Feature extraction of dual-pol SAR imagery for sea ice image segmentation[☆]

Peter Yu[1,*], A. K. Qin[2], David A. Clausi[1]

[1]*Department of Systems Design Engineering, University of Waterloo, Waterloo, ON, Canada*
[2]*MISTIS Team, INRIA Grenoble Rhône-Alpes, ZIRST, 655 avenue de l'Europe, Montbonnot, 38334 Saint Ismier Cedex, FRANCE*

Abstract

Dual-polarization synthetic aperture radar (SAR) image data, such as that available from RADARSAT-2, provides additional information for discriminating sea ice types compared to single-polarization data. A thorough investigation of published feature extraction and fusion techniques for making optimal use of this additional information for unsupervised sea ice image segmentation has been performed. Segmentation was performed by transforming the dual-pol data (a) into a new two channel feature space (multivariate) and (b) into a fused single channel feature space (univariate). Both real and synthetic dual-polarization SAR sea ice images were transformed using a variety of methods and segmented using a recognized SAR segmentation algorithm (IRGS). The results indicate that the untransformed data provides consistent and high segmentation accuracy, avoids feature extraction pre-processing, and is thus recommended for SAR sea ice image segmentation using dual-pol imagery.

1. Introduction

Sea ice mapping is an important application of remote sensing systems, essential for understanding the Arctic climate system (Johannessen et al., 2004) and for safe navigation of ships in waters where sea ice can form (Wilson et al., 2004). Single-polarization images from the Canadian RADARSAT-1 (RS-1) synthetic aperture radar (SAR) are an important source of information for operational sea ice mapping (Flett, 2003). RADARSAT-2 (RS-2), launched in 2007, is a Canadian SAR satellite that offers an operationally useful dual-polarization mode that is expected to improve discrimination of water from ice (Ramsay et al., 2004) and better distinguish between different types of ice (Scheuchl et al., 2004b). These enhancements are important because under certain but common circumstances, interpreting the various types of ice and water in single-polarization images can be ambiguous.

This paper investigates the advantages of using dual-polarization SAR image data from RS-2 for the purposes of automatic, unsupervised sea ice image segmentation. Current operational sea ice maps are produced manually by human analysts. This process is subjective and labor-intensive, so automated ice mapping algorithms are desired. Consideration of the advantages of dual-polarization data for automated sea ice mapping algorithms is motivated. This paper focuses on the use of RS-2's dual-polarization HH and HV magnitude data since it is the most operationally useful mode due to its 500 km swath width (Ramsay et al., 2004).

Multi-polarimetric (full- and dual-polarization) SAR imagery improves the classification and segmentation of sea ice as compared to single polarization imagery and many examples can be found in the literature. Fully-polarimetric SAR imagery, consisting of HH, HV, VH and VV channels, was assessed in (Manore et al., 2001), which found that using either the cross-polarization channels or the co-polarization ratio HH/VV can improve ice-water discrimination. The HH/VV ratio can also be used to estimate sea ice thickness (Nakamura et al., 2005). Similarly, The cross-polarization ratio HV/HH improves sea ice discrimination (Scheuchl et al., 2004b) and is one of the principal multi-polarization parameters (Collins and Livingstone, 1996). Fully polarimetric or HH and VV data has also been used for other applications, such as land cover classification (Park and Chi, 2008), ship detection (Li and Chong, 2008) and crop monitoring (Bouvet et al., 2009).

Dual-polarization ENVISAT ASAR data, which is similar to RADARSAT-2 data, improve the separation of water and ice when used with an unsupervised segmentation algorithm (Scheuchl et al., 2004a). RGB composite images of HH and HV data significantly improved visual discrimination of open water and ice (De Abreu et al., 2003). Another work tested dual-polarization data from a Ku-band SAR with three different classifiers (Orlando et al., 1990). The classification accuracy of first year ice, multi-year ice and icebergs was improved. Feature extraction using the co-polarized and cross-polarized channels was also performed using principal component analysis (PCA), which improved the visual distinction of the different ice types but did not improve classification accuracy.

In addition to direct usage of multi-polarization data, or fusing the multi-polarization bands with polarimetric band ratios, generic image fusion techniques have also been considered for combining multi-polarization imagery into single-band images for visualization or processing. Wavelet image fusion, which decomposes images into multi-resolution wavelet coefficient bands and combines the coefficients from multiple images according to defined rules (Pajares and de la Cruz, 2004), has been used for multi-polarization SAR image fusion (Simone et al., 2002; Hong et al., 2002; Jin et al., 2006; Zhang et al., 2007). Other image fusion techniques

[☆]Citation: Peter Yu, A.K. Qin and David A. Clausi. Feature extraction of dual-pol SAR imagery for sea ice image segmentation. Canadian Journal of Remote Sensing. DOI: 10.5589/m12-028.

[*]Corresponding author (Website: http://www.peteryu.ca)

are based on similar principles of multi-resolution decomposition (Zhang and Blum, 1999; Yang et al., 2009). The results of these image fusion techniques have generally been assessed by their visual appearance, either subjectively or with some objective measure of image quality, rather than by the efficacy of an image segmentation or classification algorithm applied to the fused result.

The research literature indicates that there is potential for improvements in sea ice discrimination from multi-polarization data, but we know of no papers that specifically test RS-2's dual-polarization magnitude data with a state-of-the-art sea ice image segmentation algorithm. Although previous papers have assessed the improvements of using dual-polarization data for sea ice discrimination, the assessment was either a visual assessment (Manore et al., 2001; De Abreu et al., 2003) or using traditional, pixel-based image classification schemes (Scheuchl et al., 2004a; Orlando et al., 1990). Therefore, this paper will focus on investigating methods to most effectively use dual-polarization RS-2 data in the novel Iterative Region Growing with Semantics (IRGS) image segmentation algorithm (Yu and Clausi, 2008). This algorithm is part of a larger system called MAGIC (Map-Guided Ice Classification) (Clausi et al., 2010), which aims to provide pixel level accurate ice maps given a manually created ice map using operational data. The IRGS algorithm has been demonstrated to be an improvement over other state-of-the-art segmentation algorithms for both generic (Qin and Clausi, 2010) and SAR (Yu and Clausi, 2008, 2007) imagery. IRGS has also been successfully applied to IKONOS imagery and improves upon other algorithms for segmentation of Brazilian savannahs (Barbosa and Maillard, 2010). IRGS differs from other image segmentation methods by using a region-based segmentation scheme with full consideration of image edges as part of the segmentation model.

In this paper, the potential improvements of using dual-polarization imagery with IRGS are investigated. The experiments will determine whether the dual-polarization data is best used in a multivariate framework, as in a multi-variate extension of IRGS (MIRGS) (Qin and Clausi, 2010), or in a univariate data fusion framework in which the two polarization channels are first combined into a single band before being given to the univariate IRGS algorithm. This will address the question of whether feature extraction or image fusion is helpful for automated algorithms and, if so, which technique is best in the context of the MIRGS image segmentation algorithm. To the best of our knowledge, no such study exists even though many different feature extraction and fusion methods have been proposed. This work is available to practitioners attempting to automatically interpret dual-polarization SAR imagery.

Section 2 gives a brief overview of the IRGS algorithm that is used for the experiments. Section 3 states the research objectives that will be considered. Section 4 describes the tested image data. Section 5 describes the tested feature extraction and fusion methods and Section 6 presents the results. Finally, Section 7 concludes the paper.

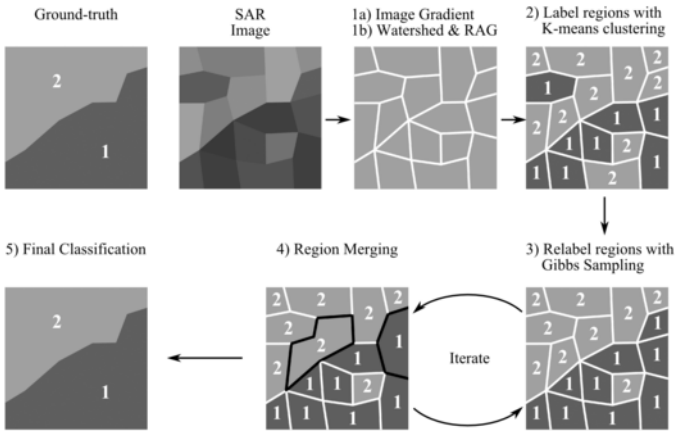

Figure 1: Major steps of the MIRGS algorithm. 1a) To initialize the system, MIRGS computes the image gradient and 1b) generates the watershed regions (with simplified shapes in this diagram; they are more arbitrary in general) and the region adjacency graph (RAG). 2) The watershed regions are then labeled with K-means clustering. 3) Regions are relabeled by Gibbs sampling (Geman and Geman, 1984). 4) Regions with the same label are merged. Steps 3 and 4 are repeated for a user-specified number of iterations. 5) The final segmentation is produced.

