Segmentation and Classification of Polarimetric SAR Data Using Spectral Graph Partitioning

by

KAAN ERSAHIN

B.Sc., Istanbul Technical University, 1999
M.Sc., Istanbul Technical University, 2001

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY in

The Faculty of Graduate Studies
(Electrical and Computer Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)
November 2009
© KAAN ERSAHIN 2009
Abstract

Polarimetric Synthetic Aperture Radar (POLSAR) data have been commercially available for the last few years, which has increased demand for its operational use in remote sensing applications. Segmentation and classification of image data are important tasks for POLSAR data analysis and interpretation, which often requires human interaction.

Existing strategies for automated POLSAR data analysis have utilized the polarimetric attributes of pixels, which involve target decompositions based on physical, mathematical or statistical models. A well-established and widely-used technique is the Wishart classifier, which is used as the benchmark in this work.

In this thesis, a new methodology is used that exploits both the polarimetric attributes of pixels, and the visual aspect of the image data through computer vision principles. In this process, the performance level of humans is desired, and several features or cues, inspired by perceptual organization, are utilized, i.e., patch-based similarity of intensity, contour, spatial proximity, and the polarimetric cue. The pair-wise grouping technique of Spectral Graph Partitioning (SGP) is employed to perform the segmentation and classification tasks based on graph cuts.

A new classification algorithm is developed for POLSAR data, where segmentation based on the contour and spatial proximity cues is followed by classification based on the polarimetric cue (i.e., similarity of coherency matrices). It offers a way to utilize the complete polarimetric information through the coherency matrix representation in the SGP framework.

The proposed unsupervised technique aims to automate the data analysis process for the mapping of distributed targets. Two fully polarimetric data sets in L-, and C-bands acquired by AIRSAR and the Convair-580, both
Abstract

containing agricultural fields, were used to obtain the experimental results and analysis. The results suggest quantitative and qualitative improvements over the Wishart classifier. This method is suitable for applications where homogeneity within each separated region is desirable, such as mapping crops or other types of terrain.

The SGP methodology used in the developed scheme is flexible in the definition of affinity functions and will likely allow further improvements through the addition of different image features and data sources.
# Table of Contents

**Abstract** ................................................................. ii

**Table of Contents** ....................................................... iv

**List of Tables** ............................................................ vii

**List of Figures** ............................................................ viii

**List of Symbols** ........................................................... xi

**Acknowledgements** ........................................................ xiv

**1 Introduction** ............................................................. 1

1.1 Background Technology .................................................. 3

1.1.1 Statistical Models for Polarimetric SAR Data .................. 4

1.1.2 Speckle Reduction Techniques for Polarimetric SAR .......... 6

1.1.3 Feature Extraction, Classification and Segmentation

1.1.4 Polarimetric SAR Data and Texture .............................. 10

1.1.5 Emerging Techniques in Computer Vision for Group-

1.2 Thesis Outline and Objective ........................................... 12

1.3 Thesis Outline and Objective ........................................... 14

**2 Spectral Graph Partitioning** ........................................... 18

2.1 Graph Partitioning with Normalized Cuts .......................... 19

2.1.1 Spectral Clustering Algorithm .................................... 21

2.1.2 Multiclass Spectral Clustering Algorithm ....................... 23
# Table of Contents

2.2 Techniques for Reduced Computational Complexity

2.2.1 Iterative Methods

2.2.2 Sparse Representation

2.2.3 The Nyström Extension Method

2.2.4 Other Techniques for $O(N^2)$ Problems

2.3 Implementation and Runtime Analysis of the Algorithms

2.4 Summary

3 Cues for Polarimetric SAR Data

3.1 Patch-based Similarity: Intensity Cue

3.2 Contour-based Similarity Cue

3.3 Similarity of Coherency Matrices: Polarimetric Cue

3.3.1 Bartlett Distance

3.3.2 Symmetric Revised Wishart Distance

3.4 Spatial Proximity Cue

3.4.1 Ramp Function

3.4.2 Step Function

3.5 Summary

4 The Proposed Method

4.1 Preliminary Research

4.1.1 Scheme S1: Patch-based Similarity & Spatial Proximity

4.1.2 Scheme S2: Contour-based Similarity & Spatial Proximity

4.1.3 Scheme S3: Contour and Spatial Proximity Followed by Intensity Cue

4.2 The Proposed Method: Contour and Spatial Proximity Followed by Polarimetric Cue

5 Experimental Results, Analysis and Discussion

5.1 Data Sets

5.1.1 Westham Island, BC, Canada

5.1.2 Flevoland, The Netherlands
# Table of Contents

5.2 Results and Analysis for Scheme S1 .................................................. 64
5.3 Results and Analysis for Scheme S2 .................................................. 66
   5.3.1 Utilizing Three Intensity Channels ............................................. 66
   5.3.2 Utilizing Co-polarized Correlation Coefficient ............................ 68
   5.3.3 Analysis for Different Number of Partitions ................................ 70
5.4 Results and Analysis for Scheme S3 .................................................. 86
5.5 Results and Analysis for the Proposed Method ................................. 88
   5.5.1 Convair-580: Westham Island Dataset ....................................... 89
   5.5.2 AIRSAR: Flevoland Dataset .................................................... 97

6 Conclusion ............................................................................................... 99
   6.1 Summary ........................................................................................ 99
   6.2 Research Contributions ................................................................. 103
   6.3 Future Work .................................................................................. 105

Bibliography .............................................................................................. 106

Appendices

A Representations of Polarimetric SAR Data ............................................. 115
   A.1 Scattering Matrix ....................................................................... 115
   A.2 Covariance and Coherency Matrices .......................................... 116

