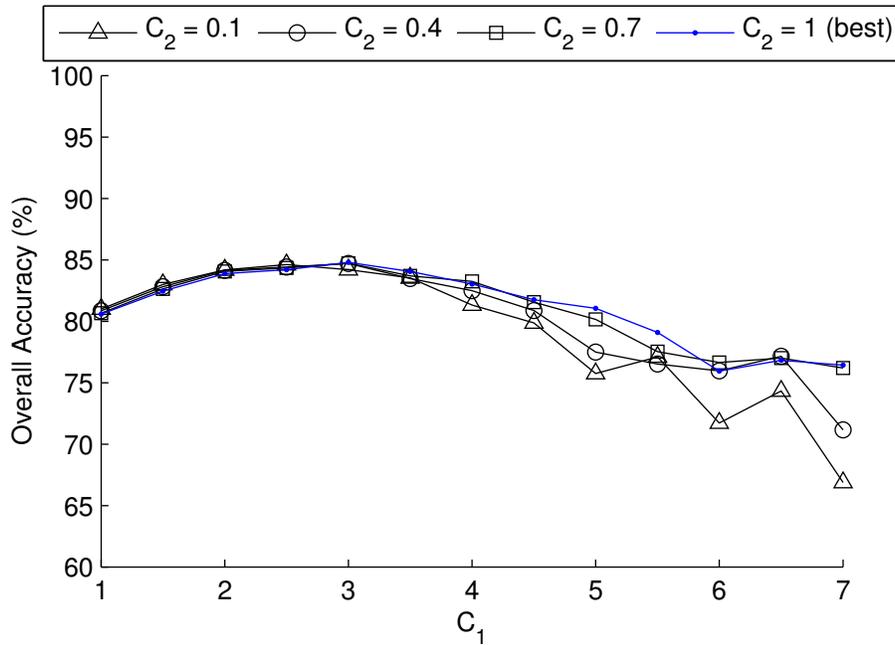
(a) Accuracy as a function of C_1 for various fixed C_2
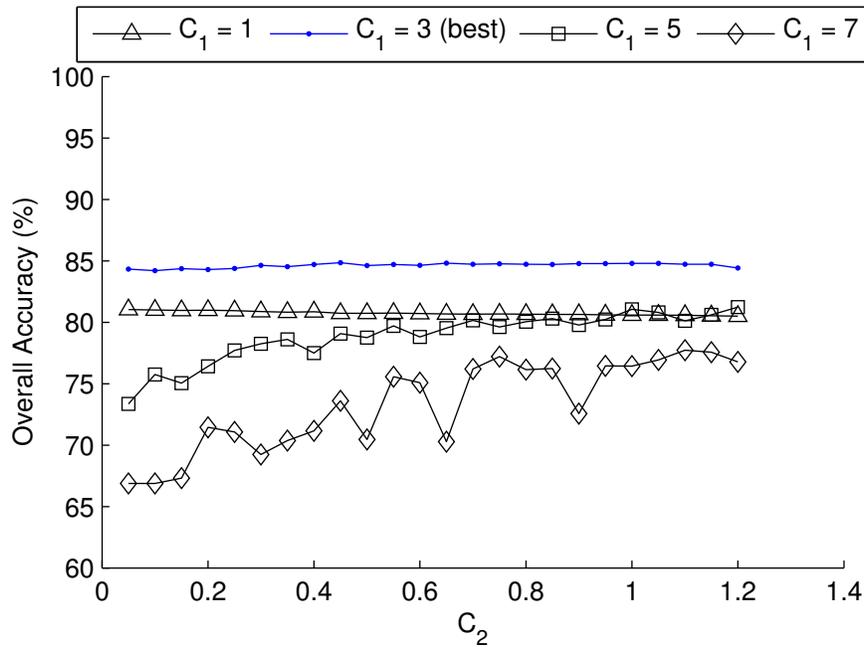
(b) Accuracy as a function of C_2 for various fixed C_1

Figure 4.3: Overall segmentation accuracy for the real image (Figure B.1) using HH & HV with VFG gradient as functions of (a) C_1 and (b) C_2. Accuracy is highest when $C_1 = 3$ and $C_2 = 0.45$. The accuracy curve is smooth when C_1 is small because the segmentation results are stable across different values of C_2. When C_1 is large, the curve is less smooth because the large size of the remaining unmerged regions causes perturbations in the stochastic segmentation to have a large effect.

Figure 4.3(a) shows the overall accuracy of the real image HH & HV, VFG gradient segmentation as a function of C_1, from $[1, 7]$ in increments of 0.5. The four curves correspond to four different values of C_2 for which the tests were performed. The blue curve indicates the C_2 that provided the highest accuracy. Accuracy is at a maximum at $C_1 = 3$, which agrees with Table 4.3. C_1 values below and above 3 have lower accuracy, corresponding to noisy and overmerged segmentations, respectively. The accuracy curve is smooth for $C_1 < 4$, which indicates that the MIRGS algorithm is stable over different ranges of C_1, with no irregularities in the segmentation optimization process. For larger C_1 values, the accuracy curve is less stable. This is expected from the MIRGS algorithm, which assigns labels to entire regions at a time: as the region size increases due to large C_1, any error in the label of a particular region has a larger effect on overall accuracy since each region comprises a larger number of pixels. Because the algorithm is partly stochastic due to Gibbs sampling [21], random label changes may also occur on a regional basis in different runs of the algorithm.

Figure 4.3(b) shows the overall accuracy of the real image (HH & HV, VFG gradient) segmentation as a function of C_2, from $[0.1, 1.2]$ in increments of 0.05. As with Figure 4.3(a), the four curves correspond to four different values of C_1 for which the tests were performed and the blue curve indicates the C_1 that provided the highest accuracy. Here, $C_1 = 3$ is the best parameter value for all values of C_2. Additionally, C_2 can be set to any value from $[0.1, 1.2]$ without affecting accuracy. As before, higher C_1 leads to accuracy curves that are less smooth due to large region sizes. Increasing C_2 can partially compensate for the increased merging caused by large C_1, which explains the rise in accuracy for the $C_1 = 7$ curve.

The experiments with the real image show that the preferred and reasonable method to use for the dual-polarization data is the HH & HV feature set with VFG gradient with $C_1 = 3$ and $C_2 = 1$.

4.1.2 Synthetic image

Table 4.4 shows the accuracy assessment results for the synthetic image (Figure B.2) for all data fusion methods at four different values of the C_1 parameter, with $C_2 = 0.4$. Similar to the results for the real image shown in Table 4.2, all four of the results obtained using the HH & HV data set have similar accuracy, with the MAX gradient performing best amongst the four. The closeness of the accuracy

Table 4.4: Accuracy statistics for tested data fusion techniques on the synthetic image (Figure B.2) at different values for the MIRGS merging parameter C_1. C_2 for each method was fixed at 0.4.

Method	$C_1 = 1$		$C_1 = 3$		$C_1 = 5$		$C_1 = 7$	
	Overall (%)	κ	Overall (%)	κ	Overall (%)	κ	Overall (%)	κ
ADWA Wavelet	50.94	0.32	56.92	0.39	*57.12*	*0.38*	54.21	0.34
HH	50.09	0.33	48.76	0.27	50.11	0.29	*78.87*	*0.70*
HH & HV, ADWA	91.38	0.88	*98.08*	*0.97*	97.07	0.96	95.86	0.94
HH & HV, VFG	**92.01**	**0.89**	*98.25*	*0.98*	97.15	0.96	96.25	0.95
HH & HV, MAX	90.95	0.88	*98.28*	*0.98*	97.24	0.96	**96.31**	**0.95**
HH & HV, CG	91.09	0.88	*98.09*	*0.97*	**97.28**	**0.96**	**96.31**	**0.95**
HV	54.97	0.41	79.99	0.72	*83.14*	*0.76*	81.79	0.74
LEIGS 1	67.26	0.56	77.39	0.70	*95.87*	*0.94*	95.27	0.93
LEIGS 1 & 2	74.18	0.65	80.67	0.74	*95.78*	*0.94*	77.20	0.68
LEIGS 2	65.45	0.55	67.99	0.55	81.75	0.74	*82.96*	*0.76*
LLE 1	77.27	0.69	90.62	0.87	*97.32*	*0.96*	96.66	0.95
LLE 1 & 2	91.22	0.88	*98.21*	*0.98*	97.60	0.97	96.42	0.95
LLE 2	62.94	0.50	89.55	0.86	*94.12*	*0.92*	93.08	0.90
LTSA 1	68.91	0.57	*93.91*	*0.92*	93.87	0.91	93.16	0.91
LTSA 1 & 2	92.13	0.89	**98.40**	**0.98**	97.23	0.96	96.15	0.95
LTSA 2	51.74	0.36	78.11	0.70	*83.40*	*0.77*	82.92	0.76
PAL	67.80	0.56	81.02	0.74	86.02	0.80	*86.59*	*0.81*
PCA	69.25	0.58	82.66	0.76	*88.44*	*0.84*	88.14	0.83
WA Wavelet	53.25	0.35	58.53	0.41	*58.92*	*0.41*	57.96	0.39
HV / HH	68.34	0.57	93.46	0.91	93.46	0.91	*93.61*	*0.91*

Bold Best result for given C_1 *Italic* Best result for given method   **Bold** Best result

values again shows that MIRGS is not sensitive to the gradient combination method. Multivariate feature sets LLE 1 & 2 and LTSA 1 & 2 also have similarly high accuracies, with LTSA 1 & 2 achieving the highest accuracy obtained from all the methods. As with the Table 4.2, the best performance for multivariate feature sets is achieved when $C_1 = 3$.

The performance of the single-polarization channel HH is affected by the low separability of the ice types at the mid-range incidence angles that the synthetic image represents. In contrast to the results with the real image, the HV channel provides better accuracy. Neither of the single-polarization channels approach the best multivariate feature sets in terms of accuracy. The single channel images created by using NLDR techniques for fusion (LEIGS 1, LLE 1, LTSA 1), PCA, PAL and HV / HH all provide better accuracy than the HH or HV channels. The NLDR techniques approach the multivariate level of accuracy. As with the real image, PCA outperforms PAL. Less successful with this image are the wavelet fusion methods, both of which have poor performance. ADWA wavelet fusion in particular was designed to take advantage of the higher backscatter level of open water compared to first year ice in the HH channel to increase image contrast. At