2. Iterative Region Growing with Semantics algorithm

The MIRGS algorithm (Qin and Clausi, 2010), which is the multivariate version of the original univariate IRGS algorithm (Yu and Clausi, 2008, 2007), is the unsupervised image segmentation algorithm that is applied to the dual-polarization RS-2 data. "Unsupervised" indicates that no *a priori* assumptions of the class distributions is used to drive the segmentation process; "supervised" methods typically use training data to estimate the class distributions. Since MIRGS is unsupervised, it does not have the necessary information to perform labeling with specific ice class names. In the MAGIC system, labeling is performed after segmentation to complete the classification process (Clausi et al., 2010). Automatic labeling is another line of research (Ochilov et al., 2010).

IRGS has been successfully applied to single-polarization RS-1 HH imagery. The results have been evaluated by Canadian Ice Service (CIS) experts and found to outperform other methods (Clausi et al., 2010; Yu and Clausi, 2007). As such, MIRGS is the segmentation method used in this paper and this section summarizes the published description of MIRGS (Qin and Clausi, 2010).

Fig. 1 shows the major steps of MIRGS. The algorithm accepts as input an image with at least one image channel. The image is first segmented with a watershed algorithm (Vincent and Soille, 1991) that divides the image into many small regions each with relatively uniform backscatter. The image gradient must be computed to generate the watershed segmentation and it is straightforward to calculate the gradient of a single channel image. However, for multichannel images, MIRGS uses a vector field gradient (VFG) approach (Lee and Cok, 1991) to calculate the joint image gradient from all channels since the edge content in each channel may be different.

The computed gradient is normalized so that the largest gradient in the scene has a value of 1.0 to ensure that MIRGS scales properly for scenes with different dynamic ranges. Once the watershed (Vincent and Soille, 1991) is generated from the normalized image gradient, the image is represented by a region adjacency graph (RAG) (Li, 2001) data structure where each node represents a watershed region and where each graph edge connects spatially adjacent regions.

Each watershed region is assigned an initial label via a K-means algorithm (Duda et al., 2001) to initialize MIRGS. The region-based K-means algorithm used in MIRGS is described in (Qin and Clausi, 2010). MIRGS then enters an iterative phase to find a configuration of labels for the regions that globally minimizes a cost function. At each iteration, a labeling process is performed with Gibbs sampling (Geman and Geman, 1984) to move the segmentation towards the optimal configuration. After each iteration, regions with the same labels are merged to reduce the number of nodes in the RAG by combining adjacent regions, which makes subsequent iterations more efficient as fewer nodes have to be considered.

The cost function that MIRGS minimizes to produce the optimal segmentation consists of a feature space model and a spatial context model (Qin and Clausi, 2010). The cost function considers a segmentation more likely to be "true" when the regions assigned to each class are similar to each other in feature space and when spatially adjacent regions belong to the same class if the edge strength between them is weak. This is similar to the Markov random field (MRF) based multilevel logistic (MLL) segmentation model (Derin and Elliott, 1987) but MLL does not consider the edge strength in its spatial context model. The MIRGS model agrees more closely with intuition: if there is a strong edge between two regions, they are more likely to be from different classes than when there is no edge. MLL, in contrast, makes no such distinction and favours results where adjacent regions are assigned to the same class regardless of the edge strength between them.

3. Research objectives

To generate an accurate and consistent segmentation, the MIRGS algorithm requires data with sufficient feature space separability for different ice classes, i.e. it should be possible to discern a difference in feature space values between different classes (for example, one class might appear darker than another class in the image). Additionally, MIRGS requires the proper generation of the initial watershed and an image gradient that presents strong boundaries between regions of different ice classes. As will be seen in Section 4, information from both dual-polarization channels (HH and HV) are necessary. Many strategies exist to use dual-polarization RS-2 data to satisfy these requirements. The objective of this study is to determine which of these strategies is the most effective. The following three strategies will be tested:

1. *Direct MIRGS implementation*: The most basic strategy is to use the backscatter values from the HH and HV channels directly in the multivariate formulation of MIRGS, using the VFG gradient method that is already implemented to create the watershed and image gradient.
2. *Gradient combination*: While the feature space separability provided by the dual-polarization data is fully utilized by Strategy 1, the VFG image gradient was not designed with domain knowledge of dual-polarization data. VFG tends to assign the highest strength only to edges that are strongest in *both* the HH and HV channels while strong edges that appear in only *one* of the two channels are assigned a lower value. However, strong edges that appear in at least one of the channels are equally meaningful as they denote a boundary between ice classes. Thus, there is motivation for testing and comparing various gradient generation strategies that combine information from both channels.
3. *Feature extraction and image fusion*: Another strategy for making use of dual-polarization data is to fuse the information from both channels into a single image first with feature extraction or image fusion techniques before segmentation in MIRGS. If feature space separability can be maintained between all ice classes after mapping each two dimensional feature vector to a one dimensional value, then both the separability and the image gradient requirements can be satisfied: all ice classes will have a different brightness in the fused image, which will naturally cause edges between them to appear in the fused image.

The experiments will test whether the basic multivariate strategy, a modified gradient combination approach or feature extraction will give the best results for the RS-2 data.

4. Data

ScanSAR Wide A has a pixel resolution of 100 m × 100 m, with a pixel spacing of 50 m × 50 m. The full 500km swath width spans approximately 10000 × 10000 pixels. The CIS expects to use data from the co-polarization (σ°_{HH}) and the cross-polarization (σ°_{HV}) channels for their operations and has provided real-valued RS-2 imagery for testing. Each pixel in the image is represented by a two dimensional feature vector whose elements are σ°_{HH} and σ°_{HV}. The HH channel contains the same information as that available from the single-polarization RADARSAT-1 (RS-1) satellite. Complex-valued images are not considered in this paper because these are not used operationally by CIS.

A Gulf of St. Lawrence scene recorded on February 25, 2008 was tested in this paper. CIS provided operational ice charts for the area on this date, which were created from RS-1 data since CIS had not yet integrated RS-2 imagery into their operational pipeline at that time. A manually segmented ground-truth image was produced based on the ice charts for a small part of the RS-2 scene (depicting an area north of Anticosti Island) to use for validation purposes (Fig. 2). This image represents ice appearance for an incidence angle range of less than 10°.

There are still ambiguities in the manual segmentation because certain ice types cannot always be reliably identified from the backscatter images alone and because each polygon in the CIS ice chart contains a mix of ice types but not the exact pixel location of each type. However, there are small patches of the original, full RS-2 scene where the ice type is known, such as within polygons that have only one ice type or a mix of distinctive ice types. Although the ice type in

these areas is known more reliably, they occur in small, isolated patches, which is not ideal for testing a segmentation algorithm.

To overcome the above difficulties, a synthetic image was created as a second test image (Fig. 3). The synthetic image consists of artificially created shapes, which are each assigned to one of four ice classes. For each ice class, a patch that unambiguously represents that class was identified in the full RS-2 scene and the shapes were filled with textures synthesized from the corresponding patch. Therefore, the ice class of each shape in the synthetic image is known unambiguously, which creates a more reliable ground-truth than the manual segmentation of the image shown in Fig. 2. The IceSynth II algorithm (Wong et al., 2010) synthesized the textures. IceSynth II extends the small patches of known class into a larger synthetic patch with the same textural and backscatter characteristics. In the synthetic image, the joint statistical properties of the HH and HV channels are maintained because IceSynth II draws data from the same locations in both channels as it synthesizes the image. Thus, the synthetic image has the same type of intra-class variation and inter-class contrast characteristics as the original RS-2 image for the selected ice types. All of the patches of known ice type occurred at incidence angles between 30° to 35°.

Although the geometric appearance of the synthetic image is not fully realistic, this does not have a large effect on the segmentation process as MIRGS only takes into account edge strength between regions but not the boundary shape. Since the synthetic image provides a more reliable ground-truth and the real image provides more realistic shapes, both images are tested in this paper to obtain more reliable conclusions.

Both the real and synthetic image show that neither the HH nor the HV channel alone clearly distinguishes between all the ice types within each image. In the real image, water and first year ice look similar in HH while smooth ice, gray ice and water are ambiguous in HV. In the synthetic image, which represents ice at a different incidence angle range, the water, gray ice and first year ice are indistinguishable in HH while smooth ice and water are poorly separated in HV. Therefore, information from both channels is needed for accurate segmentation.