B Test Schemes Using the Polarimetric Cue ............................................ 117
   B.1 Scheme S4: Polarimetric Cue & Spatial Proximity Cue .................. 117
      B.1.1 Results and Analysis for Scheme S4 ...................................... 118
   B.2 Scheme S5: Polarimetric Cue & Contour Cue & Spatial Proximity
      Cue ................................................................................................. 121
      B.2.1 Results and Analysis for Scheme S5 ...................................... 123
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1</td>
<td>Overall classification accuracy (in %) obtained by S1 for patch size, $d \in {11, 15, 21}$, and local scaling neighborhood size, $n_{LS} \in {5, 25}$ for two scenarios: with and without spatial proximity cue</td>
<td>64</td>
</tr>
<tr>
<td>5.2</td>
<td>Overall classification accuracy (in %) obtained by S2 for elongation, $\lambda^2 \in {5, 9}$, and number of orientations, $ori \in {4, 6}$ for Test Regions # 1 and # 2 of the Westham Island data</td>
<td>68</td>
</tr>
<tr>
<td>5.3</td>
<td>Overall accuracy and Kappa coefficient obtained by the proposed method for Test Region # 1. $\lambda^2 \in {3, 5, 9}$; $ori \in {4, 6}$; $\sigma \in {1, 2}$</td>
<td>90</td>
</tr>
<tr>
<td>5.4</td>
<td>Overall accuracy and Kappa coefficient obtained by the proposed method for Test Region # 2. $\lambda^2 \in {3, 5, 9}$; $ori \in {4, 6}$; $\sigma \in {1, 2}$</td>
<td>90</td>
</tr>
</tbody>
</table>
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Three-circle data: K-means vs. Spectral Clustering</td>
<td>23</td>
</tr>
<tr>
<td>2.2</td>
<td>Three-circle data: Spectral Clustering with local scaling</td>
<td>24</td>
</tr>
<tr>
<td>3.1</td>
<td>Rotated copies of the filter used in OE computation</td>
<td>34</td>
</tr>
<tr>
<td>3.2</td>
<td>Orientation energy as a measure of contour information</td>
<td>35</td>
</tr>
<tr>
<td>4.1</td>
<td>Progression of research ideas over several years</td>
<td>44</td>
</tr>
<tr>
<td>4.2</td>
<td>Block diagram for Scheme S1</td>
<td>45</td>
</tr>
<tr>
<td>4.3</td>
<td>Block diagram for Scheme S2</td>
<td>47</td>
</tr>
<tr>
<td>4.4</td>
<td>Block diagram for Scheme S3</td>
<td>48</td>
</tr>
<tr>
<td>4.5</td>
<td>Block diagram for the proposed method.</td>
<td>51</td>
</tr>
<tr>
<td>5.1</td>
<td>Westham Island Scene - Color Composite (HH-HV-VV)</td>
<td>59</td>
</tr>
<tr>
<td>5.2</td>
<td>Westham Island scene: color composite, ground truth, and</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Wishart classifier results for Test Regions # 1, # 2, and # 3</td>
<td>60</td>
</tr>
<tr>
<td>5.3</td>
<td>Westham Island Scene - Ground truth and field photographs</td>
<td>61</td>
</tr>
<tr>
<td>5.4</td>
<td>Flevoland L-band data: color composite (750 × 1024 pixels)</td>
<td>62</td>
</tr>
<tr>
<td>5.5</td>
<td>Flevoland data subset with ground truth information</td>
<td>63</td>
</tr>
<tr>
<td>5.6</td>
<td>S1 classification results of ROI for different patch sizes</td>
<td>65</td>
</tr>
<tr>
<td>5.7</td>
<td>S1 classification result and the first $k$ eigenvectors of $L$</td>
<td>66</td>
</tr>
<tr>
<td>5.8</td>
<td>S2 Results: Westham Island - Test Region # 1</td>
<td>67</td>
</tr>
<tr>
<td>5.9</td>
<td>S2 Results: Westham Island - Test Region # 2</td>
<td>68</td>
</tr>
<tr>
<td>5.10</td>
<td>S2 Results: Region # 1. (a) with filtering, $n_b = 3$; (b) with filtering, $n_b = 4$; (c) without filtering, $n_b = 3$; (d) without filtering, $n_b = 4$</td>
<td>69</td>
</tr>
</tbody>
</table>
List of Figures

5.11 S2 Results: Region # 2. (a) with filtering, \( n_b = 3 \); (b) with filtering, \( n_b = 4 \); (c) without filtering, \( n_b = 3 \); (d) without filtering, \( n_b = 4 \) .................................................. 69

5.12 S2 Results: Region # 3. (a) with filtering, \( n_b = 3 \); (b) with filtering, \( n_b = 4 \); (c) without filtering, \( n_b = 3 \); (d) without filtering, \( n_b = 4 \) .................................................. 70