64
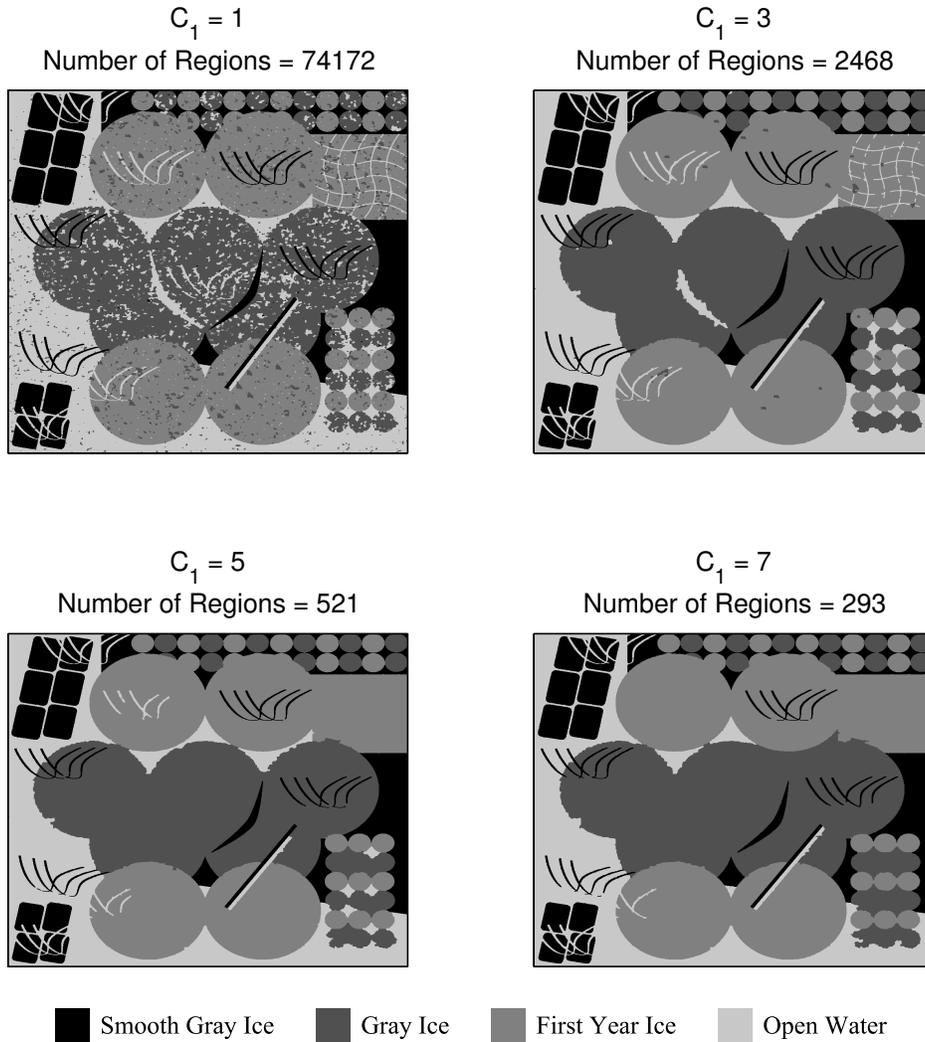
Figure 4.4: Segmentation results obtained for the synthetic image (Figure B.2) using LTSA 1 & 2 (the best feature set according to Table 4.4) at different values of C_1. The effect of larger C_1 is to increase the amount of merging. C_2 was fixed at 0.4.

mid-range incidence angles, the backscatter of open water is closer to first year ice so the rule does not perform as well.

Figure 4.4 shows the LTSA 1 & 2 segmentation results for the synthetic image at different values of C_1. As before, increasing C_1 causes increased merging. Once $C_1 = 3$, the noise in the segmentation is gone. Higher levels of C_1 causes detail to be lost.

Table 4.5 shows the accuracy assessment results for the synthetic image (Figure B.2) for all data fusion methods at four different values of the C_2 parameter, with C_1 fixed at the best values obtained in Table 4.4. The table shows that $C_2 = 0.4$ is the best value for most of the multivariate feature sets. Figure 4.5 shows the

Table 4.5: Accuracy statistics for tested data fusion techniques on the synthetic image (Figure B.2) at different values for the MIRGS merging parameter C_2. C_1 for each method was fixed at the best value for the method as indicated by Table 4.4.

Method	$C_2 = 0.1$		$C_2 = 0.4$		$C_2 = 0.7$		$C_2 = 1$	
	Overall (%)	κ	Overall (%)	κ	Overall (%)	κ	Overall (%)	κ
ADWA Wavelet	53.75	0.33	57.12	0.38	*58.87*	*0.41*	58.18	0.40
HH	53.81	0.34	*78.87*	*0.70*	53.62	0.33	52.84	0.35
HH & HV, ADWA	*98.12*	*0.97*	98.08	0.97	98.08	0.97	97.96	0.97
HH & HV, VFG	98.15	0.97	*98.25*	*0.98*	98.05	0.97	97.85	0.97
HH & HV, MAX	98.24	0.98	*98.28*	*0.98*	98.08	0.97	97.61	0.97
HH & HV, CG	98.08	0.97	*98.09*	*0.97*	98.03	0.97	97.63	0.97
HV	82.76	0.76	83.14	0.76	83.18	0.77	*83.35*	*0.77*
LEIGS 1	95.64	0.94	*95.87*	*0.94*	93.56	0.91	85.36	0.80
LEIGS 1 & 2	*96.17*	*0.95*	95.78	0.94	95.14	0.93	94.12	0.92
LEIGS 2	*95.44*	*0.94*	82.96	0.76	82.87	0.75	82.42	0.75
LLE 1	97.21	0.96	*97.32*	*0.96*	97.17	0.96	96.57	0.95
LLE 1 & 2	*98.29*	*0.98*	98.21	0.98	98.13	0.97	97.97	0.97
LLE 2	93.89	0.92	*94.12*	*0.92*	*94.12*	*0.92*	93.00	0.90
LTSA 1	*94.37*	*0.92*	93.91	0.92	92.41	0.90	90.65	0.87
LTSA 1 & 2	98.38	0.98	**98.40**	**0.98**	98.35	0.98	98.21	0.98
LTSA 2	82.83	0.76	*83.40*	*0.77*	83.03	0.76	83.35	0.77
PAL	86.69	0.81	86.59	0.81	*87.10*	*0.82*	86.69	0.81
PCA	*89.21*	*0.85*	88.44	0.84	85.70	0.80	84.99	0.79
WA Wavelet	58.15	0.40	58.92	0.41	*59.98*	*0.43*	59.60	0.42
HV / HH	93.01	0.90	93.61	0.91	93.05	0.90	*94.38*	*0.92*

Bold Best result for given C_2 *Italic* Best result for given method   **Bold** Best result

effect of varying C_2 on the segmentation result. Larger C_2 leads to less merging (as indicated by the increasing number of regions remaining in the RAG) but, as with the real image case, this is not apparent visually.

Figure 4.6(a)-(b) show the accuracy of the synthetic image segmentation (using LTSA 1 & 2) results as functions of C_1 and C_2. The accuracy is highest at $C_1 = 3$, with C_2 not having a large effect unless C_1 is larger. This is similar to the real image case. Unlike the real image case, all the accuracy curves are smooth regardless of the value of C_1. The synthetic image is easier to segment than the real image due to the clean outlines so regions are not as likely to be labeled incorrectly by the stochastic optimization process. This means that even with larger regions, accuracy does not fluctuate too much. Figure 4.6(b) shows that when $C_1 = 5$ or $C_1 = 7$, increasing C_2 makes the accuracy approach that of the $C_1 = 3$ case. This is a case of C_2 being able to compensate for the extra merging caused by large C_1.
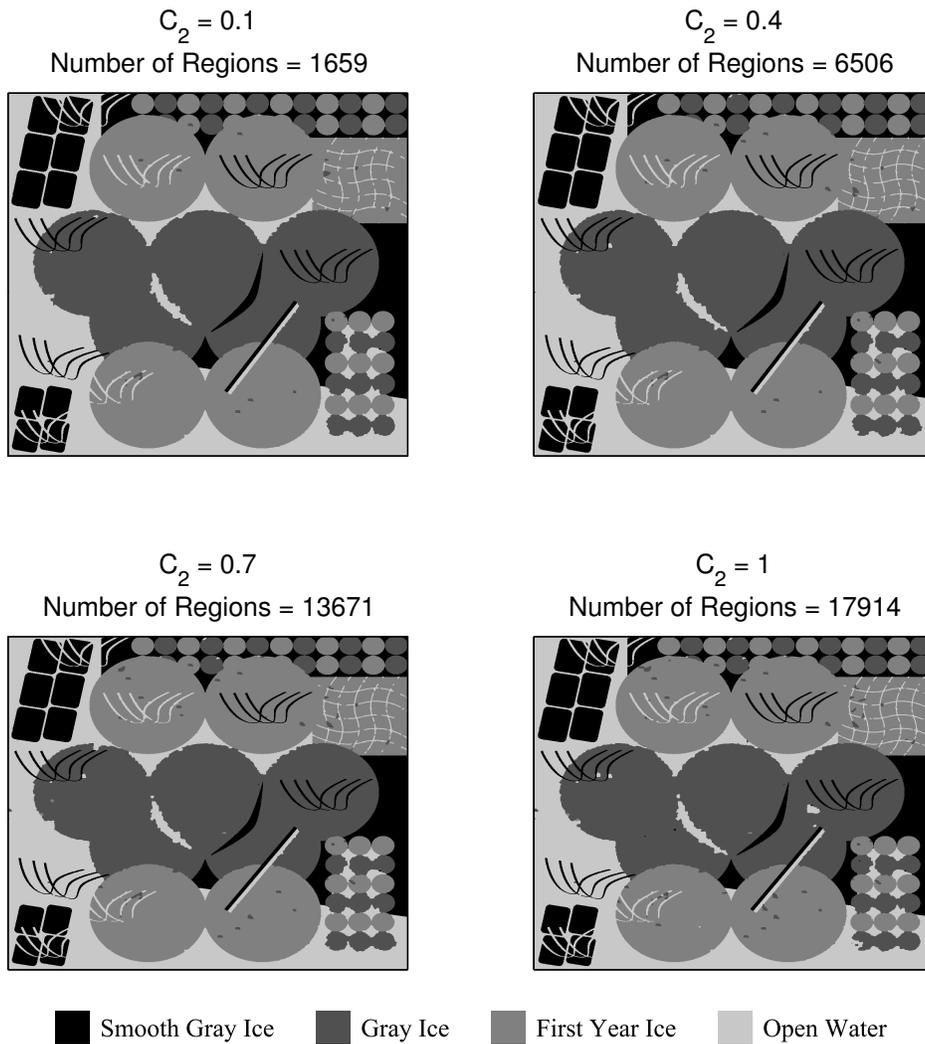
Figure 4.5: Segmentation results obtained for the synthetic image (Figure B.2) using LTSA 1 & 2 (the best feature set according to Table 4.5) at different values of C_2. The effect of larger C_2 is to decrease the amount of merging. C_2 was fixed at 3.
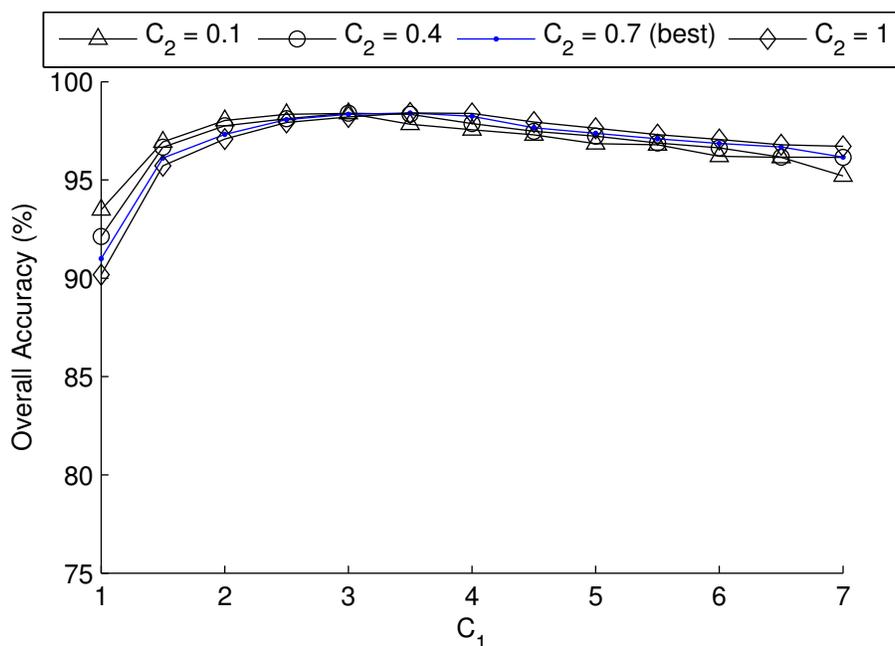
(a) Accuracy as a function of C_1 for various fixed C_2
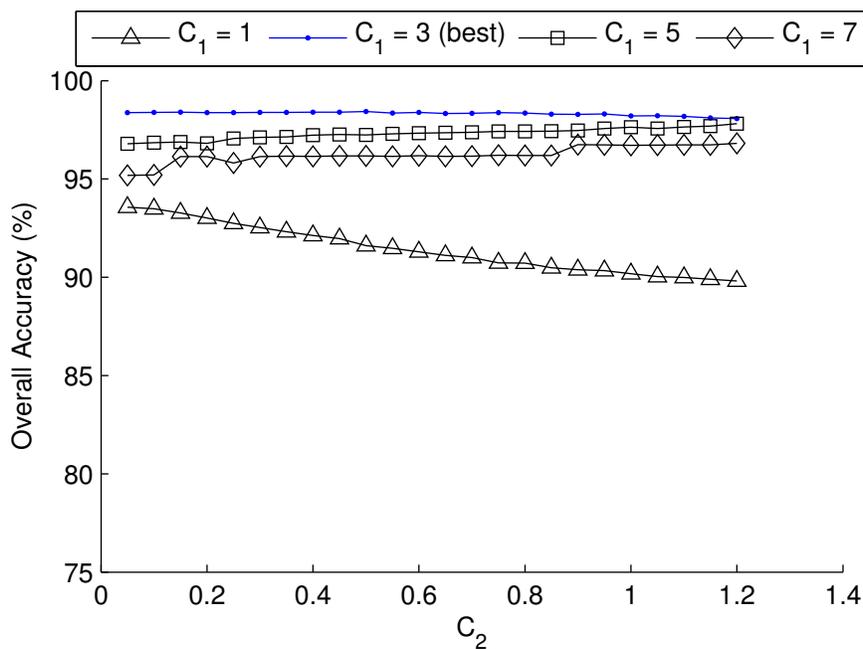
(b) Accuracy as a function of C_2 for various fixed C_1

Figure 4.6: Overall segmentation accuracy for the synthetic image (Figure B.2) using LTSA 1 & 2 as functions of (a) C_1 and (b) C_2. Accuracy is highest when $C_1 = 3.5$ and $C_2 = 0.7$. The accuracy curve is smooth because the segmentation results are stable across different values of C_1 and C_2. The effect of C_2 is much smaller than that of C_1.