The images tested are purposely limited in incidence angle range to exclude the effects of incidence angle related appearance changes (Drinkwater, 1989; Ulaby et al., 1986) of ice types (especially open water) within the same image. This is similar to how the MAGIC system performs segmentation only within operator created polygons and regions of interest to minimize the effects of incidence angle (Clausi et al., 2010). The HV channel is less sensitive to incidence angle effects (Scheuchl et al., 2004a), but it cannot distinguish all ice types, as explained earlier. Thus, incidence angle effects are not considered in this paper. While this precludes fully automatic sea ice segmentation, the procedure is still extremely valuable since it enables creation of pixel-resolution sea ice maps once an operator created polygon is available.

All image data are provided by CIS as 8-bit digital numbers (DN) that represent the backscatter value at each pixel with a range-dependent gain applied for operational visual analysis. In this paper, the backscatter value in decibels (dB) is recovered with a supplied lookup table prior to use. To facilitate display and data normalization conventions expected by the tested algorithms, the range $[-35, -5]$ dB of the backscatter values has been mapped to $[0, 255]$ for processing. Since the original data provided by CIS are 8-bit DNs, this process does not cause loss of any significant figures. All internal calculations, however, are done with floating point arithmetic.

5. Methods

5.1. Gradient combination

As discussed in Section 3, the multivariate VFG gradient calculation method used by MIRGS in Strategy 1 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 are equally meaningful and should also be assigned a high edge strength. Therefore, three different gradient calculation rules are tested to evaluate Strategy 2. In this section, let $\nabla_s^{(HH)}$ be the gradient at pixel s from the HH image alone, $\nabla_s^{(HV)}$ be the gradient for the HV image alone and ∇_s^{VFG} be the VFG (Lee and Cok, 1991) that MIRGS uses. Let S be the set of all pixels in the image.

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

$$\nabla_s^{MAX} = \max\left\{\nabla_s^{(HH)}, \nabla_s^{(HV)}, \nabla_s^{VFG}\right\} \tag{1}$$

The MAX rule ensures that the ∇_s always reflects the maximum edge strength available from all channels at each pixel. 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.

In the absence of noise, the MAX rule ensures that any edge that is strong in at least one of the channels has high edge strength in the combined gradient. However, for sites with a non-zero gradient due to noise, the MAX rule chooses the largest gradient value and amplifies the effect of noise. Thus, two alternative gradient combination rules are also tested. The first is an Absolute Difference Weighted Average (ADWA) gradient rule. This rule is defined as follows:

$$\nabla_s^{ADWA} = w_s^{ADWA}\nabla_s^{MAX} + (1 - w_s^{ADWA})\nabla_s^{VFG} \tag{2}$$

where:

$$w_s^{ADWA} = \frac{|\nabla_s^{(HH)} - \nabla_s^{(HV)}|}{\max\limits_{s\in S}\left\{|\nabla_s^{(HH)} - \nabla_s^{(HV)}|\right\}} \tag{3}$$

In the ADWA rule, the combined gradient weights heavily toward ∇_s^{MAX} if the difference in gradient magnitude between HH and HV is large. This occurs when one of the channels has a strong edge and the other does not, which is when the MAX rule is appropriate. If the difference in gradient magnitude is small, it is either a pixel with noise (under the assumption that noise has a smaller magnitude than true edges) or the edge is strong in both individual channels. In this latter case,

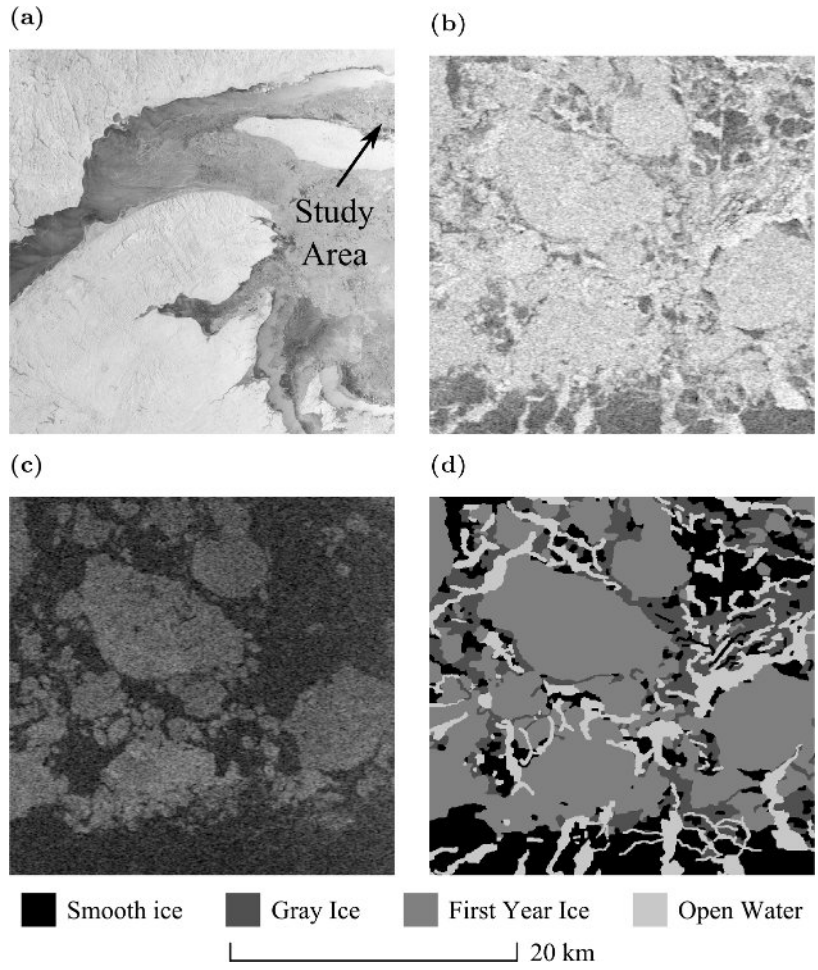

Figure 2: (a) The real dual-polarization image and its associated ground-truth image is a subscene of an area north of Anticosti Island taken from a February 25, 2008 scene of the Gulf of St. Lawrence. (b) Study area HH. (c) Study area HV. (d) Study area ground-truth. The scale and legend for subfigures (b)-(d) is also shown.

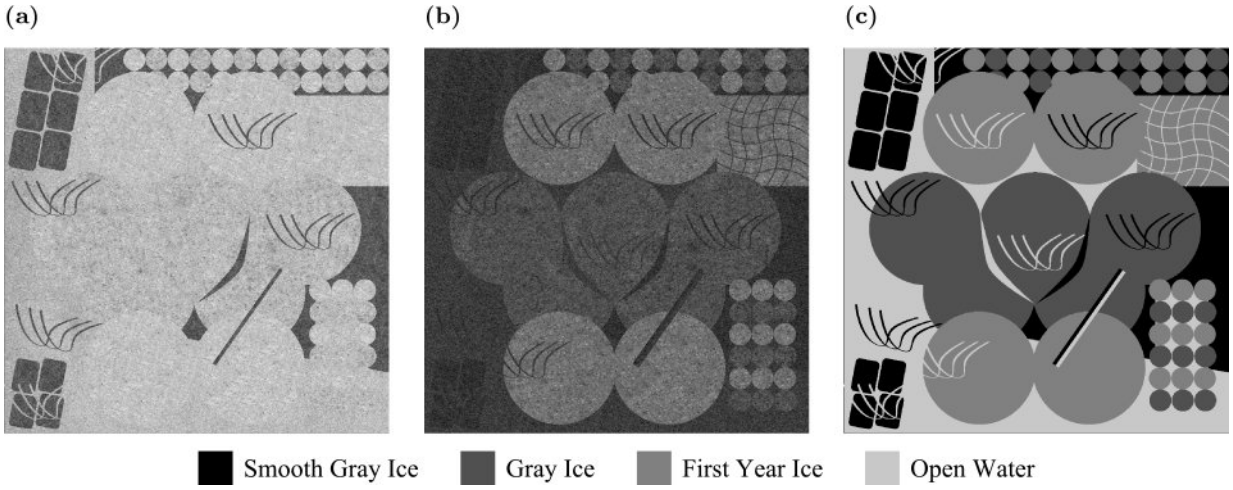

Figure 3: The synthetic dual-polarization image was synthesized from samples taken from the February 25, 2008 Gulf of St. Lawrence image. (a) Synthesized HH. (b) Synthesized HV. (c) Ground-truth.