5.13 Test Region # 1: before and after speckle filtering ............... 71

5.14 Test Region # 1: partitioning with individual inputs ............. 71

5.15 Test Region # 1: partitioning with combined inputs ............. 72

5.16 Test Region # 1: individual inputs after filtering .............. 73

5.17 Test Region # 1: combined inputs after filtering .............. 73

5.18 Test Region # 2: before and after speckle filtering ............ 74

5.19 Test Region # 2: partitioning with individual inputs ........... 75

5.20 Test Region # 2: partitioning with combined inputs ........... 75

5.21 Test Region # 2: individual inputs after filtering ............. 76

5.22 Test Region # 2: combined inputs after filtering .............. 76

5.23 Test Region # 3: before and after speckle filtering ............ 78

5.24 Test Region # 3: partitioning with HH ......................... 79

5.25 Test Region # 3: partitioning with HV ......................... 79

5.26 Test Region # 3: partitioning with VV ......................... 80

5.27 Test Region # 3: partitioning with \( \rho_{HHV} \) .................. 80

5.28 Test Region # 3: partitioning with HH, HV, and VV .......... 81

5.29 Test Region # 3: partitioning with HH, HV, VV, and \( \rho_{HHV} \) 81

5.30 Test Region # 3: using HH after speckle filtering ............ 83

5.31 Test Region # 3: using HV after speckle filtering ............ 83

5.32 Test Region # 3: using VV after speckle filtering ............ 84

5.33 Test Region # 3: using \( \rho_{HHV} \) after speckle filtering ............... 84

5.34 Test Region # 3: using HH, HV, and VV after speckle filtering 85

5.35 Test Region # 3: using HH, HV, VV, and \( \rho_{HHV} \) after filtering 85

5.36 S3 Results: Test Region # 1 (nc=8; k=6) ....................... 86

5.37 S3 Results: Test Region # 2 (nc=10; k=8) ....................... 86

5.38 S3 Results: Test Region # 3 (nc=18; k=13) ....................... 87

5.39 Proposed Method Results: Test Region # 1 ...................... 91
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.40</td>
<td>Proposed Method Results: Test Region # 2</td>
<td>91</td>
</tr>
<tr>
<td>5.41</td>
<td>Proposed Method: Test Region # 3 segmentation for different nc values</td>
<td>92</td>
</tr>
<tr>
<td>5.42</td>
<td>Proposed Method: Test Region # 3 segmentation for different nc, d values</td>
<td>93</td>
</tr>
<tr>
<td>5.43</td>
<td>Proposed Method: Test Region # 3 classification using SRW distance</td>
<td>94</td>
</tr>
<tr>
<td>5.44</td>
<td>Proposed Method: Overall Accuracy (OA) vs. Number of partitions (nc)</td>
<td>95</td>
</tr>
<tr>
<td>5.45</td>
<td>Proposed Method: Westham Island segmentation using block processing</td>
<td>96</td>
</tr>
<tr>
<td>5.46</td>
<td>Flevoland subset data and results</td>
<td>98</td>
</tr>
<tr>
<td>B.1</td>
<td>S4 Results with Bartlett distance - Test Region # 1 (d = 7) for s ∈ {1, 0.1, 0.05, 0.01, 0.005, 0.0005, 0.0001}</td>
<td>119</td>
</tr>
<tr>
<td>B.2</td>
<td>S4 Results with Bartlett distance - Test Region # 2 (d = 7) for s ∈ {1, 0.1, 0.05, 0.01, 0.005, 0.0005, 0.0001}</td>
<td>119</td>
</tr>
<tr>
<td>B.3</td>
<td>S4 Results with Bartlett distance - Test Region # 3 (d = 7) for s ∈ {1, 0.1, 0.05, 0.01, 0.005, 0.0005, 0.0001}</td>
<td>120</td>
</tr>
<tr>
<td>B.4</td>
<td>S4 Results with Symmetric Revised Wishart distance - Test Region # 1 (d = 7) for s ∈ {1, 0.1, 0.05, 0.01, 0.005, 0.0001}</td>
<td>120</td>
</tr>
<tr>
<td>B.5</td>
<td>S4 Results with Symmetric Revised Wishart distance - Test Region # 2 (d = 7) for s ∈ {1, 0.1, 0.05, 0.01, 0.005, 0.0001}</td>
<td>121</td>
</tr>
<tr>
<td>B.6</td>
<td>S4 Results with Symmetric Revised Wishart distance - Test Region # 3 (d = 7) for s ∈ {1, 0.1, 0.05, 0.01, 0.005, 0.0001}</td>
<td>121</td>
</tr>
<tr>
<td>B.7</td>
<td>S5 Results: Bartlett distance - Test Region # 1 (s = 0.05)</td>
<td>124</td>
</tr>
<tr>
<td>B.8</td>
<td>S5 Results: Bartlett distance - Test Region # 2 (s = 0.05)</td>
<td>125</td>
</tr>
<tr>
<td>B.9</td>
<td>S5 Results: Bartlett distance - Test Region # 3 (s = 0.05)</td>
<td>126</td>
</tr>
<tr>
<td>B.10</td>
<td>S5 Results: Bartlett distance - Test Region # 3 (s = 0.005)</td>
<td>126</td>
</tr>
<tr>
<td>B.11</td>
<td>S5 Results: SRW distance - Test Region # 1 (s = 0.05)</td>
<td>127</td>
</tr>
<tr>
<td>B.12</td>
<td>S5 Results: SRW distance - Test Region # 2 (s = 0.05)</td>
<td>128</td>
</tr>
<tr>
<td>B.13</td>
<td>S5 Results: SRW distance - Test Region # 3 (s = 0.05)</td>
<td>129</td>
</tr>
<tr>
<td>B.14</td>
<td>S5 Results: SRW distance - Test Region # 3 (s = 0.005)</td>
<td>129</td>
</tr>
</tbody>
</table>
List of Symbols

\( A \)  
affinity matrix

\( b \)  
input channel index

\( d \)  
spatial proximity neighborhood size, pixels

\( \chi^2(h_i, h_j) \)  
chi-square distance measured between \( h_i \) and \( h_j \)

\( d(s_i, s_j) \)  
Euclidean distance measured between \( s_i \) and \( s_j \)

\( d_B(T_i, T_j) \)  
Bartlett distance measured between \( T_i \) and \( T_j \)

\( d_{SRW}(T_i, T_j) \)  
Symmetric Revised Wishart distance between \( T_i \) and \( T_j \)

\( D \)  
diagonal matrix, where entries are the sum of the rows of \( W \)

\( D_A \)  
diagonal matrix, where entries are the sum of the rows of \( A \)

\( ev \)  
edge variance, a constant used to relate the orientation energy to the scaling parameter for the contour cue

\( E \)  
edges in an undirected graph

\( G \)  
undirected graph

\( h_i \)  
histogram of the edge aligned patch associated with pixel \( i \)

\( I \)  
identity matrix

\( k \)  
number of classes or clusters

\( K \)  
Kappa coefficient

\( \lambda \)  
elongation factor of the filter used to calculate OE

\( L \)  
normalized affinity matrix

\( \mathcal{L} \)  
normalized graph Laplacian
List of Symbols

\( M \) \hspace{1cm} \text{number of bins in histogram calculation of } \chi^2\text{-distance}

\( N \) \hspace{1cm} \text{number of pixels in an image}

\( nc \) \hspace{1cm} \text{number of partitions}

\( nsb \) \hspace{1cm} \text{number of segments per block}

\( Nblk \) \hspace{1cm} \text{number of blocks used to cover the entire scene}

\( n_b \) \hspace{1cm} \text{number of input channels}

\( n_{LS} \) \hspace{1cm} \text{number of nearest neighbors used for the estimation of the local scaling parameter}

\( N(s_i) \) \hspace{1cm} \text{the set of nearest neighbors of the instance } s_i \text{ used for the estimation of } \sigma_i

\( OA \) \hspace{1cm} \text{overall classification accuracy}

\( OE \) \hspace{1cm} \text{orientation energy}

\( ori \) \hspace{1cm} \text{number of orientations used in OE calculation}

\( r \) \hspace{1cm} \text{radius of the neighborhood used in the spatial proximity cue}