4.1.3 Discussion

Considering the results from both the real image (Figure B.1) and the synthetic image (Figure B.2), HH & HV with any of the gradient combination rules and LLE 1 & 2 consistently provide the best or near the best performance. The best accuracy is achieved when C_1 is close to 3 to 3.5. When C_1 is set correctly, C_2 can be set to nearly any value within $[0.1, 1]$ without affecting accuracy. The multivariate output from the NLDR transforms do not alter the feature space separability as compared to HH & HV for either image.

The most consistently performing NLDR method is LLE. LLE 1 consistently achieves the highest univariate accuracy for both synthetic and real images and LLE 1 & 2 performs very closely to HH & HV. LTSA only performs well for the synthetic image and LEIGS does not perform particularly well for either image. The difference in LTSA's performance between the two images appears to be related to the image resampling process that determines the training feature vectors. This is made clear in Figure 4.7. The figure shows LTSA 1 for the whole image under different sets of training feature vectors obtained from different image resampling schemes. In the top row, a pseudorandom resampling scheme was used where pixels were selected to be used as training feature vectors from a regular grid. The grid was offset by a random amount for each example shown in the top row so that the training vectors were different each time. It is clear that the fused image result from LTSA is different for every set of training vectors. The bottom row shows similar results, except the training vectors were obtained by bicubic sampling the original image to have the indicated number of pixels. Figure 4.8 and Figure 4.9 show the same experiment for LEIGS and LLE, respectively. The results for LEIGS and LLE are much more consistent over different sets of training vectors, with the bicubic sampling scheme showing greatest consistency. These results show that LTSA is not robust to different sets of training vectors. Based on the above considerations, the LLE method appears to be the best NLDR technique for the purposes of fusing R2 data.

The wavelet techniques, PAL, PCA and HV / HH do not perform particularly well for either real or synthetic images and should not be considered as an image fusion method for input into MIRGS.
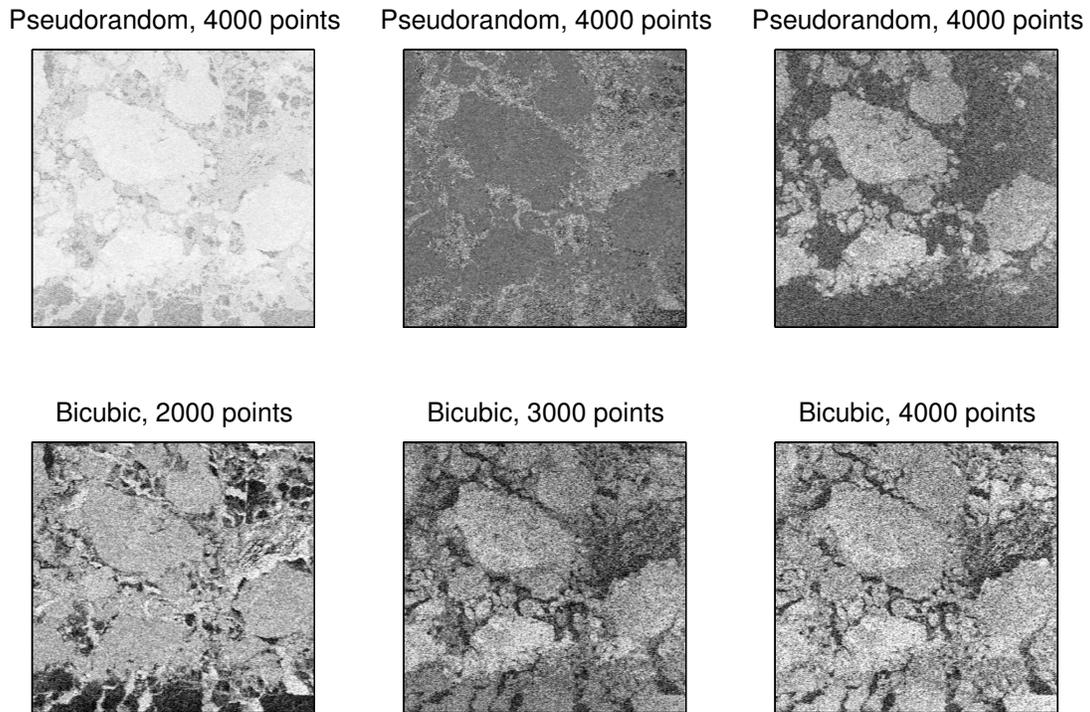
Figure 4.7: LTSA is not robust to picking different subsets of samples for use during the training phase since each sample subset produces a different fused image result.
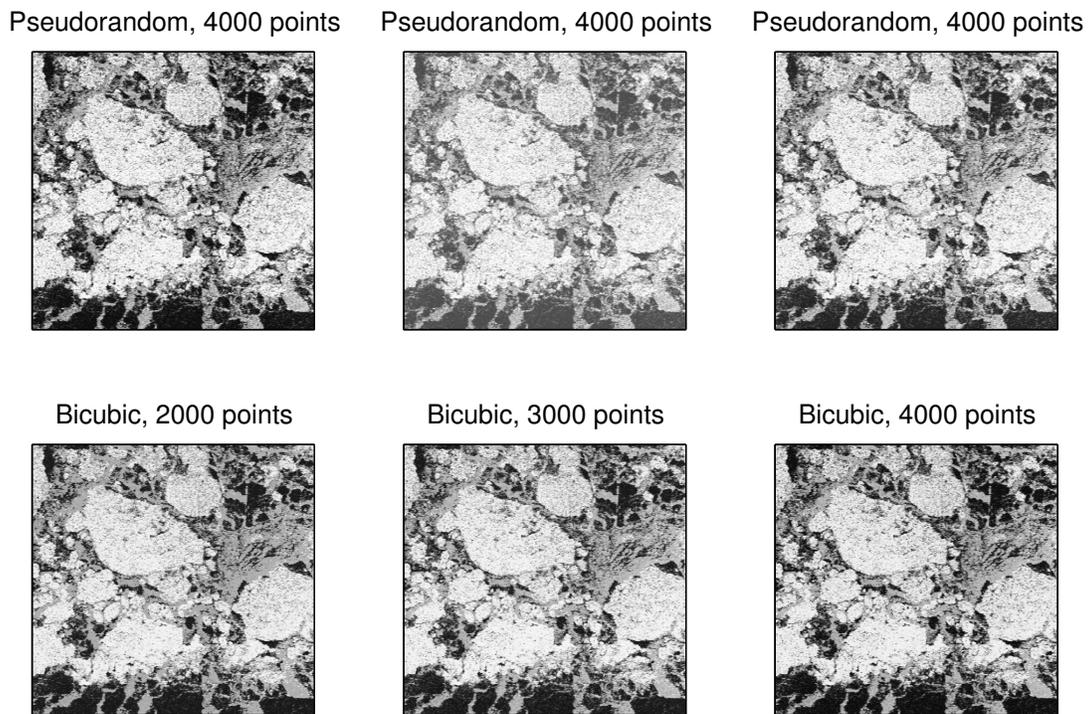

Figure 4.8: LEIGS produces images that are similar to each other under different training sample selection techniques.
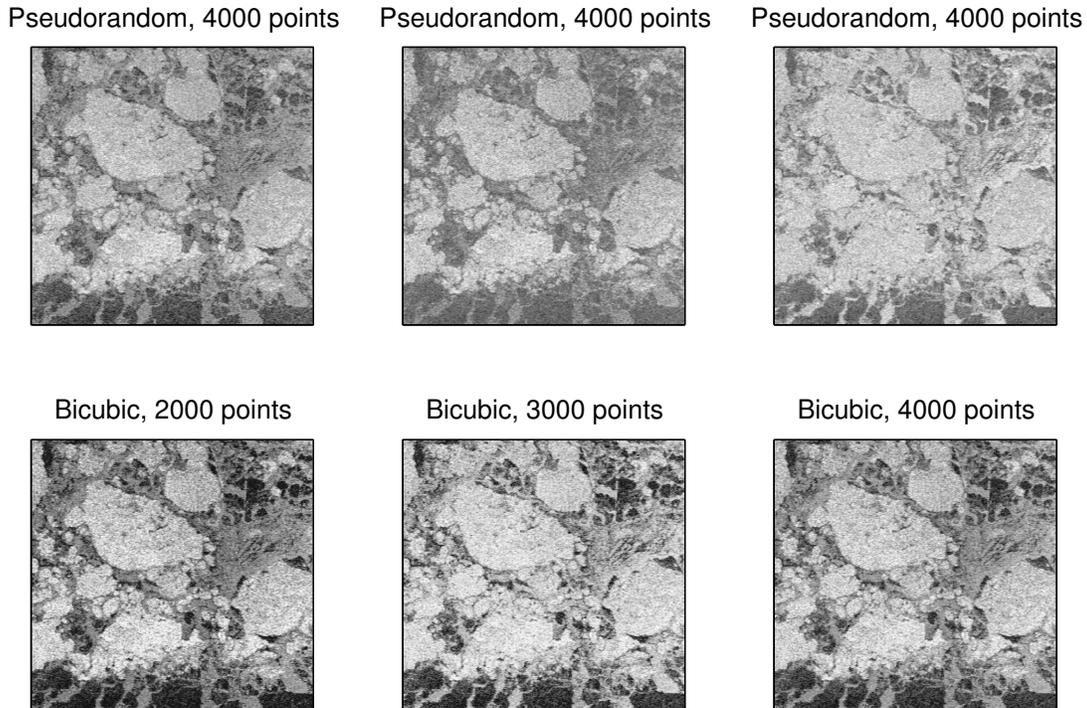
Figure 4.9: LLE produces images that are similar to each other under different training sample selection techniques.

4.2 Hierarchical segmentation

The hierarchical segmentation results are presented in this section. All experiments follow the same procedure. First, the SITH is defined, either manually or with Auto-SITH. The HV channel is then segmented using MIRGS into the number of classes specified by the hierarchy, with $C_1 = 5$ and $C_2 = 0.4$ (set at the best values obtained previously). Each segment from the HV channel is then segmented using the HH channel using the number of classes specified by the hierarchy, with $C_1 = 3$ and $C_2 = 0.4$. C_1 is set to 3 in order to retain the best level of detail in the segmentation. This is different from the results obtained previously in Tables 4.2 and Tables 4.4 where $C_1 = 7$ gave the best results for HH because the feature space separability. For hierarchical segmentation, such a high C_1 is not necessary because the HV channel has already been used to segment out classes that are difficult to distinguish in HH. Whenever excess classes must be merged, calculation of the Fisher criterion was tested for three cases: with the HH only (HH merge), HV only (HV merge) or HH & HV (HH & HV merge) feature sets.