VFG is appropriate since it minimizes noise (Lee and Cok, 1991) and properly addresses strong edges that appear in both channels.

The second alternative to the MAX rule is the Canny gradient combination rule (CG). This rule is identical to ADWA but the weight w_s^{ADWA} is replaced with:

$$w_s^{CG} = \begin{cases} 1 & \text{if site } s \text{ is a local maximum} \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

The local maximum is defined as in the Canny edge detection algorithm (Canny, 1986): a pixel 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}.

5.2. Image fusion

Two separate image fusion techniques, the HV/HH band ratio and wavelet image fusion, were tested as methods to implement Strategy 3. Wavelet image fusion is performed with several different image fusion rules. Let $D_{HH}(p)$ and $D_{HV}(p)$ represent the wavelet decomposition of the HH and HV channels, where $p = (r, c, k, l)$ indicates the wavelet coefficient in row r and column c of the k^{th} decomposition level for the l^{th} direction. 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. Let $D_F(p)$ be the wavelet decomposition of the fused image. $D_F(p)$ is created by combining the corresponding coefficients in $D_{HH}(p)$ and $D_{HV}(p)$ with fusion rules. Following (Pajares and de la Cruz, 2004), each detail coefficient of $D_F(p)$ is fused by taking the maximum of the coefficients in the corresponding position in $D_{HH}(p)$ and $D_{HV}(p)$.

Two fusion rules for the approximation coefficients are tested: the weighted-average (WA) rule, first introduced in (Burt and Kolczynski, 1993) and used for multi-polarization SAR fusion in (Simone et al., 2002), and an Absolute Difference Weighted Average (ADWA) rule that has been formulated to take advantage of differences in the HH and HV channels. The WA rule was found to give the best results in (Pajares and de la Cruz, 2004). It is formulated as follows (Burt and Kolczynski, 1993):

$$D_F(p) = w_1(p)D_{HH}(p) + w_2(p)D_{HV}(p) \quad (5)$$

where w_1 and w_2 are weights that are assigned based on a match measure between the coefficients at each location. If the match measure indicates similarity between the coefficients, the weights will average the two coefficients. If the matching is poor, the weights will choose the more salient coefficient. The match measure is the local normalized correlation for a small neighbourhood around each location while the salience calculation is a measure of the local variance in each coefficient image (Burt and Kolczynski, 1993). For this work, the match measure and salience at each location p was computed by considering a 3×3 window based on promising results in initial tests.

The ADWA rule is a simpler fusion rule. To implement ADWA, the weights $w_1(p)$ and $w_2(p)$ in Eq. 5 are set to the following:

$$w_1(p) = \frac{|D_{HH}(p) - D_{HV}(p)|}{\max_{r,c}\{|D_{HH}(p) - D_{HV}(p)|\}} \quad (6)$$

$$w_2(p) = 1 - w_1(p) \quad (7)$$

where the maximum value of $|D_{HH}(p) - D_{HV}(p)|$ is taken over all rows and columns at the particular decomposition level. 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. The reasoning for this approach comes from the characteristics of dual-polarization data. The backscatter difference between HH and HV is small for first year ice, so the ADWA rule makes the first year ice darker in the fused image by emphasizing the darker HV band. For open water at near range incidence angles, the HH band is much brighter than the HV band and will remain bright in the fused image by emphasizing the HH band. This fusion technique increases contrast between first year ice and open water in the fused image, which helps to resolve the ambiguities observed in Fig. 2 between first year ice and water.

The wavelet basis used 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. For all tests, four levels of decomposition were used as this gave the best results during initial testing. Rather than using the discrete wavelet transform (DWT), as in (Simone et al., 2002) and (Pajares and de la Cruz, 2004), the stationary wavelet transform (SWT) (Nason and Silverman, 1995) was used since it produced fewer visual artifacts in the fused images, which is a result noted in (Rockinger, 1997).

5.3. Feature space fusion with feature extraction techniques

Strategy 3 can also be implemented by feature extraction, where the original feature vectors transformed by considering the properties of the feature space. Fig. 4 shows a feature space plot of the real image (Fig. 2) with the dual-polarization HH & HV feature set. HH and HV values in DN of each feature vector are mapped to the $y^{(HH)}$ and $y^{(HV)}$ axes, respectively. 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. The feature extraction methods tested in this paper attempt to transform the 2-D feature space shown in Fig. 4 into a 1-D feature space that retains class separability. Hence, feature extraction is a form of dimensionality reduction. When applied to image data, it can also be considered a form of image fusion. Three categories of feature extraction methods are tested: principal component analysis (PCA), parabolic arc-length projection (PAL) and non-linear dimensionality reduction (NLDR).

Principal component analysis (PCA) has been used as a feature extraction technique for multichannel image data (Orlando et al., 1990; Collins and Livingstone, 1996; Yu et al., 2009; Pohl and Genderen, 1998). PCA fuses the HH and HV

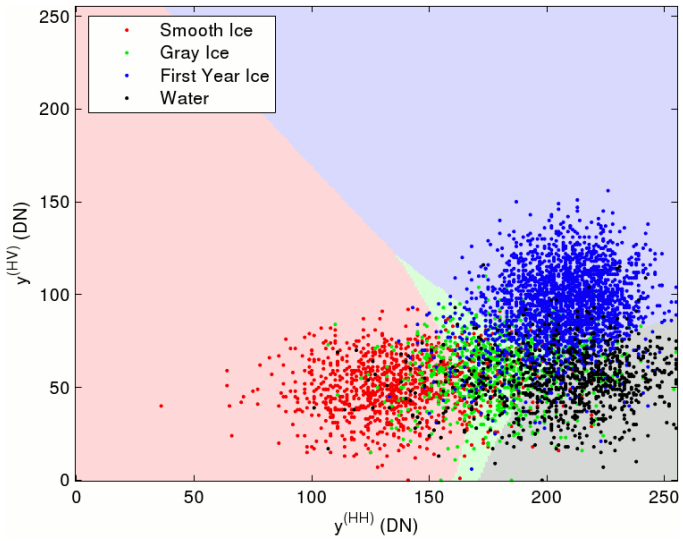

Figure 4: The HH & HV feature space plot of the real image (Fig. 2). Light shades in the background indicate the Gaussian maximum likelihood decision boundaries.

channels by projecting the 2-D feature vectors (denoted by $\mathbf{y}_s$) onto the axes in the direction of maximum variance of the data (the principal component direction) (Schowengerdt, 1997). Mathematically, the fused PCA channel for each feature vector at pixel s can be calculated as follows (Schowengerdt, 1997):

$$y_s^{PC} = \mathbf{e}^{\mathbf{T}}(\mathbf{y}_s - \bar{\mathbf{y}}) \tag{8}$$

where $\mathbf{y}_s$ is the feature vector at pixel 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 covariance matrix.

Fusion with feature extraction methods is not limited to linear projections. The feature space distribution of class clusters in Fig. 4 suggests that a simple linear projection may not maintain full class separability: the class means are not distributed along a straight line but rather a curve. This is the case for all RS-2 image data that were investigated. Projection of the points onto a coordinate system defined by a non-linear curve may be a more useful way of fusing the two channels while maintaining feature space separability. A parabolic curve in the dual-polarization feature space of Fig. 4 can be defined by:

$$y^{(HV)} = a(y^{(HH)})^2 + c \tag{9}$$

The constants a and c are found by least squares fitting Eq. 9 to the data points. Projection of each data point to the arc-length coordinate of this parabolic curve is a non-linear transformation that may be able to "unwrap" the feature space into a single dimension while maintaining class separability. The projection is accomplished as follows. Let $\mathbf{y}'_s$ be the point on the parabola closest to the feature vector being transformed ($\mathbf{y}_s$):

$$\mathbf{y}'_s = \arg\min_{\mathbf{y}_p} ||\mathbf{y}_s - \mathbf{y}_p|| \tag{10}$$

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, which results in a cubic equation that can be solved analytically (Cardano, 1545). 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 by using the standard arc-length formula (Bradley and Smith, 1995) with the quadratic curve in Equation 9:

$$y_s^{PAL} = \int_0^{y'^{(HH)}_s} \sqrt{1 + (2ay^{(HH)})^2}\, dy^{(HH)} \tag{11}$$

where a is the least-squares fitted coefficient from Equation 9, $y'^{(HH)}_s$ is the HH component of $\mathbf{y}'_s$ and y_s^{PAL} is the Parabolic Arc-Length coordinate of pixel s. The arc-length is measured from $y^{(HH)} = 0$. Equation 11 is integrated analytically.