\( s \) \hspace{1cm} \text{a constant that relates the maximum distance measured in the spatial proximity neighborhood to the scaling parameter}

\( sr \) \hspace{1cm} \text{sampling rate parameter used for sub-sampling data pairs to decrease the computational cost in } W \text{ calculation}

\( \sigma \) \hspace{1cm} \text{scale parameter of the elongated Gaussian filter used in OE calculation}

\( \sigma_c \) \hspace{1cm} \text{scaling parameter used in the affinity function for the contour cue}

\( \sigma_B \) \hspace{1cm} \text{scaling parameter estimated through local scaling when Bartlett distance is used in polarimetric cue}

\( \sigma_i \) \hspace{1cm} \text{scaling parameter estimated through local scaling for the instance } i

\( \sigma_{SRW} \) \hspace{1cm} \text{scaling parameter estimated through local scaling when Symmetric Revised Wishart distance is used in polarimetric cue}
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_i$</td>
<td>polarimetric coherency matrix of an instance $i$</td>
</tr>
<tr>
<td>$V$</td>
<td>nodes in an undirected graph</td>
</tr>
<tr>
<td>$\omega(u, v)$</td>
<td>weight that represents similarity of two nodes $u$ and $v$</td>
</tr>
<tr>
<td>$W$</td>
<td>similarity matrix</td>
</tr>
<tr>
<td>$W^b$</td>
<td>similarity matrix for the $b^{th}$ input channel</td>
</tr>
<tr>
<td>$W^C$</td>
<td>similarity matrix for the contour cue</td>
</tr>
<tr>
<td>$W^B$</td>
<td>similarity matrix for the polarimetric cue formed using the Bartlett distance</td>
</tr>
<tr>
<td>$W^P$</td>
<td>similarity matrix for the spatial proximity cue formed using the ramp function</td>
</tr>
<tr>
<td>$W^{SRW}$</td>
<td>similarity matrix for the polarimetric cue formed using the Symmetric Revised Wishart distance</td>
</tr>
<tr>
<td>$W^{tot}$</td>
<td>similarity matrix for the combined cues</td>
</tr>
</tbody>
</table>
Acknowledgements

I am grateful to my research supervisor, Dr. Ian G. Cumming for his continuous encouragement, help and guidance throughout my research program. I also would like to express my sincere appreciation to Dr. Rabab K. Ward, who has been generous in showing her support as my co-supervisor in the last 3.5 years of this journey.

I acknowledge research funding from Natural Sciences and Engineering Research Council of Canada (NSERC) through the research grants of Dr. Cumming and Dr. Ward. I also appreciate the partial scholarship provided by the Turkish Education Foundation (TEV) for the first year of my graduate studies at the University of British Columbia (UBC). Without this financial support I would not find the courage to move to Canada to further my education.

I acknowledge the owners of the Westham Island data set used in this thesis: The Canadian Space Agency (CSA) with the cooperation of Environment Canada, Natural Resources Canada (CCRS), Defense Research and Development Canada - Ottawa, and MacDonald Dettwiler and Associates Ltd. Geospatial Services (MDA GS, formerly known as Radarsat International – RSI). I also thank S. N. Anfinsen for providing the digitized ground truth map for the Flevoland data set.

I do sincerely thank my supervisory committee members, the members of the examining committee, and the external examiner for their detailed comments and recommendations, which were very helpful in improving the presentation of this thesis.

I would like to thank my friends and colleagues from Radar Remote Sensing Group (RRSG) and Image and Signal Processing Lab at UBC, who have always inspired and supported me to continue this long journey.
Acknowledgements

I wish to thank Verda, who did not hesitate to follow me to Canada for sharing a dream, and stood beside me for many years as I pursued several social and academic endeavors.

Among many great people I have connected with in Canada, there are a few that have been very important in bringing myself up to the place I stand today, without whom I would not enjoy this journey: My friend and colleague Flavio, has always been there when he was most needed. Former students of the RRSG, Millie, Bernd, Yew Lam, and fellow members of the Signal and Image Processing Lab, Mahsa, Ilker, and Rupin have shared the joys and frustrations of the academic life, and marked some of many unforgettable moments in my life. Rupin and Flavio have offered their insight and kept me company during long nights spent in the lab with brief trips to UBC Village to re-fuel. During tough times, Aliye helped me to remind myself how powerful it is to be able to dream and run after those dreams, helped me to re-discover my passions that define who I am. Demet has been a role model for perseverance, has always been motivational, realistic, and supportive. Edeer, Karacabeyli and Ergudenler families, Flavio & Kelly, became my family in Vancouver in the last few years. They have played vital roles in my social support system along with my dear friends, Esra Burcu, Emine & Feti, Emrah & Seniz, Ozge & Cheng, Recep & Pinar, Rubab, and Tolga. Last but not the least, Burhan has been a key player, often behind the scenes, in helping me turn this year into a remarkable one from social, personal and academic aspects.

Finally, I wish to thank my family, especially to my parents Hulya & Ergun Ersahin, who have raised me up to whom I am today, who have unconditionally supported me, trusted my decisions, and who will always love me.

October 2009

KAAN ERSAHIN

Vancouver, BC
Chapter 1

Introduction

Synthetic Aperture Radar (SAR) is an active microwave sensor and plays an important role in Earth observation systems. Unlike optical instruments, radar imaging is not limited to daytime and good weather conditions. Therefore, SAR systems enable continuous monitoring of the terrain for time-critical applications such as crop and sea-ice monitoring. The information provided by SAR is quite different from optical remote sensing data, since radar is sensitive to different properties than optical sensors - radar is sensitive to electrical and structural properties (e.g., dielectric constant, surface roughness).

For more than two decades, radar remote sensing has been used for many applications (e.g., oil-slick detection, sea-ice monitoring, ship detection, crop mapping). However, single polarization data is often insufficient for resolving ambiguities among targets of interest. The dimensionality of the observation can be increased by employing:

- multiple incidence angles
- multiple frequencies
- multiple polarizations
- multi-temporal acquisitions.

Using fully polarimetric observations (i.e., HH, HV, VV, VH channels) the information that can be obtained about a target is more complete than the single polarization case.