In addition to performing the experiments on the full scene image, the experiments were also performed on both the real and synthetic images to determine

Table 4.6: Accuracy statistics for tested sea ice type hierarchies.

(a) Real Image (Figure B.1)

Hierarchy	Overall (%)	κ
User-Defined SITH	**82.6**	**0.75**
SITH 2-3, HH Merge	69.77	0.59
SITH 2-3, HV Merge	70.64	0.59
SITH 2-3, HH & HV Merge	78.69	0.70
SITH 3-2, HH Merge	69.08	0.57
SITH 3-2, HV Merge	57.61	0.40
SITH 3-2, HH & HV Merge	71.05	0.60
Auto-SITH	Same as SITH 3-2	

(b) Synthetic Image (Figure B.2)

| Hierarchy	Overall (%)	κ
User-Defined SITH	97.08	0.96
SITH 2-3, HH Merge	76.02	0.67
SITH 2-3, HV Merge	76.02	0.67
SITH 2-3, HH & HV Merge	76.02	0.67
SITH 3-2, HH Merge	66.67	54
SITH 3-2, HV Merge	**97.67**	**0.97**
SITH 3-2, HH & HV Merge	**97.67**	**0.97**
Auto-SITH	Same as SITH 3-2	

Bold Best result

whether hierarchical segmentation is a technique to make use of dual-polarization data even for images without the incidence angle effect.

4.2.1   Real and synthetic images

The user defined SITHs for the real image (Figure B.1) and the synthetic image (Figure B.2) are shown in Figure 4.10. These are generated by visually inspecting the images and determining a reasonable number of classes for each channel. A number of different hierarchies were defined and tested. Only the hierarchies with the best results in these initial tests are presented here as the user defined hierarchies.

Table 4.6 shows accuracy statistics for the user defined SITH and the generic SITH 2-3 and SITH 3-2 hierarchies. For the generic SITHs, the accuracy for the three cases of excess class merging (HH merge, HV merge and HH & HV merge) are shown. In the case of both images, Auto-SITH generated a hierarchy identical to SITH 3-2, so the SITH 3-2 results apply. The best results are obtained with the user defined SITHs, although for the synthetic image, SITH 3-2 generated a slightly better result. This indicates that Class 3 or Class 4 in Figure 4.10(b) contains small areas of either Class 1 or Class 2 which are properly segmented by SITH 3-2 and then merged with Class 1 or Class 2 in the post-processing. It should be noted that neither Auto-SITH or the generic SITH works consistently for both images. The DB index used by Auto-SITH likely generated an incorrect number of clusters for the HV image: the DB index indicated three clusters while the HV image can only distinguish two clusters. This is supported by visually inspecting the image and by observing that the user-defined hierarchy obtained the best results. However, the DB index is correct for the synthetic image. The best feature set to use for post-process merging was also not consistent between the two images or even
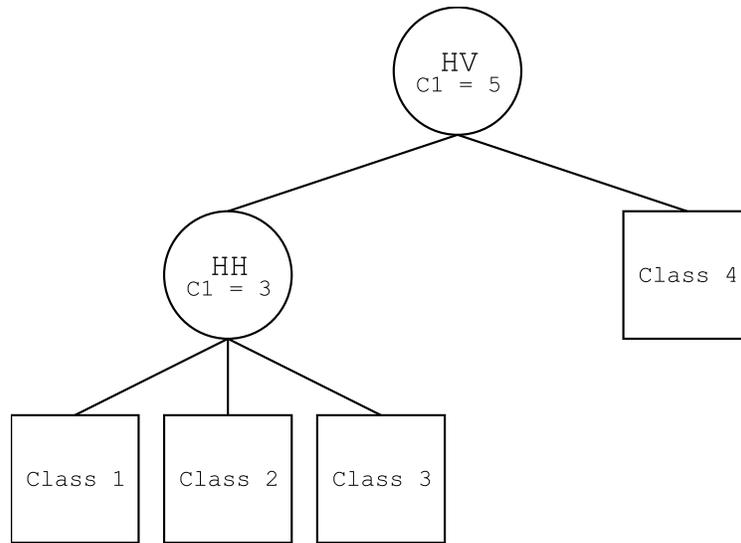
(a) Real image
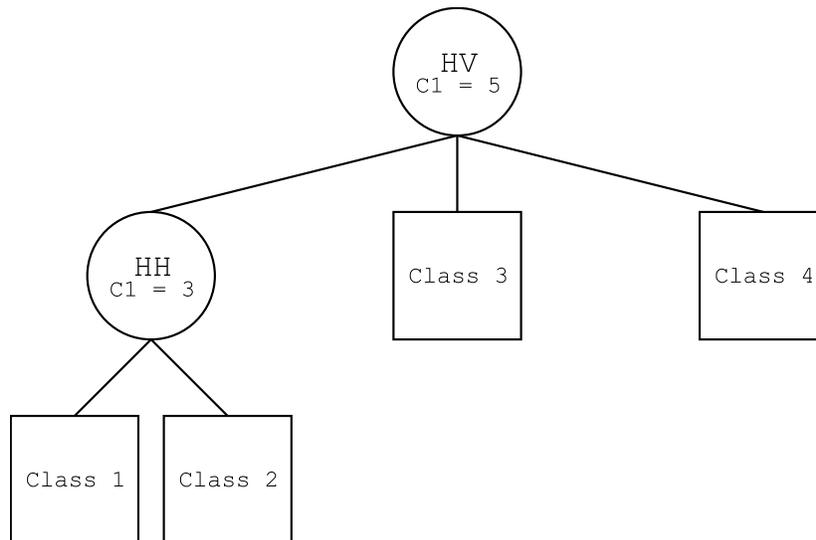
(b) Synthetic image

Figure 4.10: User defined segmentation hierarchies for the real (Figure B.1) and synthetic (Figure B.2) images.

(a) HH & HV, VFG   (b) User-defined SITH   (c) Ground-truth

■ Smooth ice   ■ Gray Ice   ■ First Year Ice   ■ Open Water
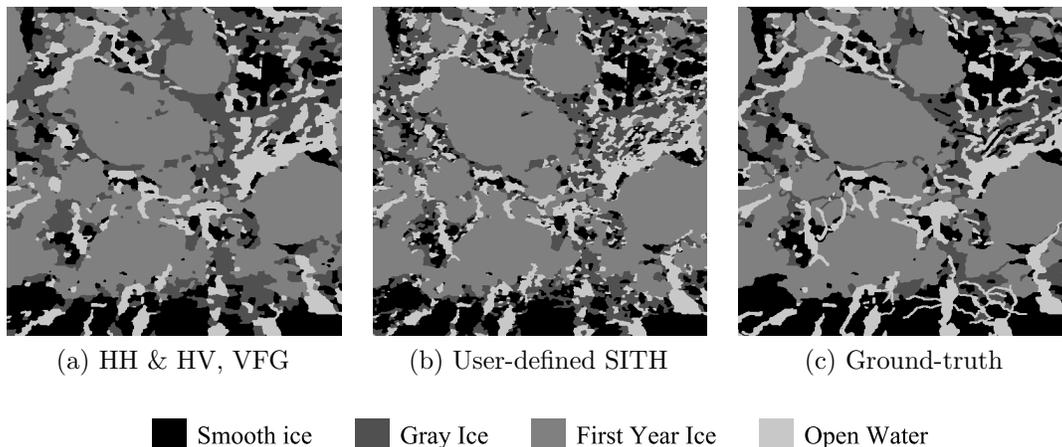
Figure 4.11: Comparison between the best regular segmentation results and the best SITH results for the real image (Figure B.1).

between the two generic SITHs that were tried for each image.

The best accuracies obtained using the SITH methods are slightly lower than the best accuracies obtained from the best segmentations obtained in the Section 4.1. Figures 4.11 and 4.12 show the best regular segmentations compared to the best hierarchical segmentations for the real and synthetic images, respectively. The SITH results for the real image are slightly noisier than the HH & HV segmentation. The SITH results for the synthetic image have lost some detail compared to the LTSA 1 & 2 segmentation. These experiments have shown that hierarchical segmentation is slightly worse than regular segmentation when the incidence angle effect is not present.

4.2.2 Full scene image

Figure 4.13 shows the user-defined and Auto-SITH hierarchies for the full scene image (Figure B.3). Both of these hierarchies contain an extra class that needs to be merged during post-processing. Class 1 and Class 2 in the user-defined hierarchy both correspond to open water. Parts of the open water are grouped with Class 3 (gray ice) and Class 4 (first year ice) during the initial HV segmentation, so that the entire segment must be separated into three classes to recover the open water. The segments that correspond to Class 1 and Class 2 must then be merged. This is an example where merging based on a criterion that includes spatial context may not work: Class 1 and Class 2 do not actually touch and yet they are the classes

(a) LTSA 1 & 2   (b) SITH 3-2, HH & HV   (c) Ground-truth
                            merge

■ Smooth Gray Ice   ■ Gray Ice   ■ First Year Ice   ■ Open Water
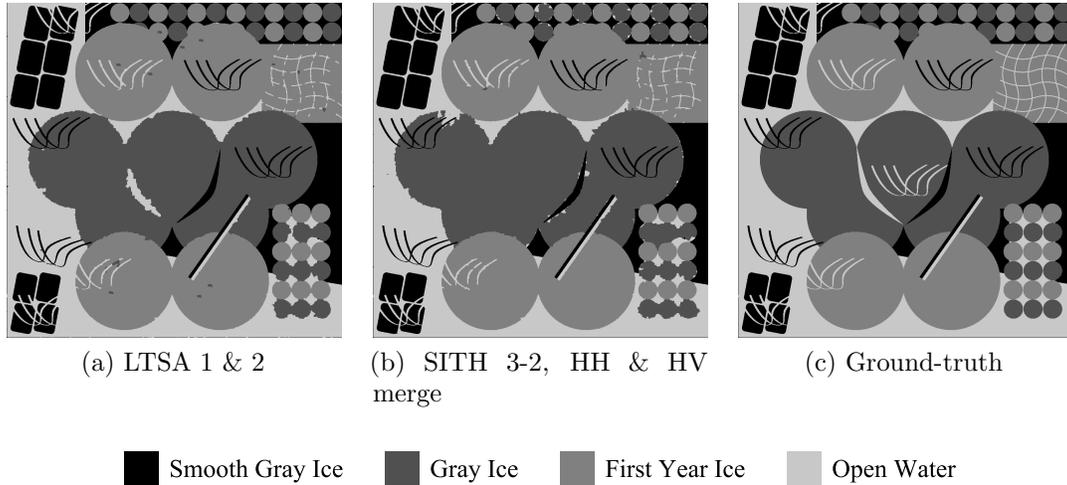
Figure 4.12: Comparison between the best regular segmentation results and the best SITH results for the synthetic image (Figure B.2).

that should be merged together.

Figure 4.14 shows the segmentation results for the user-defined SITH and the generic SITHs, with merging based on the three different feature sets. Only the user-defined and SITH 2-3 hierarchies produce the expected segmentation. The user-defined SITH can be merged using both HH & HV and HV only, while the SITH 2-3 only works with HV merging. In general, any merging that involves the HH band does not give the proper segmentation result if the open water portion from the initial HV segmentation was segmented into more than one class, as in SITH 2-3. This is because the open water in the HH band varies dramatically in appearance as a result of incidence angle. In fact, the only reason that the user-defined SITH produces an acceptable segmentation when merged using HH & HV is because the main open water portion was not segmented further. SITH 3-2 with HV merging does not have any of the incidence angle related segmentation problems, but much of the gray ice is confused with open water. First year ice can almost always be separated from open water when using hierarchical segmentation.