Other families of curves can also be used; the parabolic curve is tested here as a representative of non-linear projections using curve fitting because there are analytical solutions to the above equations.

In addition to the PCA and PAL techniques, three non-linear dimensionality reduction (NLDR) techniques were also used for feature extraction: locally linear embedding (LLE) (Roweis and Saul, 2000), Laplacian eigenmaps (LEIGS) (Belkin and Niyogi, 2003) and local tangent-space alignment (LTSA) (Zhang and Zha, 2003). These techniques map the data to coordinates of a low dimensional "manifold" embedded in the higher dimensional feature space. Unlike the PAL technique, which assumes a parabolic curve or manifold, NLDR techniques learn the manifold coordinates from the data, with no assumption for a particular manifold shape. NLDR techniques are normally used for reducing the dimensionality of very high-dimension data but they are used here as non-linear generalizations of the common PCA feature extraction technique to determine whether dual-polarization data can be better fused with non-linear projections, as suggested by the feature space plot in Fig. 4.

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

$$\mathbf{y}_s = f(\mathbf{\Psi}_s) + \epsilon \tag{12}$$

where f is some unknown non-linear mapping, $\mathbf{\Psi}_s$ are the manifold coordinates of pixel s and ϵ is noise (Zhang and Zha, 2003). 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 found that preserve the local geometry for all the original feature vectors. The three methods mainly differ in the nature of the local geometry that they preserve.

LLE (Roweis and Saul, 2000) assumes that the manifold is locally linear and that each data point can be reconstructed by a linear combination of its neighbours. An optimal set of weights for this reconstruction can be found. A weight matrix

$\mathbf{W}$ is computed by minimizing the following reconstruction error $\mathcal{E}(\mathbf{W})$:

$$\mathcal{E}(\mathbf{W}) = \sum_s |\mathbf{y}_s - \sum_{j=1}^{k} W_{sj}\mathbf{y}_{s_j}|^2 \tag{13}$$

where $\mathbf{y}_{s_j}$ refers to one of the k nearest neighbours of the feature vector $\mathbf{y}_s$. 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 are invariant to translation, rotation and scaling (Roweis and Saul, 2000), the same set of weights will also optimally reconstruct the lower-dimensional manifold coordinates of each point from its neighbours in the manifold space. Therefore, choosing the $\mathbf{\Psi}$ (the set of all $\mathbf{\Psi}_s$ for every site s in the scene) to minimize the cost function $\Phi(\mathbf{\Psi})$:

$$\Phi(\mathbf{\Psi}) = \sum_s |\mathbf{\Psi}_s - \sum_{j=1}^{k} W_{sj}\mathbf{\Psi}_{s_j}|^2 \tag{14}$$

will give the desired manifold coordinates. In Equation 14, the W_{sj} are fixed to the values found in Eq. 13 and the optimization is performed to find a set of $\mathbf{\Psi}_s$ that globally minimizes the expression. 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 (Belkin and Niyogi, 2003) 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 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 lower dimensional manifold coordinates $\mathbf{\Psi}$ that minimizes the cost function $\Phi(\mathbf{\Psi})$:

$$\Phi(\mathbf{\Psi}) = \sum_{ij} (\mathbf{\Psi}_i - \mathbf{\Psi}_j)W_{ij} \tag{15}$$

where the summation is done over all pairs of nodes in the graph. The idea behind LEIGS is that points which are neighbours in the original feature space are mapped to points that are close together in the manifold coordinate space. This is ensured by the cost function while constraints described in (Belkin and Niyogi, 2003) ensure that the mapped coordinates do not collapse into a single point or a subspace with fewer dimensions than desired. The only parameter is k, the number of nearest neighbours. This value was again set to $k = 8$ after initial testing from 4 to 16 revealed little difference in the results.

LTSA (Zhang and Zha, 2003) uses the k nearest neighbours of a feature vector and uses the best fit hyperplane of those points as an estimate of the local manifold tangent. The k nearest neighbours are then converted to local tangent space coordinates $\boldsymbol{\theta}_{s_j}$, where $j = 1 \ldots k$ to indicate each of the nearest neighbour points. LTSA assumes that there is an affine transformation L_s that approximately transforms the tangent space coordinates to manifold coordinates:

$$\mathbf{\Psi}_{s_j} = \bar{\mathbf{\Psi}}_{s_j} + L_s\boldsymbol{\theta}_{s_j} + \epsilon_{s_j} \tag{16}$$

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 finds the set of L_s and $\mathbf{\Psi}_{s_j}$ that minimizes the total reconstruction error over all of the original data points:

$$\sum_s \sum_{j=1}^{k} \epsilon_{s_j} = \sum_s \sum_{j=1}^{k} \mathbf{\Psi}_{s_j} - \bar{\mathbf{\Psi}}_{s_j} - \mathbf{L}_s\boldsymbol{\theta}_{s_j} \tag{17}$$

The above problem is algebraically converted to an eigenvalue problem and the optimal manifold coordinates $\mathbf{\Psi}_s$ can be found. As with LLE and LEIGS, the only parameter that needs to be set is the number of nearest neighbours k and the same value of $k = 8$ was used.

The three NLDR techniques require solving eigenvalue problems for matrices with $(MN)^2$ entries, where M is the image width in pixels and N is the image height. 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, which is the largest number of points that could be handled without exhausting available memory (2 GB) on the test system. The 4096 feature vectors form the training set for learning the manifold. The NLDR technique is applied to the resampled image and the feature vector $\mathbf{y}_t$ at each pixel t of the resampled image will be mapped to manifold coordinates $\mathbf{\Psi}_t$.

This produces manifold coordinates only for the training feature vectors. If the training feature vectors ($\mathbf{y}_t$) are well-sampled from the underlining manifold and reasonably describe its shape, the original image feature vectors $\mathbf{y}_s$ can be mapped to coordinates on the learned manifold, using a generic method to estimate the manifold coordinates of feature vectors based on a training subset (Li et al., 2005). The affine transform $\mathbf{L}_t$ that maps each HH-HV feature vector in the training set to its manifold coordinate is found and feature vectors from the original image are mapped to manifold coordinates by using the $\mathbf{L}_t$ of the nearest (in the HH-HV feature space) training feature vector.

The resulting NLDR feature vectors in our implementation have two dimensions. The final 1D manifold coordinates are obtained by discarding the second dimension, which does not preserve as much of the local geometry of the original feature space. These 1D manifold coordinates represent the desired fused image that is input into MIRGS. MIRGS was also tested with the 2D NLDR feature vectors as a multichannel image since the transformed feature space may be incidentally beneficial for segmentation accuracy. The gradient for this multichannel image was computed with the VFG gradient technique for use with MIRGS. For completeness, the second dimension of the NLDR feature vectors was also tested as a univariate image.

The implementation of each of the three NLDR methods was obtained from (Wittman, 2005), which implements LLE, LEIGS and LTSA as described in (Roweis and Saul, 2000), (Belkin and Niyogi, 2003) and (Zhang and Zha, 2003), respectively.

5.4. Experiments

Each method described in Sections 5.1 to 5.3 was tested on the two images mentioned in Section 4. Each technique was applied to the calibrated dual-polarization images and MIRGS was used to obtain a segmentation result. The segmentation result is then compared to the ground-truth image. Two measures (Qin and Clausi, 2010) were used for measuring segmentation accuracy: the overall accuracy (Acc.), which is the percentage of pixels correctly segmented and the κ coefficient (Bishop et al., 1975; Congalton et al., 1983). The κ coefficient is defined as follows (Richards and Jia, 2006):

$$\kappa = \frac{P\sum_k x_{kk} - \sum_k x_{k+}x_{+k}}{P^2 - \sum_k x_{k+}x_{+k}} \tag{18}$$

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 a accuracy assessment measure that ranges from $[-1, 1]$ which 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.

6. Results

Table 2 shows the accuracy assessment results for the real and synthetic images (Figs. 2 and 3) for all methods. The MIRGS parameter C_1 (Qin and Clausi, 2010) that gave the best results is also listed. C_1 controls the strength of the spatial context model, with larger C_1 implying a greater reliance on spatial context rather than the feature model. In general terms, a larger C_1 leads to greater merging and smoother segmentation results. The entries identified as 'HH & HV' indicate methods where the two bands are used directly with MIRGS. There are four such entries corresponding to each method of gradient combination. HH & HV with VFG gradient is the basic MIRGS algorithm (Qin and Clausi, 2010) applied to dual-polarization data. Each NLDR technique is listed three times with suffixes '1', '1 & 2' and '2'. The '1' corresponds to the fused image formed by retaining only the first NLDR dimension. The '1 & 2' corresponds to the multivariate feature set consisting of both NLDR dimensions. The '2' corresponds to the image formed by retaining the second NLDR dimension. The acronyms used in this section are summarized in Table 1.