In the last two decades, SAR polarimetry has shown promise through airborne research campaigns, leading to the design and recent launch of fully
polarimetric space-borne missions (i.e., RADARSAT-2, ALOS-PALSAR, and TerraSAR-X). The Canadian RADARSAT-2 was launched in December 2007 and is one of the most advanced space-borne SAR systems in current operation. As new space-borne data becomes available, users of remote sensing technology can experiment with polarimetric SAR (POL-SAR) data in their analysis tasks. New application areas are being developed and large data volumes may soon require routine daily interpretation. These advances motivate the development of automated data analysis procedures.

Automated analysis and interpretation of remote sensing data typically involve segmentation and/or classification tasks. Classification of such data is challenging due to technical and conceptual issues. Researchers mostly focus on the technical issues (i.e., feature extraction and classification techniques). However, conceptual issues are usually masked by classification errors and the effects of limited system resolution. For example, a single class may actually be a mixture of targets (e.g., urban: buildings, roads, and trees). Apart from these aspects, radar image analysis has its own challenges due to speckle, incidence angle effects, layover and shadowing.

In the radar polarimetry literature, there have been many attempts to understand the information content of the data and improve the accuracy of the classification task. Different representations of polarimetric information have been used and many features have been defined to discriminate between targets. These techniques are based on utilizing polarimetric measurements and scattering mechanism information at the pixel level.

Although radar image segmentation is a challenging task for automated systems, for trained technicians segmentation and visual interpretation of such data is relatively easy. This is also true for many problems in computer vision, and usually the ultimate goal is to achieve the performance of human experts, which is possible through a good understanding of how the human perceptual system handles this task.

For humans, an image represents more than a collection of pixels: it is a meaningful organization of objects and patterns. The Gestalt Psychologists\textsuperscript{1} postulated that the brain is holistic, parallel, and analog, with self-organizing tendencies, or that...
studied this important phenomenon in the late 1930s and listed various factors (e.g., proximity, similarity, closure, symmetry, and continuity) that contribute to the perceptual grouping process. These factors are known as cues\(^2\).

Over the last few decades, research in the area of computer vision has sought principled ways to utilize these ideas. A promising technique for grouping problems, known as spectral graph partitioning, has emerged over the last few years. It is based on combining low-level cues (e.g., brightness, texture, proximity) in a single pairwise similarity matrix and solving the corresponding graph partitioning problem using a spectral clustering algorithm that provides k-way partitioning in an arbitrary feature space. This approach has been shown to solve the perceptual grouping problem and achieve results on a par with human perception.

Segmentation and visual interpretation of polarimetric SAR image data by human experts are common practice due to the fact that automated techniques are often not able to provide satisfactory results. The goal of reaching the performance level of humans, as in other computer vision applications, motivates the use of the spectral graph partitioning approach for SAR image segmentation and classification.

1.1 Background Technology

The potential of polarimetric radar imaging became evident in the late 1980s. However, the extent of the available information was not known and relevant techniques had not yet been developed \(^1\). The fundamental issues to be addressed were: (i) analysis of the statistical properties of the data, (ii) development of speckle reduction techniques, (iii) definition of rel-

---

\(^2\)A sensory cue is a statistic or signal that can be extracted from the sensory input by a perceiver, that indicates the state of some property of the world that the perceiver is interested in perceiving. Sensory cues include visual cues, auditory cues, tactile cues, and haptic cues. These cues play an important role in theories of perception, especially theories of appearance (how things look). In this work we refer to visual cues.
evant features for discrimination, and (iv) development of segmentation and classification methods. Since then, theoretical and empirical studies have explored various schemes to find relevant approaches for specific problems.

1.1.1 Statistical Models for Polarimetric SAR Data

Scattering objects (i.e., targets) can be grouped into two main categories: (i) distributed targets, and (ii) point targets. Most of the natural targets (e.g., crop fields, forests, and ocean) are distributed targets and show stochastic behavior. They can be characterized using covariance matrix or coherency matrix representations of the polarimetric measurements, which are more meaningful when averaged over an homogeneous area. In contrast, characterization of man-made targets (e.g., ships, houses) is best performed at the pixel level [2]. For this work our interest is mainly in natural targets, therefore the following discussion will focus on techniques relevant to this target group.

SAR is a coherent imaging system, and therefore suffers from a noise-like phenomenon, known as speckle, which can be explained using the following model. For distributed targets, each resolution cell contains a large number of discrete scatterers. However, the contributions of individual scatterers are unobservable, since they are on a smaller scale than the resolution of the radar. The total return from a resolution cell is the vector sum of the backscattered wave contributions from all the scatterers in that resolution cell. Since each scatterer has a different location, they will contribute to the total response with different phases. As a result, the total backscatter will be affected by constructive or destructive interference, which depends upon the random distribution of the phase terms. To infer this distribution, one should note that slant range resolution is typically many wavelengths across. Hence, scatterers located at different parts of the resolution cell have very different phase terms, even if their scattering behavior is identical. In practice the phase of the contribution from individual scatterers is considered to be uniformly distributed within $[-\pi, \pi]$, and is independent from

---

3Covariance and coherency matrix representations carry the complete and equivalent polarimetric information on distributed targets (for details see Appendix A.2)
the amplitude. Therefore, this coherent summation at every resolution cell looks like a random walk in the complex plane. The analysis reveals that for a large number of statistically identical scatterers, the observed in-phase (I) and quadrature (Q) components of the complex measurement will be independent, identically distributed, and zero-mean Gaussian random variables.

For polarimetric SAR data, this speckle model generalizes to a multivariate complex Gaussian distribution, and its covariance matrix has a complex Wishart distribution [3]. Under the assumption that the radar cross-section (RCS) for a distributed target (e.g., field of wheat) is constant, the polarimetric measurements can be represented by the *multiplicative speckle model* (i.e., image intensity is given by a constant RCS multiplied by the speckle) [4]. However, if there is spatial variation of the RCS across the target, also known as texture, this can be accounted by using the *product model*, where the radar backscatter is the product of two independent processes (i.e., a speckle model for the zero-mean multivariate complex Gaussian data and a univariate gamma distribution for the texture).