Figure 4.15 shows the Auto-SITH segmentation results. Like the previous cases, any merging performed using a feature set that involves the HH band produces a poor result. Merging with HV eliminates the incidence angle-induced segment but much of the open water embedded in the first year and gray ice regions is confused with gray ice. This is because there were only two classes in each segment produced by the initial HV segmentation. In this case, the DB index correctly indicated that
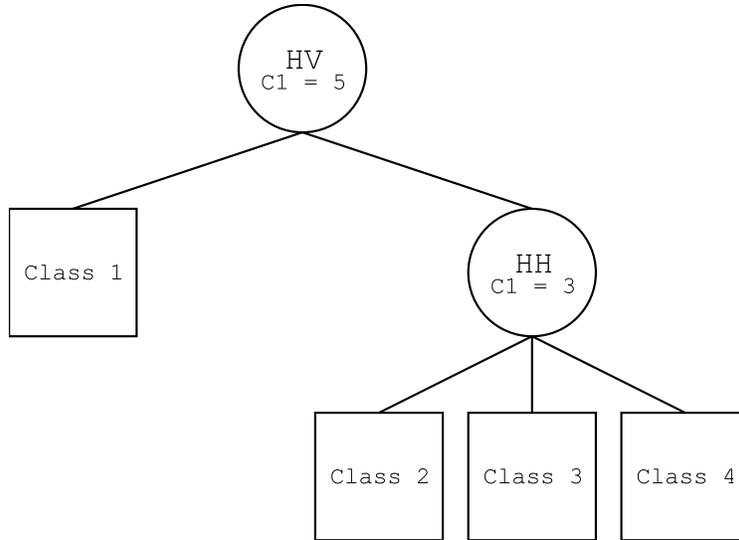
(a) User defined hierarchy
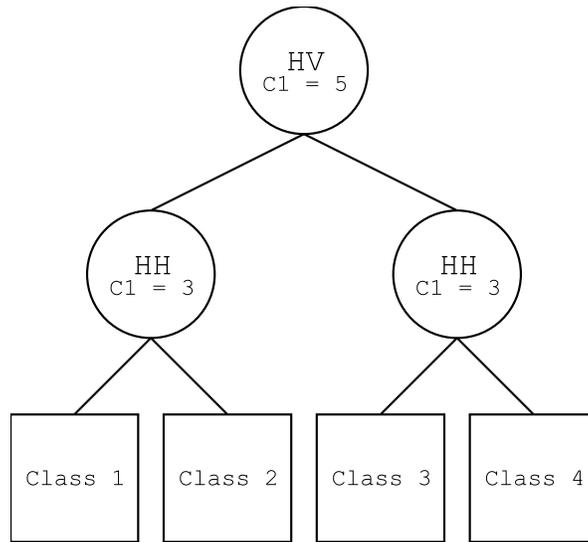
(b) Auto-SITH hierarchy

Figure 4.13: (a) The user-defined SITH for the full scene image (Figure B.3) contains an extra class because the initial HV segmentation does not fully separate open water from ice (class 1 and class 2 are both open water). The user-defined SITH therefore requires post-process merging. (a) The Auto-SITH hierarchy also contains an extra class that must be merged.

(a) User-defined SITH, HH merge

(b) User-defined SITH, HV merge

(c) User-defined SITH, HH & HV merge

(d) SITH 2-3, HH merge

(e) SITH 2-3, HV merge

(f) SITH 2-3, HH & HV merge

(g) SITH 3-2, HH merge

(h) SITH 3-2, HV merge

(i) SITH 3-2, HH & HV merge

(j) Expected Segmentation (derived from (b)

| ■ Land | ■ Gray Ice | ■ First Year Ice | ■ Open Water |
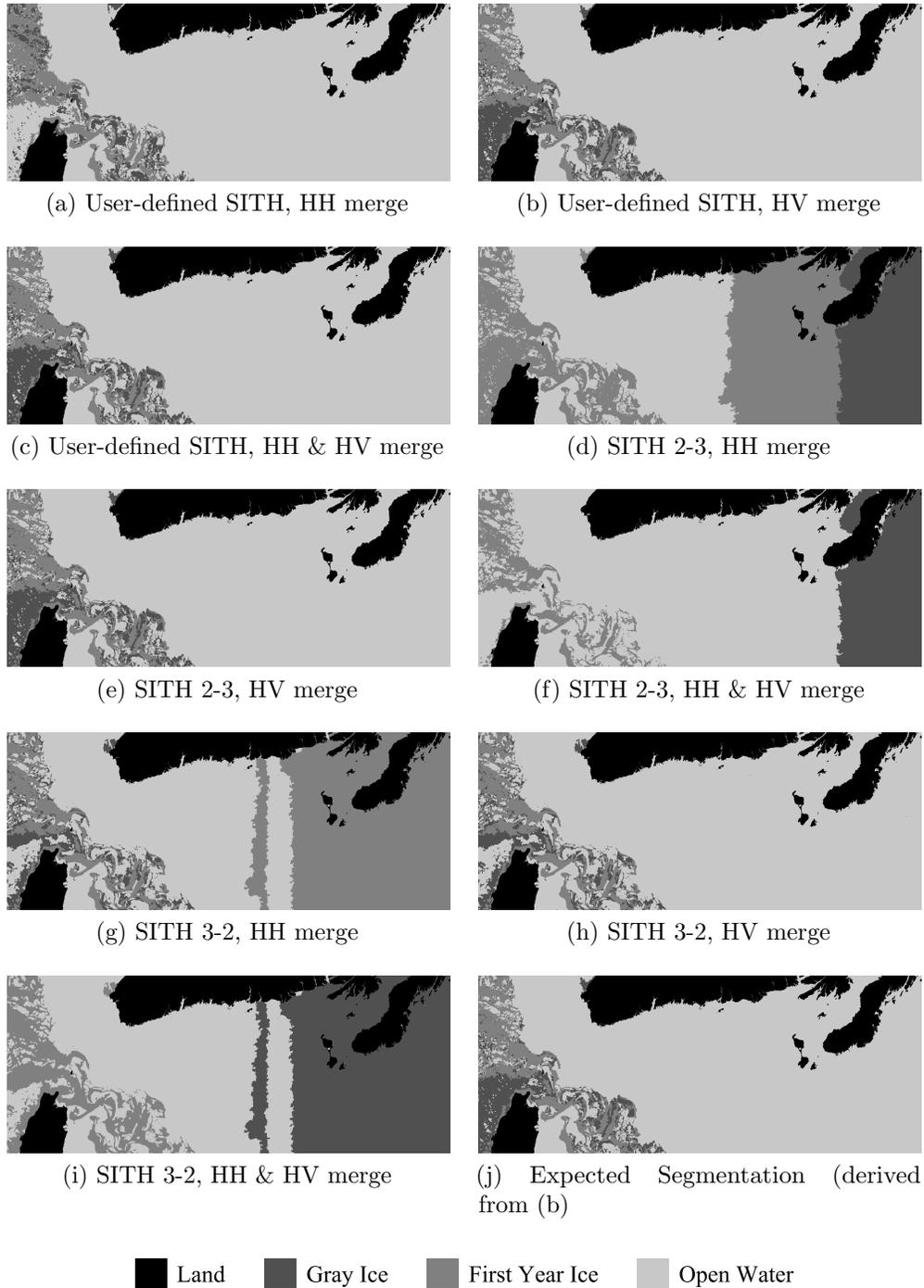
Figure 4.14: The incidence angle variation of open water can be handled by the selection of an appropriate sea ice typing hierarchy followed by merging of extra segments using the Fisher criterion in an appropriate feature set. In the full scene image (Figure B.3), the user-defined hierarchy followed by merging using (b) HV only or (c) HH & HV works well. SITH 2-3 with merging on (e) HV also works. (j) The expected answer for comparison is derived from the user-defined hierarchy result.

(a) Auto-SITH, HH merge   (b) Auto-SITH, HV merge

(c) Auto-SITH, HH & HV merge   (d) Expected Segmentation

■ Land   ■ Gray Ice   ■ First Year Ice   ■ Open Water
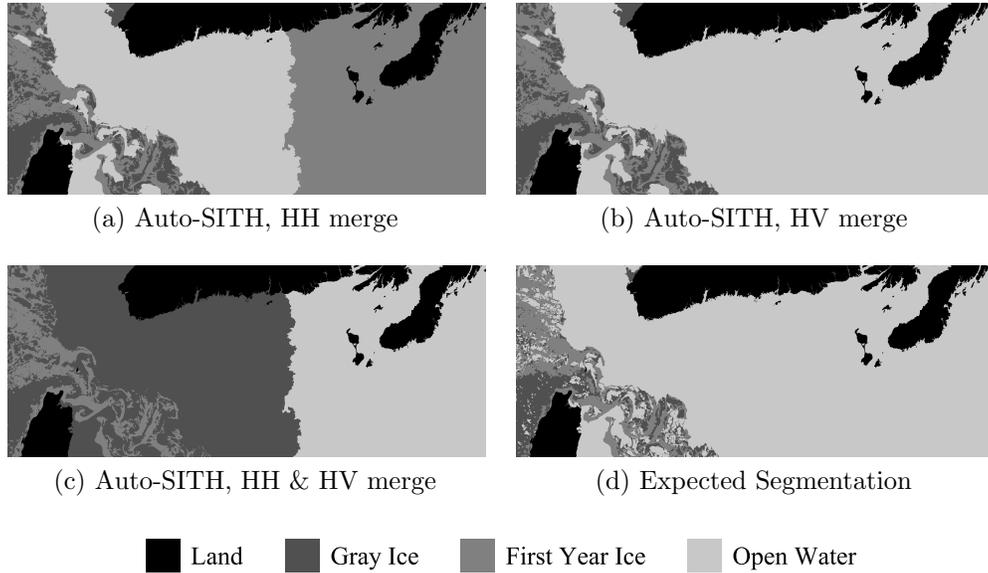
Figure 4.15: Only the HV merged Auto-SITH result comes close to the expected segmentation for the full scene image (Figure B.3), but it has lost most of the small areas of open water on the left hand side of the image.

the HV channel can be divided into two segments, but failed to correctly identify the number of classes in each of those two segments.

4.2.3   Discussion

The preceding results indicate that hierarchical segmentation is not the ideal method of dual-polarization segmentation when incidence angle effects are not present. The use of regular segmentation is thus motivated for the current MAGIC use cases, where limited range CIS polygons with no incidence-angle-based variations for interior pixels are segmented by MIRGS to a pixel resolution level. For full scene automatic segmentation, hierarchical segmentation appears to be a promising approach but only when the hierarchies are manually defined. Furthermore, if the hierarchies involve excess classes, which may be unavoidable in some cases, HV is the only feature set that can be used for merging. Although not tested here because there were no examples in the available R2 data sets, this merging scheme may be problematic if some of the segments are ice types (such as gray ice, water and new ice) in which the HV band has poor separability. In those cases, merging on HV may erroneously merge different classes together. It should be possible for an ice expert to manually define the merging in the same way the segmentation hierarchy was defined, either

by using a user interface or by setting up a set of merging steps. This will increase the workload required over a fully automated sea ice segmentation algorithm, but will still be less laborious than manual segmentation.

Some of the merging problems may be solvable if the classification process that was partially developed in [64] is used to label some of the segments. The classification process could use shape, texture and tone information to help identify the ice class of a particular segment, which will help the merging process.

Auto-SITH is not entirely reliable. The estimation method for the number of clusters seems to give values that are not what is expected for the given image. The DB index estimate of the number of classes is often in error by one class (e.g. it indicates 2 when the number is 3 or vice versa). Since SITH segmentation requires a precise number of classes to work properly, Auto-SITH with the DB index is not a recommended method for automatically generating the segmentation hierarchies.