In Table 2, the feature sets that resulted in the highest accuracy are denoted by a $\star$ symbol. The highest accuracy feature set and those which are comparable to the most accurate results are listed in bold text. The grouping of results into those that are comparable to the best results was done by visual assessment and by overall accuracy, rather than by considering statistical significance. It was found that *any* two results for the same image had a statistically significant difference in accuracy according to McNemar's test (Foody, 2004), even for results that differ very little in accuracy or κ. This was due to the large number of pixels for each class in the image. Since the groupings amongst the results were very obvious qualitatively, statistical significance was not used.

For the real RS-2 image, the best accuracies were obtained with multivariate feature sets (HH & HV and LLE 1 & 2). All of these performed 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. MIRGS is not very sensitive to the gradient combination method and all four gradient generation methods produced very similar results. Many of the image fusion and feature extraction techniques (ADWA Wavelet, WA Wavelet, LEIGS 1, LLE 1) produced accuracy results better than the best HH or HV channels alone. This was expected as they each attempt to fuse the dual-polarization information into a single band but the fused images were unable to match the accuracy of the dual-polarization feature set HH & HV, indicating that some information had been lost after fusion. The best results obtained with LTSA 1, PAL and PCA were unable to improve upon the best HH results. The second dimension obtained by the NLDR techniques generally performed worse than the first dimension, which is expected because it does not preserve as much of the local geometry of the original feature space. The only exception is LTSA 2, which performed slightly better than LTSA 1 for the real image. However, its performance is still amongst the worst obtained for the real image.

All four of the segmentation results obtained for the synthetic image using the HH & HV data had similar accuracy, with the MAX gradient performing best amongst the four. The closeness of the accuracy values again shows that MIRGS is not sensitive to the gradient combination method, which is agrees with the results for the real RS-2 image. 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. The performance of the single polarization channel HH is affected by the low feature space class separability between 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. The single channel images created by using NLDR techniques (LEIGS 1, LLE 1, LTSA 1), PCA, PAL and HV / HH all provide better accuracy than the individual HH or HV channels. In fact, the univariate NLDR feature sets 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 was designed to take advantage of the high backscatter level of open water compared to first year ice in the HH channel to increase image contrast. However, at mid-range incidence angles, the backscatter of open water is closer to first year ice so the ADWA fusion does not perform as well. In general, the segmentation accuracies obtained from the synthetic image were higher because there are fewer thin or small details in the synthetic image.

Considering the results from both the real RS-2 image and the synthetic image, HH & HV with any of the gradient combination rules and LLE 1 & 2 consistently provide the best or near the best performance, with very little difference in accuracy or visual appearance of the segmentation. The multivariate outputs from the NLDR transforms do not perform

Table 1: Summary of acronyms and naming conventions for the image fusion and feature extraction methods tested.

Gradient Generation Methods	
ADWA Gradient	Gradient based on Eq. 3.
VFG Gradient	Existing VFG gradient (Lee and Cok, 1991).
MAX Gradient	Gradient based on Eq. 1.
CG Gradient	Gradient based on Eq. 4.
Image Fusion Techniques	
ADWA Wavelet	Wavelet fusion using Eq. 6.
WA Wavelet	Wavelet fusion using Eq. 5.
HV / HH	Band ratio image.
Feature Extraction Techniques	
LEIGS	Laplacian Eigenmaps (Belkin and Niyogi, 2003).
LLE	Locally Linear Embedding (Roweis and Saul, 2000).
LTSA	Local Tangent Space Alignment (Zhang and Zha, 2003).
PAL	Fusion via Eq. 11.
PCA	Principal Component Analysis (Schowengerdt, 1997).
Suffixes following LEIGS, LLE & LTSA, e.g. 'LEIGS 1 & 2'	
1	First channel from dimensionality reduction.
1 & 2	Both channels from dimensionality reduction.
2	Second channel from dimensionality reduction.

Table 2: Segmentation overall accuracy (Acc.) and κ coefficient for each of the tested methods on the (a) real and (b) synthetic images. The number of bands input into MIRGS is shown to distinguish between multivariate and univariate methods. The MIRGS parameter C_1 that gave the best results is also listed. Unless noted in the method name, VFG was used to compute the image gradient for all multivariate techniques.

(a) Real image

Method	Bands	Acc. (%)	κ	C1
ADWA Wavelet	1	75.24	0.62	5
HH	1	72.14	0.57	7
HH & HV, ADWA	**2**	**83.92**	**0.77**	**3**
HH & HV, VFG⋆	**2**	**84.70**	**0.78**	**3**
HH & HV, MAX	**2**	**84.19**	**0.77**	**3**
HH & HV, CG	**2**	**83.92**	**0.77**	**3**
HV	1	48.37	0.30	5
LEIGS 1	1	73.91	0.64	7
LEIGS 1 & 2	2	74.97	0.65	7
LEIGS 2	1	60.52	0.46	7
LLE 1	1	75.59	0.66	5
LLE 1 & 2	**2**	**84.13**	**0.77**	**3**
LLE 2	1	61.51	0.42	5
LTSA 1	1	49.30	0.31	5
LTSA 1 & 2	2	62.88	0.50	5
LTSA 2	1	50.38	0.29	5
PAL	1	68.78	0.55	7
PCA	1	70.55	0.58	7
WA Wavelet	1	72.52	0.59	3
HV / HH	1	57.51	0.41	5

(b) Synthetic image

Method	Bands	Acc. (%)	κ	C1
ADWA Wavelet	1	57.12	0.38	5
HH	1	78.87	0.70	7
HH & HV, ADWA	**2**	**98.08**	**0.97**	**3**
HH & HV, VFG	**2**	**98.25**	**0.98**	**3**
HH & HV, MAX	**2**	**98.28**	**0.98**	**3**
HH & HV, CG	**2**	**98.09**	**0.97**	**3**
HV	1	83.14	0.76	5
LEIGS 1	1	95.87	0.94	5
LEIGS 1 & 2	2	95.78	0.94	5
LEIGS 2	1	82.96	0.76	7
LLE 1	1	97.32	0.96	5
LLE 1 & 2	**2**	**98.21**	**0.98**	**3**
LLE 2	1	94.12	0.92	5
LTSA 1	1	93.91	0.92	3
LTSA 1 & 2⋆	**2**	**98.40**	**0.98**	**3**
LTSA 2	1	83.40	0.77	5
PAL	1	86.59	0.81	7
PCA	1	88.44	0.84	5
WA Wavelet	1	58.92	0.41	5
HV / HH	1	93.61	0.91	7

⋆ Best result for given C_1

Results comparable to best result are listed in **bold**.

better than HH & HV for either image. Although LTSA 1 & 2 provides the best accuracy for the synthetic image, its performance is close to the HH & HV results and it performs poorly on the real image. The inconsistency in performance may be related to the fact that LTSA outputs different results for the same image when different sets of training feature vectors are used to learn the manifold (Yu, 2009). The high performance of LTSA 1 & 2 for the synthetic image is thus purely incidental. Among the univariate fused images obtained from NLDR techniques, LLE 1 consistently achieves the highest univariate accuracy for both synthetic and real images. LTSA 1 only performs well for the synthetic image and LEIGS 1 is not particularly noteworthy for either image. Since only LLE is applicable for both images tested, it appears to be the best NLDR technique for the purposes of fusing RS-2 data. However, neither the multivariate nor univariate output from LLE improve upon the results obtained by using the HH & HV feature set directly in MIRGS.

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 RS-2 SAR sea ice image segmentation. The above observations indicate that the basic MIRGS algorithm with HH & HV data remains the best method for using the dual-polarization data. Due to the lack of accuracy difference between the various types of gradient calculation methods, there is no motivation for replacing VFG with any of the other tested methods.

Figs. 5 and 6 compare the segmentation results using HH data only and using the HH & HV data (with VFG gradient) for both the real image and the synthetic image. The dual-polarization clearly improves the discrimination of the four ice types as shown by the images and the accuracy statistics in Table 2.