To model fully polarimetric high-resolution radar data, Yueh *et al.* [5] used the multivariate K-distribution that follows from the *product model*, where texture is assumed to be independent from polarization. Quegan and Rhodes [6] evaluated both the Gaussian and K-distribution models for different vegetation types, and reported that at C-band, the Gaussian model agrees well with the data. However, at L- and P-bands, the K-distribution was a better fit. Lee *et al.* [7] studied the statistical characteristics of multi-look processed data, and using the complex Wishart distribution, derived a maximum likelihood (ML) classifier [8]. Hoekman *et al.* [9] studied the classification of agricultural fields and argued that although a homogeneous area may be described accurately by the complex Wishart distribution, this is not necessarily true for a collection of homogeneous areas of the same class due to, for example, variation in biophysical parameters. De Grandi *et al.* [10] showed that texture is dependent on polarization (i.e., the assumption made by the product model is invalid) and argued for the need for improved models that account for the polarization dependence of texture. Bombrun and Beaulieu used the Fisher distribution for texture modeling and reported
improvements over the Wishart distribution [11].

Doulgeris et al. [12] presented a generalized Wishart classifier derived from a non-Gaussian model for POLSAR data. They argued that the scale of mixture of Gaussian distribution (SMoG) model is suitable for modeling POLSAR data and derived a solution for one particular case, i.e., multivariate K-distribution. Sample covariance matrix based on this distribution called K-Wishart distribution is used in a Bayesian classification scheme.

1.1.2 Speckle Reduction Techniques for Polarimetric SAR

To improve classification accuracy, speckle reduction techniques are usually required. Novak and Burl [13] proposed the polarimetric whitening filter (PWF), which combines the elements of the scattering matrix and forms a single-channel image with reduced speckle. Based on the product model, Lopes and Sery [14] suggested a maximum a posteriori (MAP) algorithm to estimate the local texture coefficient for a given processing window. Lee et al. [3] used the multiplicative speckle model and filtered the diagonal terms of the covariance matrix, which are real and represent the channel powers. This approach was later generalized to all the elements for single-look and multi-look cases [15],[16]. The technique in [16] suffers in practice from the difficulty of filtering the cross-product terms. Lee et al. [17] circumvented this difficulty by simplifying the problem to one dimension in the POLSAR filter. The speckle filtering technique known as Lee filter [4], that was originally designed for single polarization data is used on the elements of the covariance matrix independently. This filter is based on minimizing the mean squared error to estimate the unspeckled intensity value at every pixel. The POLSAR filter combines the speckle reduction procedure with an edge-aligned windowing scheme to prevent the edges from smearing [17]. This technique is widely used in the recent radar polarimetry literature and is acknowledged to be the state-of-the-art speckle reduction technique for polarimetric SAR data.
1.1. Background Technology

1.1.3 Feature Extraction, Classification and Segmentation Strategies

Early work in polarimetry was based on experiments with several features such as channel intensities, ratios and phase differences, in order to assess their discriminative power. The objective was to understand the information content of polarimetric data, and find the most suitable combination for representing this valuable information in a concise way [18]. Among the polarimetric measurements, the phase and magnitude of the cross-polarized correlation coefficients (i.e., $\phi_{HVHH}$, $\rho_{HVHH}$, $\phi_{VVHV}$, and $\rho_{VVHV}$) are not expected to carry useful information [18], [19]. Van Zyl et al. [20] introduced the polarization signature (the normalized backscattered power as a function of ellipticity and orientation angles), and the coefficient of variation as polarimetric discriminators.

Kong et al. [21] introduced a supervised classification method for single-look polarimetric data, that follows from the multiplicative speckle model. The multivariate complex Gaussian distribution was used to derive the optimum classification rule based on the likelihood ratio test. Lee et al. [8] used the same approach and extended the Maximum Likelihood (ML) classifier to the multi-look case, where the covariance matrix representation of polarimetric measurements have a complex Wishart distribution. Based on this approach, a distance measure was derived, which simplifies to (1.1) under the assumption that classes have equal prior probabilities.

$$d(C, \omega_m) = \ln |C_m| + Tr(C_m^{-1} C)$$ (1.1)

where $C$ is the covariance matrix representation of the polarimetric measurements for every pixel in a multi-look processed data set, and $C_m$ represents the mean covariance matrix for the class $\omega_m$. In Lee’s classification scheme [8], a pixel is assigned to the class whose mean is closest. In a supervised scenario, a set of training points from each class can be used to estimate these class means.

Van Zyl [19] was the first to introduce an unsupervised classification scheme based on the relation between the number of bounces (i.e., reflec-
1.1. Background Technology

ations that the radar signal experiences) and the co-polarized phase difference ($\phi_{HHVV}$). This method was shown to be useful in identifying three scattering types: (i) odd bounce, (ii) even bounce, and (iii) diffuse scattering.

Rignot et al. [22] suggested a supervised technique for segmentation of Polarimetric SAR data based on a model for the conditional distribution of the polarimetric complex data that is combined with a Markov Random Field (MRF) representation for the distribution of the region labels to obtain the posterior distribution. Optimal region labeling is then given by the maximum a posteriori (MAP) estimate.

Rignot et al. [23] also suggested an unsupervised technique for segmentation of Polarimetric SAR data using multi-dimensional fuzzy clustering of the logarithm of the elements in the covariance matrix. This approach was tested on Polarimetric SAR data of lava flows and of sea-ice acquired by the NASA/JPL airborne sensor (AIRSAR).

Cloude and Pottier [24] derived an unsupervised method based on Cloude’s target decomposition theory [25]. The eigenvalue decomposition of the polarimetric covariance matrix is used to represent the scattering mechanism by three orthogonal contributions. Each eigenvector represents one of the three bases, and the corresponding eigenvalues quantify the strength of the contributions. The eigenvalues and eigenvectors are used to define entropy ($H$), anisotropy ($A$) and alpha angle ($\alpha$). $H$ and $\alpha$ span a two-dimensional feature space that is partitioned into zones representing different scattering behaviors. The $\alpha$-axis discriminates between surface, volume, and double-bounce scattering, while the $H$-axis represents low, medium, and high entropy. $H = 0$ indicates a single dominant mechanism, and $H = 1$ corresponds to a random mixture of mechanisms with equal probabilities. Anisotropy, $A$, provides information on the relative importance of second and third scattering mechanisms, and is useful for discrimination within the intermediate entropy region. Partitioning the $H/A/\alpha$ feature space was also used to determine the number of classes and to initialize cluster centers in Lee’s iterative Wishart classifier [26]. This iterative scheme, based on

---

4A Markov random field is a graphical model in which a set of random variables have a Markov property described by an undirected graph.
the complex Wishart distribution, has been used for sea-ice classification by Scheuchl et al. [27]. An extension of the iterative Wishart classification scheme to multi-frequency data sets is provided by Ferro-Famil et al. [28].