Chapter 5

Conclusions

5.1 Summary

The use of dual-polarization RADARSAT-2 SAR sea ice imagery to improve sea ice segmentation has been investigated in this thesis. An initial investigation of using dual-polarization RADARSAT-2 in the MIRGS algorithm has shown that it can substantially improve the segmentation accuracy, particularly with regards to separating open water from ice at near range incidence angles. This has confirmed that automated algorithms can benefit from the improvements provided by dual-polarization data, just as other researchers have found that the dual-polarization data have been useful for human interpretation.

The evaluation of RADARSAT-2 data with the standard MIRGS algorithm revealed that some segmentation detail was still lost and that full scene segmentation was still not possible due to the appearance of open water changing across the range of incidence angles. To improve the accuracy of MIRGS, several data fusion schemes for the dual-polarization data were investigated. These included classic feature space fusion approaches such as principal components analysis (PCA), as well as non-linear dimensionality reduction (NLDR) techniques and gradient combination techniques that attempted to combine the edge information from the dual-polarization channels in an intelligent manner. Image fusion by means of a dual-polarization channel ratio (HV / HH) and wavelet methods were also investigated. MIRGS parameter selection was also investigated to determine the best set of parameters to use for dual-polarization data. A hierarchical segmentation scheme was also introduced to take advantage of the HV channel's insensitivity to incidence angle in order to

reduce errors caused by incidence angle related changes in sea ice appearance.

Experiments were performed on three images: a limited incidence angle range real RADARSAT-2 image with manual ground-truth (Figure B.1), a synthetic limited incidence angle range RADARSAT-2 image (Figure B.2) and a full scene RADARSAT-2 image with incidence angle effects (Figure B.3). According to the experiments, the best data fusion method for dual-polarization data is the standard MIRGS algorithm with no changes. Intelligent gradient combination rules had very little effect on overall accuracy, while feature space and image fusion approaches did not retain all the separability information when the two channels were combined into one channel. The best fusion technique, a NLDR technique called locally linear embedding (LLE), consistently produced the best single channel image segmentation results from the dual-polarization data but was still unable to match the segmentation performance of the dual-polarization feature set.

The best values for the two MIRGS parameters C_1 and C_2 were found to be $C_1 \in [3, 3.5]$, with C_2 having very little influence if the C_1 value was set correctly. The default MIRGS parameters are in this range and produced results that were very close to the best results obtained. Use of the standard MIRGS algorithm is recommended for dual-polarization imagery as it performs as well as or better than any of the proposed methods, with greater than 80% segmentation accuracy on both the real and synthetic images.

The hierarchical segmentation scheme was able to overcome the incidence angle variation of the appearance of open water. It involves first segmenting the HV channel to separate open water and ice, each of which is then segmented independently using information from the HH channel. However, automatic generation of the segmentation hierarchies was not successful because it requires prior estimation of the number of clusters distinguishable in each channel. The estimation technique used here, the Davies Bouldin (DB) index, was frequently in error by one class in terms of its estimation of the number of classes, affecting the accuracy of the segmentation results. It is therefore recommended that the hierarchies be constructed by trained users until a solution to this is found. An additional problem with the hierarchical segmentation scheme is that if the open water portion of the segmentation needs to be further segmented using the HH band, the incidence angle effect can still confuse the segmentation algorithm. Finally, the hierarchical segmentation technique should not be used on images without the incidence angle effect because the best accuracy obtained is lower than that of using the standard MIRGS algorithm with the correct

parameters.

The answers posed in Section 2.9 can now be answered:

1. What methods can be used to increase the accuracy in the dual-polarization segmentation results?

   MIRGS does not benefit from intelligent gradient combination and data fusion of the dual-polarization data. The standard MIRGS algorithm consistently produces the highest segmentation accuracy. The default parameters are already near the optimal values as determined by the experiments.

2. Can the incidence angle effect that causes confusion between open water and ice be eliminated by using dual-polarization data?

   The hierarchical segmentation scheme is able to eliminate the confusion between open water and ice, but only when the hierarchies are user-defined. There remains some question as to whether the hierarchical segmentation works for all scenes, specifically those where ice types that are hard to separate in the HV channel are all present.

5.2 Future Work

Several lines of future work can be identified. First, the tests should be performed with additional CIS validated data to strengthen the conclusions. This would require creation of additional ground-truth images and / or synthetic images from additional operational RADARSAT-2 data sets with the help of expert sea ice analysts. Much of the future work should be focused on solving the issues with the incidence angle effect, as this is one of the obstacles to fully automatic, full scene segmentation. Because none of the data fusion methods improved on the performance of the standard MIRGS algorithm, the next step would be to improve the model used by MIRGS, perhaps to take into account the incidence angle as part of the segmentation process. Another possibility is to incorporate an incidence angle aware classification step after MIRGS is used in order to deal with the segmentation errors caused by the incidence angle.

Finally, the hierarchical segmentation process could be improved. Better ways to automatically generate the hierarchies, such as investigating other methods for

determining the number of clusters than can be distinguished, would reduce the user input requirements of the method. The DB index could be modified so that the index is aware of the spatial context model. Such an index can be used to evaluate a MIRGS segmentation result as opposed to the K-means segmentation result. While the K-means segmentation result is similar to the MIRGS result, it is sufficiently different that the number of clusters estimation may also be different. Other areas for investigation include whether the post-process merging criterion can be improved and whether some other feature set could be added to the hierarchy to resolve remaining ambiguities. For example, incorporation of passive microwave image data or texture features derived from the dual-polarization imagery might be considered.

References

[1] R. D. Abreu, D. Flett, B. Scheuchl, and B. Ramsay. Operational sea ice monitoring with RADARSAT-2 - a glimpse into the future. In *Proc. IEEE International Geoscience and Remote Sensing Symposium*, volume 2, pages 1308 – 1310, July 2003. 26

[2] T. Acharya and A. K. Ray. *Image Processing: Principles and Applications*. Wiley-Interscience, 2005. 34

[3] C. M. Bachmann, T. L. Ainsworth, and R. A. Fusina. Exploiting manifold geometry in hyperspectral imagery. *IEEE Trans. Geosci. Remote Sens.*, 43(3):441 – 454, Mar. 2005. 27

[4] G. I. Belchanskya and D. C. Douglas. Seasonal comparisons of sea ice concentration estimates derived from SSM/I, OKEAN, and RADARSAT data. *Remote Sensing of Environment*, 81(1):67 – 81, July 2002. 2

[5] M. Belkin and P. Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Computation*, 15(6):1373–1396, Mar. 2003. xiii, 43, 44, 45

[6] J. Bentley. Multidimensional binary search trees used for associative searching. *Communications of the ACM*, 18(9):509 – 517, 1975. 47

[7] T. Bishop, S. Fienberg, and P. Holland. *Discrete Multivariate Analysis - Theory and Practice*. MIT Press, Cambridge, 1975. 56

[8] A. Bouvet, T. L. Toan, and N. Lam-Dao. Monitoring of the rice cropping system in the Mekong Delta using ENVISAT/ASAR dual polarization data. *IEEE Trans. Geosci. Remote Sens.*, 47(2):517 – 526, Feb. 2009. 26

[9] P. J. Burt and R. J. Kolczynski. Enhanced image capture through fusion. In *Proc. Fourth International Conference on Computer Vision*, pages 173 – 182, Berlin, May 1993. 36

[10] Canadian Ice Service. *Manual of Standard Procedures for Observing and Reporting Ice Conditions*. Meteorological Service of Canada, Ottawa, 9th edition, 2005. 4, 5, 6

[11] J. Canny. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.*, 8(6):679 – 698, Nov. 1986. 31

[12] D. Clausi, A. Qin, M. Chowdhury, P. Yu, and P. Malliard. MAGIC: MAp-Guided Ice Classification System. *Canadian Journal of Remote Sensing*, (accepted, April 2009). xiii, 2, 14, 29

[13] D. A. Clausi. K-means iterative Fisher (KIF) unsupervised clustering algorithm applied to image texture segmentation. *Pattern Recognition*, 35:1959 – 1972, 2002. 53

[14] M. J. Collins and C. E. Livingstone. On the dimensionality of multiparameter microwave image data from the thin sea ice in the Labrador Sea. *IEEE Trans. Geosci. Remote Sens.*, 34(1):114 – 136, Jan. 1996. 27, 32

[15] R. G. Congalton, R. G. Oderwald, and R. A. Mead. Assessing Landsat classification accuracy using discrete multivariate analysis statistical techniques. *Photogram. Eng. Remote Sens.*, 49(12):1671–1678, Dec. 1983. 56

[16] D. L. Davies and D. W. Bouldin. A cluster separation measure. *IEEE Trans. Pattern Anal. Mach. Intell.*, 1(2):224–227, Apr. 1979. 51

[17] M. R. Drinkwater. LIMEX 87 ice surface characteristics: Implications for C-Band SAR backscatter signatures. *IEEE Trans. Geosci. Remote Sens.*, 27(5):501–513, Sept. 1989. 11

[18] R. Duda, P. Hart, and D. Stork. *Pattern Classification*. Wiley, New York, 2nd edition, 2001. 2, 19, 52, 53

[19] D. G. Flett. Operational use of SAR at the Canadian Ice Service: present operations and a look to the future. In *Proc. 2nd Workshop on Coastal and Marine Applications of SAR*, Svalbard, Norway, Sept. 2003. 1, 96

[20] D. G. Flett. private communication, 2006. 20

[21] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Mach. Intell.*, 6(6):721 – 741, Nov. 1984. 19, 63

[22] X. He and P. Niyogi. Locality preserving projections. In S. Thrun, L. Saul, and B. Schölkopf, editors, *Advances in Neural Information Processing Systems 16*. MIT Press, Cambridge, MA, 2004. 41

[23] S. Hong, W. M. Moon, H.-Y. Paik, and G.-H. Choi. Data fusion of multiple polarimetric SAR images using Discrete Wavelet Transform (DWT). In *Proc. IEEE International Geoscience and Remote Sensing Symposium*, volume 6, pages 3323 – 3325, Nov. 2002. 27

[24] A. K. Jain and R. C. Dubes. *Algorithms for Clustering Data*. Prentice Hall, New Jersey, 1988. 48, 52, 53

[25] O. M. Johannessen, L. Bengtsson, M. W. Miles, S. I. . Kuzmina, V. A. Semenov, G. V. Alekseev, A. P. Nagurnyi, V. F. Zakharov, L. P. Bobylev, L. H. Pettersson, K. Hasselmann, and H. P. Cattle. Arctic climate change: observed and modelled temperature and sea-ice variability. *Tellus - Series A - Dynamic Meteorology and Oceanography*, 56(5):328 341, Aug. 2004. 1

[26] J. A. Karvonen and M. Similae. Pulse-coupled neural network for sea ice SAR image segmentation and classification. In T. Lindblad, M. L. Padgett, and J. M. Kinser, editors, *Proc. Ninth Workshop on Virtual Intelligence/Dynamic Neural Networks*, volume 3728, pages 333–350, Stockholm, Sweden, June 1999. SPIE. 11

[27] H.-C. Lee and D. R. Cok. Detecting boundaries in a vector field. *IEEE Trans. Signal Process.*, 39(5):1181 – 1194, May 1991. xiv, 16, 23, 30, 31

[28] H. Li, L. Teng, W. Chen, and I.-F. Shen. Supervised learning on local tangent space. In *Advances in Neural Networks ISNN 2005*, volume 3496 of *Lecture Notes in Computer Science*, pages 546 – 551. Springer-Verlag, 2005. 46

[29] X. Li and J. Chong. Processing of Envisat alternating polarization data for vessel detection. *IEEE Geoscience and Remote Sensing Letters*, 5(2):271 – 275, Apr. 2008. 26

[30] S. G. Mallat. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 11(70):674 – 693, July 1989. 34

[31] M. Manore, D. G. Flett, R. A. D. Abreu, B. R. Ramsay, and J. van der Sanden. Multi-polarization SAR data for operational ice monitoring. In *Proc. IEEE International Geoscience and Remote Sensing Symposium*, volume 3, pages 1246 – 1248, Sydney, Australia, July 2001. 10, 26, 27