To obtain accurate results with only the HH channel, the spatial context model had to be weighted highly with a large C_1 value in MIRGS since the HH channel does not provide sufficient feature space class separability for the different ice types. This illustrates the role of the C_1 parameter: poor separability between classes leads to a noisy segmentation where some regions are incorrectly labeled. This can be partially compensated by an increased C_1 value, which gives greater emphasis to the spatial context model. The spatial context model considers adjacent regions to be more likely to belong to the same class, which can correct erroneously labeled regions as long as the adjacent regions are correctly labeled. However, the increased spatial context weighting manifests itself as a loss of detail in the segmentation results. Previous work (Yu, 2009) has explored the range of accuracy that can be obtained by changing the C_1 parameter. In general, the accuracy decreases smoothly as C_1 is adjusted away from the value that gives the peak accuracy for each type of image; large C_1 values caused over-smoothing while small C_1 values led to noisy segmentations. As this paper is focused on comparing the best attainable results between different image fusion or feature extraction techniques, only the results from the C_1 values giving highest accuracy are reported.

7. Conclusion

The use of dual-polarization RADARSAT-2 SAR sea ice imagery, which will be used for operational sea ice mapping, to improve sea ice segmentation has been investigated in this paper. Experiments with both a real and a synthetic dual-polarization image generated from real sea ice imagery were performed to determine the best strategy for utilizing the dual-polarization information. Several image fusion and feature extraction schemes for the dual-polarization data were investigated and compared with using the dual-polarization data directly in the standard MIRGS segmentation algorithm. The tested methods included classic feature extraction 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.

According to the experiments, the best strategy for dual-polarization data is Strategy 1, the standard MIRGS algorithm with no changes. Gradient combination rules (Strategy 2) had very little effect on overall accuracy, while feature extraction and image fusion approaches (Strategy 3) did not retain all the feature space class separability information when the two channels were combined into one channel. The best studied fusion technique (whether image fusion or feature extraction) was an NLDR technique called locally linear embedding (LLE) which consistently produced the best single channel image segmentation results from the dual-polarization data. However, it was still unable to match the segmentation performance of the original HH & HV feature set. In terms of the benefits of adding dual-polarization information for sea ice segmentation, the experiments have shown that it can substantially improve the segmentation accuracy over segmentation with only single-polarization imagery. 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 is useful for human interpretation.

The work here has not considered the incidence angle variation of the appearance of various ice types. Methods to make use of dual-polarization data to address this problem should be investigated in future work. As additional dual-polarization sea ice data sets with ground-truth becomes available, the experimental framework established in this paper could be used to draw additional conclusions about the performance of data fusion for sea ice from different geographical locations and dates. A similar investigation of feature extraction and image fusion methods for other applications such as agriculture or forestry could also be performed.

8. Acknowledgment

The authors would like to thank Canadian Ice Service staff for data and advice. The Natural Sciences and Engineering Research Council of Canada (NSERC) and the Canadian Federal Government's International Polar Year are thanked for financially supporting this project. RADARSAT is an official mark of the Canadian Space Agency.

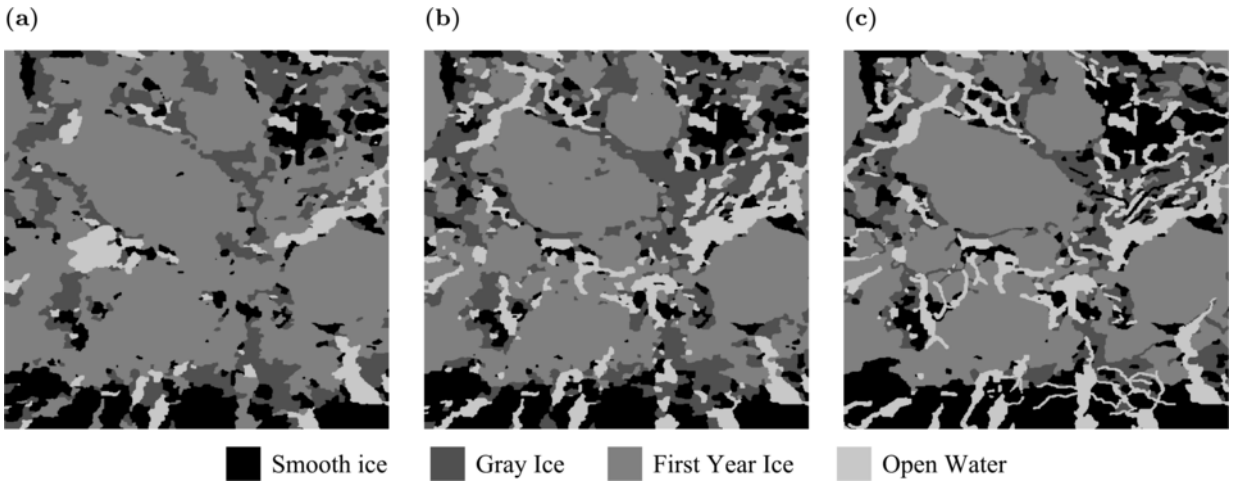

Figure 5: Comparison of real image segmentation results obtained from (a) HH alone and from (b) HH & HV with VFG gradient. (c) Ground-truth.

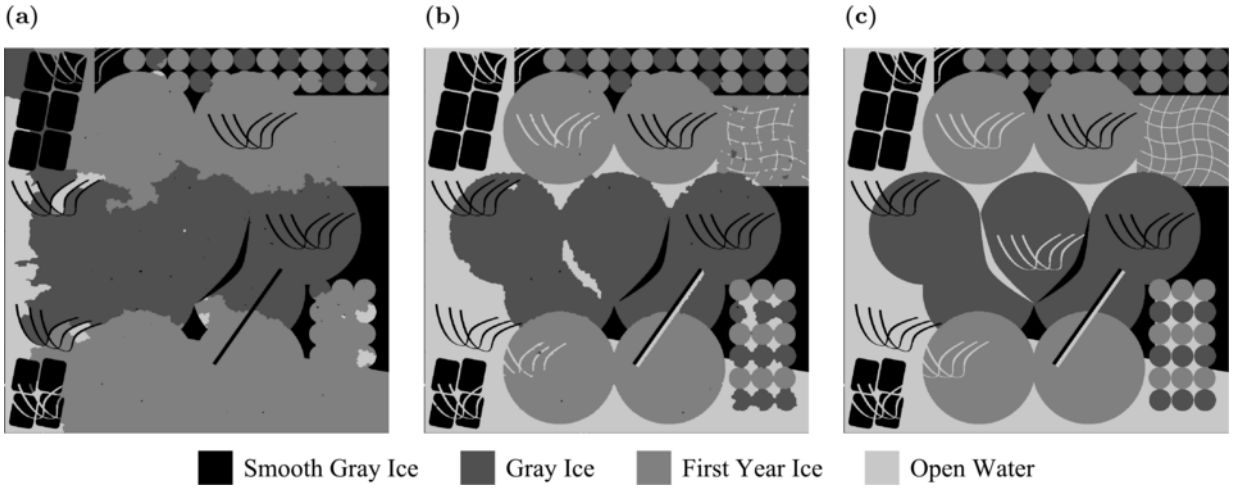

Figure 6: Comparison of synthetic image segmentation results obtained from (a) HH alone and from (b) HH & HV with VFG gradient. (c) Ground-truth.

References

Barbosa, I. S., Maillard, P., 2010. Mapping a wetland complex in the Brazilian savannah using an Ikonos image: assessing the potential of a new region-based classifier. Canadian Journal of Remote Sensing 36 (S2), S231–S242.

Belkin, M., Niyogi, P., Mar. 2003. Laplacian eigenmaps for dimensionality reduction and data representation. Neural Computation 15 (6), 1373–1396.

Bishop, T., Fienberg, S., Holland, P., 1975. Discrete Multivariate Analysis - Theory and Practice. MIT Press, Cambridge.

Bouvet, A., Toan, T. L., Lam-Dao, N., Feb. 2009. Monitoring of the rice cropping system in the Mekong Delta using ENVISAT/ASAR dual polarization data. IEEE Transactions on Geoscience and Remote Sensing 47 (2), 517 – 526.

Bradley, G. L., Smith, K. J., 1995. Calculus. Prentice Hall, New Jersey.

Burt, P. J., Kolczynski, R. J., May 1993. Enhanced image capture through fusion. In: Proceedings of the Fourth International Conference on Computer Vision. Berlin, pp. 173 – 182.

Canny, J., Nov. 1986. A computational approach to edge detection. IEEE Transactions on Pattern Analysis and Machine Intelligence 8 (6), 679 – 698.

Cardano, G., 1545. Artis magnæ, sive de regulis algebraicis liber unus. Nuremberg.

Clausi, D. A., Qin, A., Chowdhury, M., Yu, P., Malliard, P., 2010. MAGIC: MAp-Guided Ice Classification System. Canadian Journal of Remote Sensing 36 (S1), S13 – S25.