Freeman and Durden introduced an unsupervised classification scheme based on a three-component scattering model [29]. Their decomposition uses physical models for the scattering mechanisms (i.e., surface, double-bounce, and volume scattering) rather than the pure mathematical decomposition used by Cloude. The contribution of each component to the total power is used to determine the dominant scattering mechanism, and the class for each pixel. This approach was also integrated in Lee’s classification scheme [30], resulting in improved performance.

Kersten et al. [31] modified the Fuzzy C-means (FCM) clustering algorithm by using a distance measure based on the complex Wishart distribution (1.1). A comparison of fuzzy clustering and expectation maximization (EM) is given for different distance measures in [32]. The main conclusion of [32] is that both fuzzy clustering and EM perform similarly, if the distance measure based on the complex Wishart distribution is used.

Chen et al. [33] presented a supervised technique that integrates a fuzzy neural network as a classifier, the covariance matrix elements as a feature vector, and a distance measure based on the statistical properties of polarimetric data, i.e., Wishart distribution.

Ayed et al. [34] proposed a level set method for polarimetric SAR image segmentation. It consists of minimizing a functional containing an observation term and a boundary length prior. The minimization is carried out by a multi-phase method which embeds a simple partition constraint directly in curve evolution. Results are obtained for airborne SAR image data from AIRSAR and compared to Beaulieu and Touzi’s approach in [35].

Cao et al. [36] used $H/A/\alpha$ and total backscattering power (SPAN) together to initialize the Wishart classifier.

Recently, Wu et al. [37] suggested a segmentation scheme for multi-look fully polarimetric SAR images using the statistical characteristics of the pixels and the interaction between adjacent pixels. In order to use the statistical knowledge of the data and the spatial relation of neighboring pix-
1.1. Background Technology

els, the Wishart distribution of the covariance matrix is integrated with the Markov random field, then the iterated conditional modes (ICM) algorithm is used to obtain the maximum a posteriori (MAP) estimation of pixel labels. Although the ICM has fast convergence, it is affected easily by initial conditions, so the Wishart-based ML is used to obtain the initial segmentation map. Using multi-look fully polarimetric SAR images, acquired by the NASA/JPL AIRSAR sensor, they have presented comparisons with the Watershed, Gaussian MRF and pixel-based ML (a.k.a. supervised Wishart) classifier. This supervised technique called Wishart MRF was found to offer better segmentation performance and higher accuracy \cite{37, 38}.

Another promising approach to segmentation of single polarization SAR data is the iterative region growing using Semantics (IRGS) that is based on edge penalties and region growing \cite{39}. Yu and Clausi argue that IRGS is an improvement over the traditional MRF-based approaches since a more stable estimation of model parameters is achieved. This approach has also been applied by Yu and Clausi to the problem of classifying operational SAR sea-ice imagery in and the results have been validated by the Canadian Ice Service (CIS) personnel \cite{40}.

1.1.4 Polarimetric SAR Data and Texture

Almost all the methods for classification of polarimetric SAR data are based on the multivariate complex Gaussian model, and utilize polarimetric information only at the pixel level. Besides the extensive use of texture analysis techniques for single-polarization SAR image data, there has been very limited research on employing this information in the fully polarimetric case. The first attempt was based on the product model, where the variance of the gamma distributed texture was estimated \cite{41}. However, the assumption employed in this model (i.e., texture being polarization independent) is questionable according to the recent research results \cite{10}.

From an image processing perspective, texture analysis is often performed using features extracted from the Gray Level Co-occurrence Matrix (GLCM). In a processing window, a pixel pair is defined by a separation,
1.1. Background Technology

\(\delta\), and an orientation angle, \(\theta\). The gray levels associated with the pair are used to form the \(G \times G\) co-occurrence matrix, where \(G\) is the number of gray levels. The entries of this matrix are related to the occurrence probabilities of gray level pairs. Anys and He [42] studied several features extracted from the first-, second- and third-order histograms and reported improved results over the single-channel case, by utilizing all three channels (i.e., HH, VV, and HV).

Another approach to modeling texture in images is to use Markov Random Fields (MRFs). In this case, each pixel in the image is represented by a node in an undirected graph, where the interactions between pixel pairs are quantified by the clique potentials. The nodes are often locally connected to their immediate neighbors. Therefore, given its neighbors, the label of a pixel is independent from the rest of the image. A comparison of GLCM and MRF for single-polarization SAR image segmentation is given in [43], where GLCM provided improved discrimination over MRF for small window sizes; however, it was found to be more sensitive to boundary confusion.

Beaulieu and Touzi [35] argued that algorithms based on the complex Gaussian model should only be used for homogeneous areas (i.e, targets with constant RCS) since their performance degrades significantly in the presence of texture. They used the product model and showed improved results for textured areas. The technique is a bottom-up approach that starts at the pixel level and in every step merges the two segments that maximize the log-likelihood. However, two concerns remain: (i) robustness to speckle, and (ii) validity of the assumption made by the product model.

In [44], we have argued that improved classification accuracy can be obtained by using texture information in addition to the polarimetric feature vector. Texture features were extracted using GLCM, and a neural network classifier was used to obtain results from fully polarimetric airborne data in L-band. However, texture features obtained using this approach are not expected to be relevant in the case of space-borne data due to lower resolution, especially in dual-polarization ScanSAR mode of RADARSAT-2. Since our goal was to develop schemes for RADARSAT-2 data analysis, we have excluded the use of texture information in our research.
1.1. Background Technology

Recently, Bombrun and Beaulieu [11] used the Fisher distribution for texture modeling and reported improved results in comparison to the Wishart distribution and homogeneity assumption for distributed targets.