[32] MDA. RADARSAT-2 Product Format Definition. Technical Report RN-RP-51-2713, MacDonald, Dettwiler and Associates Ltd, 2008. `http://www.radarsat2.info/product/51-2713_-_RSAT-2_Product_Format_Definition_-_Iss1_7.pdf`. 91, 92

[33] P. Minton. Global GSHHS land mask in SHP format. `http://www.evs-islands.com/2007/11/data-global-gshhs-land-mask-in-shp.html`. 91

[34] D. C. Munson, Jr. and R. L. Vinsentin. A signal processing view of strip-mapping synthetic aperture radar. *IEEE Trans. Acoust., Speech, Signal Process.*, 37(7):2131 – 2147, December 1989. 8

[35] K. Nakamura, H. Wakabayashi, K. Naoki, F. Nishio, T. Moriyama, and S. Uratsuka. Observation of sea-ice thickness in the Sea of Okhotsk by using dual-frequency and fully polarimetric airborne SAR (Pi-SAR) data. *IEEE Trans. Geosci. Remote Sens.*, 43(11):2460 – 2469, Nov. 2005. 26

[36] G. P. Nason and B. W. Silverman. The stationary wavelet transform and some statistical applications. In A. Antoniadis and G. Oppenheim, editors, *Wavelets and Statistics*, volume 103 of *Lecture Notes In Statistics*, pages 281–300. Springer-Verlag, 1995. 35

[37] C. Oliver and S. Quegan. *Understanding Synthetic Aperture Radar Images*. SciTech Publishing, Raleigh, 2004. 8, 31

[38] R. G. Onstott. SAR and scatterometer signatures of sea ice. In F. D. Carsey, editor, *Microwave Remote Sensing of Sea Ice*, volume 68 of *Geophysical Monograph*, chapter 5, pages 73 – 104. American Geophysical Union, Washington, 1992. 9

[39] J. R. Orlando, R. Mann, and S. Haykin. Classification of sea-ice images using a dual-polarized radar. *IEEE Journal of Oceanic Engineering*, 15(3):228 – 237, July 1990. 26, 27

[40] G. Pajares and J. M. de la Cruz. A wavelet-based image fusion tutorial. *Pattern Recognition*, 37(9):1855 – 1872, Sept. 2004. 27, 34, 36

[41] N. W. Park and K. H. Chi. Integration of multitemporal/polarization C-band SAR data sets for land-cover classification. *International Journal of Remote Sensing*, 29(16):4667 – 4688, Jan. 2008. 26

[42] C. Pohl and J. L. V. Genderen. Multisensor image fusion in remote sensing: concepts, methods and applications. *International Journal of Remote Sensing*, 19(5):823 – 854, 1998. 27

[43] A. K. Qin and D. A. Clausi. Multivariate image segmentation using semantic region growing with adaptive edge penalty. *IEEE Trans. Image Process.*, (submitted, May 2009). xiii, 14, 18

[44] B. Ramsay, D. Flett, H. S. Andersen, R. Gill, S. Nghiem, and C. Bertoia. Preparation for the operational use of RADARSAT-2 for ice monitoring. *Canadian Journal of Remote Sensing*, 30(3):415 – 423, June 2004. 1, 10, 26

[45] J. A. Richards and X. Jia. *Remote Sensing Digital Image Analysis*. Springer, Berlin, 4th edition, 2006. 7, 56

[46] O. Rockinger. Image sequence fusion using a shift-invariant wavelet transform. In *Proc. IEEE International Conference on Image Processing*, volume 3, pages 288 – 291, Santa Barbara, CA, USA, Oct. 1997. 35

[47] S. Roweis and L. Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 22(5500):2323–2336, Dec. 2000. xiii, 43, 44

[48] B. Scheuchl, R. Caves, D. Flett, R. Abreu, M. Arkett, and I. Cumming. ENVISAT ASAR AP data for operational sea ice monitoring. In *Proc. IEEE International Geoscience and Remote Sensing Symposium*, volume 3, pages 2142 – 2145, Sept. 2004. 10

[49] B. Scheuchl, R. Caves, D. Flett, R. D. Abreu, M. Arkett, and I. Cumming. The potential of cross-polarization information for operational sea ice monitoring. In *Proc. Envisat & ERS Symposium*, Salzburg, Austria, Sept. 2004. 11, 26

[50] B. Scheuchl, D. Flett, R. Caves, and I. Cumming. Potential of RADARSAT-2 data for operational sea ice monitoring. *Canadian Journal of Remote Sensing*, 30(3):448 – 461, June 2004. 1, 10, 26, 27, 32

[51] B. Scheuchl, R.Caves, I. Cumming, and G. Staples. Automated sea ice classification using spaceborne polarimetric SAR data. In *Proc. IEEE International Geoscience and Remote Sensing Symposium*, volume 7, pages 3117 – 3119, Sydney, Australia, July 2001. 26

[52] R. Schowengerdt. *Remote Sensing: Models and Methods for Image Processing.* Academic Press, San Diego, second edition, 1997. 27, 40

[53] N. Shepherd. Extraction of Beta Nought and Sigma Nought from RADARSAT CDPF products. Technical Report AS97-5001 Rev. 4, ALTRIX Systems, April 2000. 93

[54] G. Simone, A. Farina, F. Morabito, S. Serpico, and L. Bruzzone. Image fusion techniques for remote sensing applications. *Information Fusion*, 3(1):3 – 15, Mar. 2002. 27, 34, 35, 38

[55] G. Strang. Wavelets. *American Scientist*, 82(3):250–255, 1994. 34

[56] J. B. Tenenbaum, V. de Silva, and J. C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319–2323, Dec. 2000. 47

[57] F. T. Ulaby, R. K. Moore, and A. K. Fung. *Microwave Remote Sensing: Active and Passive.* Artech House, Norwood, MA, USA, 1986. 11

[58] L. Vincent and P. Soille. Watersheds in digital spaces: An efficient algorithm based on immersion simulations. *IEEE Trans. Pattern Anal. Mach. Intell.*, 13(6):583 – 598, June 1991. 14

[59] K. J. Wilson, J. Falkingham, H. Melling, and R. D. Abreu. Shipping in the Canadian Arctic: other possible climate change scenarios. In *Proc. IEEE International Geoscience and Remote Sensing Symposium*, volume 3, pages 1853 – 1856, Sept. 2004. 1

[60] T. Wittman. MANIfold learning MATLAB demo, Apr. 2005. `http://www.math.umn.edu/~wittman/mani/`. 47

[61] A. Wong, P. Yu, W. Zhang, and D. A. Clausi. IceSynth II: Synthesis of SAR sea-ice imagery using region-based posterior sampling. *IEEE Geosci. Remote Sens. Lett.*, (submitted, June 2009). 96

[62] A. Worbya and J. Comiso. Studies of the Antarctic sea ice edge and ice extent from satellite and ship observations. *Remote Sensing of Environment*, 92(1):98 – 111, July 2004. 2

[63] P. Yu, D. A. Clausi, and S. E. L. Howell. Fusing AMSR-E and QuikSCAT imagery for improved sea ice recognition. *IEEE Trans. Geosci. Remote Sens.*, 47(7):1980–1989, July 2009. 27

[64] Q. Yu and D. A. Clausi. SAR sea-ice image analysis based on iterative region growing using semantics. *IEEE Trans. Geosci. Remote Sens.*, 45(12):3919 – 3931, December 2007. 14, 17, 79

[65] Q. Yu and D. A. Clausi. IRGS: Image segmentation using edge penalties and region growing. *IEEE Trans. Pattern Anal. Mach. Intell.*, 30(12):2126 – 2139, December 2008. 2, 13, 14, 17, 18, 19

[66] Z. Zhang and R. S. Blum. A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application. *Proc. IEEE*, 87(8):1315 – 1326, Aug. 1999. 34

[67] Z. Zhang and H. Zha. Nonlinear dimension reduction via local tangent space alignment. In *Intelligent Data Engineering and Automated Learning*, volume 2690 of *Lecture Notes in Computer Science*. Springer, Aug. 2003. xiii, 43, 45

Appendix A

Conversion and Calibration of RADARSAT Data

The procedure for converting and calibrating RADARSAT-2 data to a format readable by the MAGIC / IRGS system is described in the following sections. The corresponding process for RADARSAT-1 data, as implemented in the existing source code, is also documented here.

A.1 RADARSAT-2 data

The R2 products are described in the RADARSAT-2 Product Format Definition document [32]. R2 data products consist of several extensible markup language (XML) files containing product metadata and a set of GeoTIFF image files containing the image raster data with one GeoTIFF file for each polarization. The metadata contains ground control points (for georeferencing), satellite orbit and calibration information while the GeoTIFF files contain digital numbers (DNs) that can be converted to backscatter σ° by an appropriate function.

The `calibrateR2` program was developed to convert and calibrate R2 files into a format readable by MAGIC (a choice of BMP or BIL format). The program also rasterizes an appropriate land mask from a vector land database [33] based on the ground control points (GCPs) for the image. The following are the steps of the calibration procedure when `calibrateR2` is executed:

1. Parse the XML files included with the R2 data and extract the ground control points (GCPs) and calibration constants. The calibration constants consist of

an offset B and a list of gains $A(j)$ for each sample j in the range direction of the image.

2. Save an ASCII text file with the GCP information.

3. Based on the ground control points, generate the rasterized land mask and save it as a BMP file.

4. Convert the DNs in the GeoTIFF file to DNs representing σ° in decibels (dB).

5. Save the calibrated data as either an 8 bits per pixel (bpp) BMP or an 8 or 16 bpp BIL file.

Experiments in this thesis used the 8 bpp workflow to conserve memory due to the large size of some of the images. The DNs in the GeoTIFF files are converted to σ° by one of two functions [32]. For DNs that represent real numbers (which is the case with the dual-polarization ScanSAR Wide data used in this thesis), the function used to convert the $DN(i, j)$ at line i and range sample j is:

$$\sigma^\circ(i, j) = 10 \log_{10} \left(\frac{DN(i, j)^2 + B}{A(j)} \right) \quad [\text{dB}] \tag{A.1}$$

where $A(j)$ is the gain at range sample j and B is a constant offset. For DNs that represent complex numbers, the function used is:

$$\sigma^\circ(i, j) = 10 \log_{10} \left(\frac{|DN(i, j)|^2}{A(j)^2} \right) \quad [\text{dB}] \tag{A.2}$$

The σ° in dB is then clamped to a range of $[-35, -5]$ dB and then linearly scaled to DNs with a range of $[0, 255]$ or $[0, 65535]$ (for 8 bpp and 16 bpp output, respectively). The clamp range is user configurable; the values chosen here provide reasonable visual contrast and little clipping of the image histogram. Both HH and HV are clamped to the same range to ensure that the output DNs are the same units.

A.2 RADARSAT-1 data

R-1 data are received from the CIS in a special format that consists of AVG files. There is no documentation for this file format. The `calibrateR1` program is able to read these files and convert them into BIL files for MAGIC to read. The process is very similar to that used for calibrating R2 data, except the output is always 8 bpp and there is no automatic land mask generation. The following steps are derived from a code audit of the `calibrateR1` program:

1. Parse the SAR trailer file (with extension .SART.AVG) file included with the R-1 data and extract the orbit information and a look up table of calibration constants. The calibration constants consist of an offset B and a list of gains $A(j_{LUT})$ for some regularly spaced samples j_{LUT} in the range direction of the image.

2. Read the SAR data file (with extension .SARD.AVG) and extract ground control points. Save these to an ASCII GCP file.

3. Calculate incidence angle $I(j)$ at each range pixel and interpolate the gain look up table $A(j_{LUT})$ in order to get the gains $A(j)$ at each range pixel j.