Collins, M. J., Livingstone, C. E., Jan. 1996. On the dimensionality of multiparameter microwave image data from the thin sea ice in the Labrador Sea. IEEE Transactions on Geoscience and Remote Sensing 34 (1), 114 – 136.

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

De Abreu, R., Flett, D., Scheuchl, B., Ramsay, B., Jul. 2003. Operational sea ice monitoring with RADARSAT-2 - a glimpse into the future. In: Proceedings of the IEEE International Geoscience and Remote Sensing Symposium. Vol. 2. pp. 1308 – 1310.

Derin, H., Elliott, H., Jan. 1987. Modeling and segmentation of noisy and textured images using Gibbs random fields. IEEE Transactions on Pattern Analysis and Machine Intelligence 9 (1), 39 – 55.

Drinkwater, M. R., Sep. 1989. LIMEX 87 ice surface characteristics: Implications for C-Band SAR backscatter signatures. IEEE Transactions on Geoscience and Remote Sensing 27 (5), 501–513.

Duda, R., Hart, P., Stork, D., 2001. Pattern Classification, 2nd Edition. Wiley, New York.

Flett, D. G., Sep. 2003. Operational use of SAR at the Canadian Ice Service: present operations and a look to the future. In: Proceedings of the 2nd Workshop on Coastal and Marine Applications of SAR. Svalbard, Norway.

Foody, G. M., Nov. 2004. Thematic map comparison: Evaluating the statistical significance of differences in classification accuracy. Photogrammetric Engineering & Remote Sensing 70 (5), 627 – 633.

Geman, S., Geman, D., Nov. 1984. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. IEEE Transactions on Pattern Analysis and Machine Intelligence 6 (6), 721 – 741.

Hong, S., Moon, W. M., Paik, H.-Y., Choi, G.-H., Nov. 2002. Data fusion of multiple polarimetric SAR images using Discrete Wavelet Transform (DWT). In: Proceedings of the IEEE International Geoscience and Remote Sensing Symposium. Vol. 6. pp. 3323 – 3325.

Jin, Y., Ruliang, Y., Ruohong, H., Oct. 2006. Pixel level fusion for multiple SAR images using PCA and wavelet transform. In: International Conference on Radar, CIE '06. Shanghai, China.

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

Lee, H.-C., Cok, D. R., May 1991. Detecting boundaries in a vector field. IEEE Transactions on Signal Processing 39 (5), 1181 – 1194.

Li, H., Teng, L., Chen, W., Shen, I.-F., 2005. Supervised learning on local tangent space. In: Advances in Neural Networks - ISNN 2005. Vol. 3496 of Lecture Notes in Computer Science. Springer-Verlag, pp. 546 – 551.

Li, S. Z., 2001. Markov random field modeling in image analysis, 2nd Edition. Springer-Verlag, Tokyo, Japan.

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

Manore, M., Flett, D. G., De Abreu, R. A., Ramsay, B. R., van der Sanden, J., Jul. 2001. Multi-polarization SAR data for operational ice monitoring. In: Proceedings of the IEEE International Geoscience and Remote Sensing Symposium. Vol. 3. Sydney, Australia, pp. 1246 – 1248.

Nakamura, K., Wakabayashi, H., Naoki, K., Nishio, F., Moriyama, T., Uratsuka, S., Nov. 2005. Observation of sea-ice thickness in the Sea of Okhotsk by using dual-frequency and fully polarimetric airborne SAR (Pi-SAR) data. IEEE Transactions on Geoscience and Remote Sensing 43 (11), 2460 – 2469.

Nason, G. P., Silverman, B. W., 1995. The stationary wavelet transform and some statistical applications. In: Antoniadis, A., Oppenheim, G. (Eds.), Wavelets and Statistics. Vol. 103 of Lecture Notes In Statistics. Springer-Verlag, pp. 281–300.

Ochilov, S., , Clausi, D. A., Jun. 2010. Automated classification of operational SAR sea ice images. In: 6th Canadian Conference on Computer and Robotic Vision. Ottawa, ON, Canada.

Orlando, J. R., Mann, R., Haykin, S., Jul. 1990. Classification of sea-ice images using a dual-polarized radar. IEEE Journal of Oceanic Engineering 15 (3), 228 – 237.

Pajares, G., de la Cruz, J. M., Sep. 2004. A wavelet-based image fusion tutorial. Pattern Recognition 37 (9), 1855 – 1872.

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

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

Qin, A. K., Clausi, D. A., Aug. 2010. Multivariate image segmentation using semantic region growing with adaptive edge penalty. IEEE Transactions on Image Processing 8 (19).

Ramsay, B., Flett, D., Andersen, H. S., Gill, R., Nghiem, S., Bertoia, C., Jun. 2004. Preparation for the operational use of RADARSAT-2 for ice monitoring. Canadian Journal of Remote Sensing 30 (3), 415 – 423.

Richards, J. A., Jia, X., 2006. Remote Sensing Digital Image Analysis, 4th Edition. Springer, Berlin.

Rockinger, O., Oct. 1997. Image sequence fusion using a shift-invariant wavelet transform. In: Proceedings of the IEEE International Conference on Image Processing. Vol. 3. Santa Barbara, CA, USA, pp. 288 – 291.

Roweis, S., Saul, L., Dec. 2000. Nonlinear dimensionality reduction by locally linear embedding. Science 22 (5500), 2323–2336.

Scheuchl, B., Caves, R., Flett, D., De Abreu, R., Arkett, M., Cumming, I., Sep. 2004a. The potential of cross-polarization information for operational sea ice monitoring. In: Proceedings of the Envisat & ERS Symposium. Salzburg, Austria.

Scheuchl, B., Flett, D., Caves, R., Cumming, I., Jun. 2004b. Potential of RADARSAT-2 data for operational sea ice monitoring. Canadian Journal of Remote Sensing 30 (3), 448 – 461.

Schowengerdt, R., 1997. Remote Sensing: Models and Methods for Image Processing., 2nd Edition. Academic Press, San Diego.

Simone, G., Farina, A., Morabito, F., Serpico, S., Bruzzone, L., Mar. 2002. Image fusion techniques for remote sensing applications. Information Fusion 3 (1), 3 – 15.

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

Vincent, L., Soille, P., Jun. 1991. Watersheds in digital spaces: An efficient algorithm based on immersion simulations. IEEE Transactions on Pattern Analysis and Machine Intelligence 13 (6), 583 – 598.

Wilson, K. J., Falkingham, J., Melling, H., De Abreu, R., Sep. 2004. Shipping in the Canadian Arctic: other possible climate change scenarios. In: Proceedings of the IEEE International Geoscience and Remote Sensing Symposium. Vol. 3. pp. 1853 – 1856.

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

Wong, A., Yu, P., Zhang, W., Clausi, D. A., Apr. 2010. IceSynth II: Synthesis of SAR sea-ice imagery using region-based posterior sampling. IEEE Geoscience and Remote Sensing Letters 7 (2), 348 – 351.

Yang, S., Wang, M., Lu, Y. X., Qi, W., Jiao, L., Dec. 2009. Fusion of multiparametric SAR images based on SW-nonsubsampled contourlet

and PCNN. Signal Processing 89 (12), 2596 – 2608.
Yu, P., 2009. Segmentation of RADARSAT-2 dual-polarization sea ice imagery. Master's thesis, University of Waterloo.
Yu, P., Clausi, D. A., Howell, S. E. L., Jul. 2009. Fusing AMSR-E and QuikSCAT imagery for improved sea ice recognition. IEEE Transactions on Geoscience and Remote Sensing 47 (7), 1980–1989.
Yu, Q., Clausi, D. A., December 2007. SAR sea-ice image analysis based on iterative region growing using semantics. IEEE Transactions on Geoscience and Remote Sensing 45 (12), 3919 – 3931.
Yu, Q., Clausi, D. A., December 2008. IRGS: Image segmentation using edge penalties and region growing. IEEE Transactions on Pattern Analysis and Machine Intelligence 30 (12), 2126 – 2139.
Zhang, X., Huang, P., Zhou, P., Nov. 2007. Data fusion of multiple polarimetric SAR images based on combined curvelet and wavelet transform. In: Proceedings of the 1st Asian and Pacific Conference on Synthetic Aperture Radar. Huangshan, China, pp. 225 – 228.
Zhang, Z., Blum, R. S., Aug. 1999. A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application. Proceedings of the IEEE 87 (8), 1315 – 1326.
Zhang, Z., Zha, H., Aug. 2003. Nonlinear dimension reduction via local tangent space alignment. In: Intelligent Data Engineering and Automated Learning. Vol. 2690 of Lecture Notes in Computer Science. Springer.