4. Convert the DNs in the SAR data file to DNs representing σ° in dB.

5. Save the calibrated data as an 8 bpp BIL file.

Rather than Equation A.1, R-1 DNs are calibrated by the following expression [53]:

$$\sigma^\circ(i,j) = 10\log_{10}\left(\frac{DN(i,j)^2 + B}{A(j)}\right) + 10\log_{10}\left(\sin I(j)\right) \quad [\text{dB}] \tag{A.3}$$

The incidence angle $I(j)$ is calculated from the orbit information contained in the SAR trailer file. The calculation of the incidence angle is described in [53]. Once the σ° in dB is calculated, it is clamped to a range of $[-25, 0]$ dB and then linearly scaled to DNs with a range of $[0, 255]$.

Appendix B

Test Images

This chapter provides some details about the test images used in the thesis.

B.1 Real image

The real image is a subimage extracted from the February 25, 2009 Gulf of St. Lawrence scene. The image dimensions are 512×512 pixels. Since the image was extracted from the near range of the scene, it represents the appearance of ice for incidence angles of less than 10°. A manual ground-truth image was created based on information from the CIS ice chart, which indicated four types of ice. Manual segmentation was performed by grouping the pixels based on their appearance in an RGB composite (R=HV, G=HV, B=HH). The image and the manually created ground-truth is shown in Figure B.1.

B.2 Synthetic image

A synthetic image was created in order to test the effectiveness of each of the techniques considered in the thesis. Although a real R2 image was manually segmented to create a ground-truth segmentation for testing, the segmentation may not be fully accurate. The synthetic image solves this problem by providing an unambiguous ground-truth.

The synthetic image was created by first creating a template image corresponding to the desired ground-truth with size 1024 by 1024 pixels. Various shapes such as

(a) HH
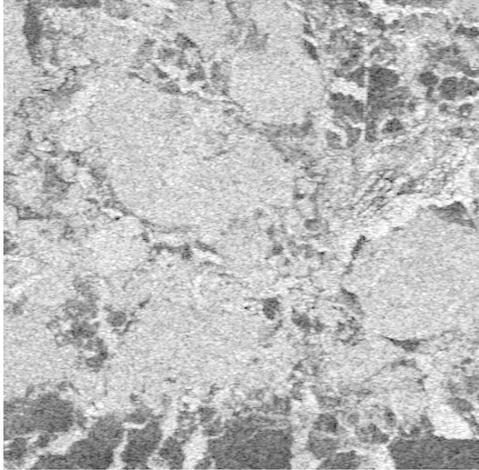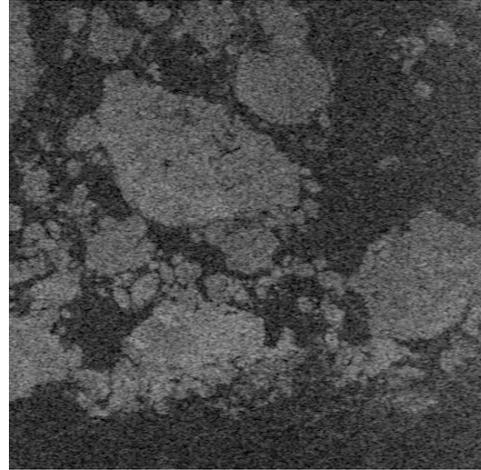
(b) HV
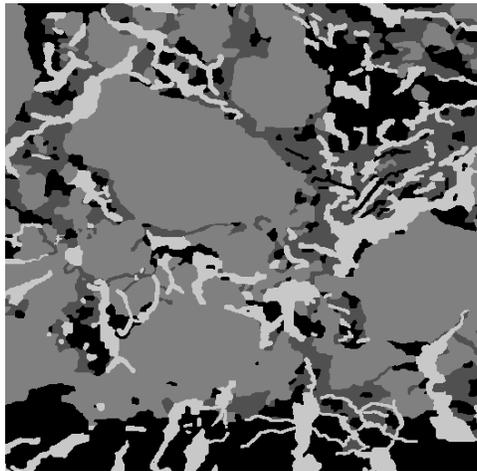
(c) Ground-truth

| ■ Smooth ice | ■ Gray Ice | ■ First Year Ice | ■ Open Water |

Figure B.1: Real image and its associated ground-truth.

circles and small lines were drawn manually in a vector drawing program to create both large and small details. A class number from 1 to 4 was assigned to each shape.

Each ice class in the ground-truth image was then filled with a texture derived from real dual-polarization imagery. The main requirement for the texture was for it to be sourced from only a single ice class. To do this, CIS ice charts were consulted to find ice polygons which were dominated by one ice class. The ice classification of these polygons can therefore be considered unambiguous. A dual-polarization subscene was extracted from 4 polygons corresponding to open water, smooth gray ice, gray ice and first year ice. Due to the highly inhomogeneous nature of sea ice, areas with only a single ice class are very rare and tend to be small. Thus, only very small subscenes could be extracted for each ice class.

The IceSynth II [61] image synthesis system was used to extend the size of the subscenes to 1024 x 1024 pixels. The final synthetic image was then composed by filling all pixels marked with a certain class by the corresponding dual-polarization texture image.

The subscenes used for texture generation were all extracted from the mid-range incidence angles (30° to 35°). In this particular incidence angle range, all of the classes except smooth ice look very similar in the HH band, which explains the low contrast nature of the HH channel.

The synthetic image and the template used to create it are shown in Figure B.2.

B.3 Full scene image

The full scene image was extracted from the March 4, 2009 Gulf of St. Lawrence scene. It covers the entire incidence angle range of ScanSAR wide. The original scene dimensions are 10000×4000 pixels but this was 4×4 block averaged to 2500×1000 pixels. The CIS performs 2×2 block averaging of all R1 scenes before processing [19]. The additional block averaging performed here is due to memory limitations as the RAG used by MIRGS requires more than the test system's 2 GB of memory for large images.

The full scene image is shown in Figure B.3. The expected segmentation was derived from a user defined SITH result that appeared reasonable. Since it is not a true ground-truth, quantitative accuracy statistics cannot be calculated from it.
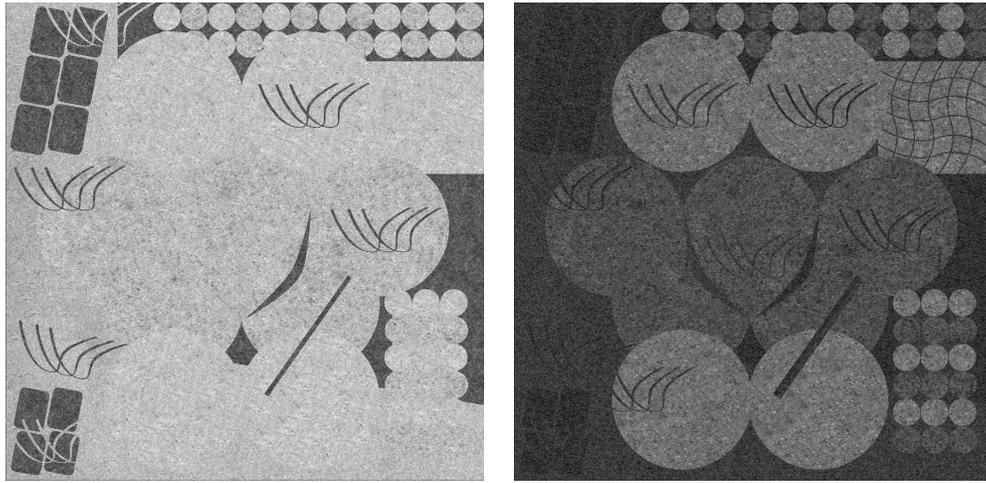
(a) Synthesized HH

(b) Synthesized HV
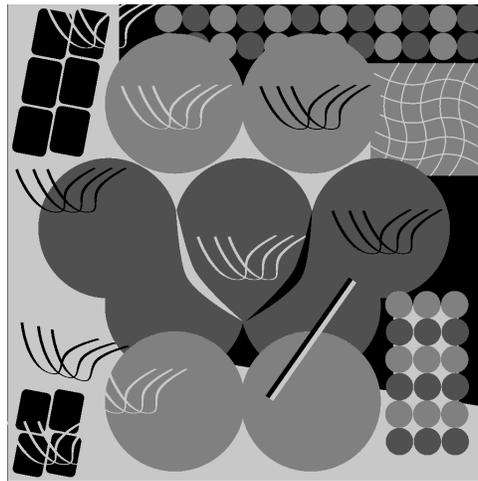
(c) Template / ground-truth

■ Smooth Gray Ice   ■ Gray Ice   ■ First Year Ice   ■ Open Water

Figure B.2: Each class in the template / ground-truth image is filled with a synthesized sea ice texture in order to create the HH and HV images.
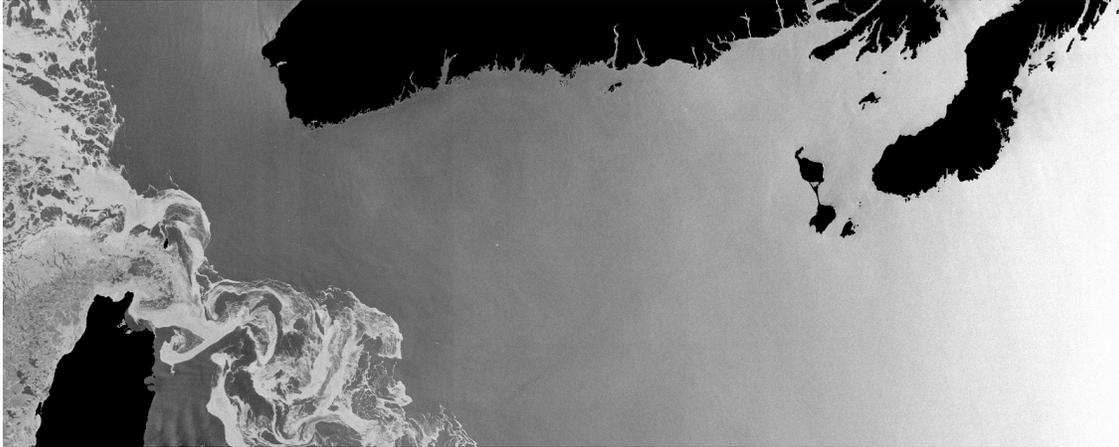
(a) HH image
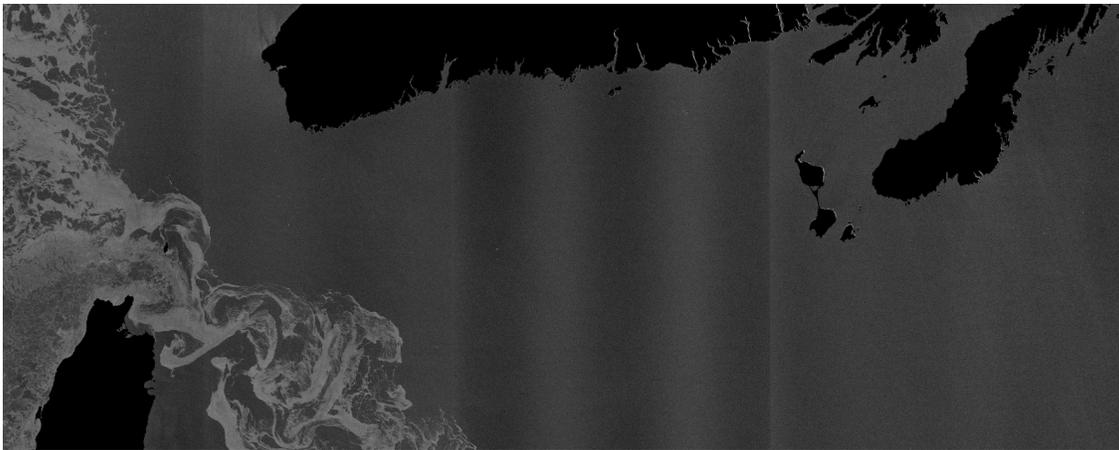
(b) HV image

(c) Expected segmentation

■ Land　　■ Gray Ice　　■ First Year Ice　　■ Open Water

Figure B.3: Full scene image and its expected segmentation.