1.1.5 Emerging Techniques in Computer Vision for Grouping Problems

For humans, an image represents a meaningful arrangement of objects and patterns, rather than a random collection of pixels. This phenomenon is a result of an important aspect of human vision system known as perceptual organization which has been studied over the last few decades in computer vision. The factors that contribute to perceptual organization (i.e., cues) have been used as guidelines for developing grouping algorithms in computer vision.

Recently proposed techniques perform grouping based on pairwise similarities of pixels. The similarity between pixel pairs is encoded in a single affinity matrix\(^5\) that contain information from one or more cues. Then the grouping is performed using spectral graph partitioning [45].

Pairwise grouping methods present an appealing alternative to central grouping\(^6\), such as k-means or mixture model fitting via expectation maximization (EM), which have the drawback of implicitly assuming a Gaussian distribution for the feature vectors. This assumption is the justification for using Euclidean or Mahalanobis distance to compare the feature vectors to the class means. Pairwise grouping methods also avoid the restriction that all points in a cluster must be close to a prototype, which allows the recovery of groups that have complicated manifold structures in the feature space. These methods also enable combination of multiple cues (e.g., brightness, texture, color, and proximity) [46], and offer flexibility in the definition of similarity between pixels. For example, if the feature vectors represent color

\(^5\)An affinity matrix (a.k.a. similarity matrix) contains the weights between the nodes of a graph. These weights are computed using the affinity functions that map the measured distances to similarities between pixels. An appropriate distance measure is used for each cue.

\(^6\)Central grouping techniques aim to cluster data points based on their distances to the cluster centers.
histograms, then using the $L_2$-norm to measure similarity is not meaningful. Instead a suitable measure such as $\chi^2$-distance can be employed.

Spectral graph partitioning has shown promise in solving grouping problems in computer vision. However, application to practical problems, especially image segmentation, has been delayed due to the computational complexity involved. Recently, a technique that provides an approximate solution has become available, which allowed spectral graph partitioning to be used for image segmentation problems [47].

Image segmentation is the most studied form of grouping, and the techniques developed for this problem can be categorized into two broad families: (i) region-based and (ii) contour-based approaches.

Region-based approaches aim to partition image pixels into sets with similar properties such as brightness, color and texture. They usually involve defining a global objective function, which ensures that decisions will be made after the information from the whole image is available. The major drawback of region-based approaches is that contour information is not utilized.

On the other hand, contour-based approaches usually perform an edge detection step followed by a linking process, that will bridge small gaps along the edges and produce extended contours. However, in general this will not result in closed and connected regions. Since the decisions on the boundaries are made locally, these methods can not easily deal with textured regions. Also, they do not recognize the fact that extended contours with low contrast are perceptually more significant than short edges with high contrast.

To benefit from both approaches, Leung and Malik proposed to use contour information in a region-based setting through the orientation energy [48]. Orientation Energy (OE) is an enhanced edge map where high value of OE along the line that connects a pair of pixels suggests the presence of a contour passing between those two pixels, therefore the similarity of that pair will be low.

The orientation energy calculation involves the second derivative of an

\footnote{In this thesis, with the word \textit{contour} refers to the boundaries of regions in the image.}
1.2 Thesis Outline and Objective

The main objective of this work is to develop an effective methodology for classification of Polarimetric Synthetic Aperture Radar (POLSAR) data. Classification techniques in the literature have only used the polarimetric attributes of pixels, and did not utilize the image aspect of such data. However, segmentation and visual interpretation of polarimetric SAR image data by human experts are common practice due to the fact that automated techniques are often not able to provide satisfactory results. We aim to improve the accuracy of polarimetric SAR data classification by utilizing the emerging spectral graph partitioning (SGP) technique. SGP can utilize visual cues to incorporate some of the advantages of the human vision system into the polarimetric analysis. This technique has been shown to achieve perceptually plausible results on other computer vision applications, which motivates us to customize it for POLSAR segmentation and classification.

The spectral graph partitioning methodology and two grouping algorithms (i.e., spectral clustering and multiclass spectral clustering) are explained in Chapter 2, which also includes a brief discussion of four approaches that can be used to reduce the computational complexity involved.

In the process of developing new schemes for POLSAR data classification, the performance level of humans is aspired to through the use of four cues: patch-based similarity, contour, spatial proximity, and similarity of coherency matrices (i.e., the polarimetric cue), which are discussed in Chapter 3. Some of these cues were customized for POLSAR data and used either individually or in groups as building blocks of the schemes tested for their utility and the proposed method.

Chapter 4 of this thesis introduces the preliminary research and the proposed method. These are explained using a roadmap, which represents the elongated Gaussian kernel and its Hilbert transform. Since the kernels are elongated, the information is integrated along the edge. Thus, extended low-contrast contours produce stronger responses than short high-contrast ones.
1.2. Thesis Outline and Objective

research progression over several years. The preliminary research includes five schemes (i.e., S1 to S5) that are used to assess the utility of either individual cues that were presented in Chapter 3 or their combinations. The analysis of experimental results from such schemes have shaped the final method proposed in this thesis.

The results and discussion corresponding to these schemes and the proposed method are presented in Chapter 5 and Appendix B. A summary of these schemes and some of the important findings from the analysis of experimental results will be outlined below.

Quantitative results are compared with the unsupervised technique known as the Wishart classifier. The Wishart classifier has been used as a benchmark, because it is a well-established algorithm and it has been widely employed for many applications using several platforms including AIRSAR, ESAR, EMISAR and ALOS-PALSAR [36], [49], [50]. The Wishart classifier is currently offered by various commercial and research software tools, such as SAR Polarimetry Workstation (SPW) of PCI Geomatics, and PolSARPro of ESA.

The first scheme, S1, utilizes both patch-based similarity and spatial proximity cues, through an adaptation of the Spectral Clustering (SC) algorithm. However, only the intensity information from polarimetric data are used, and off-diagonal terms in the coherency matrix are not utilized. This scheme outperforms the Wishart classifier which has been used as a benchmark. It also does not require any restriction on the statistical distribution of the data. However, it has an important limitation and that is the requirement of large memory allocation for computation of histogram similarities. It also does not utilize a powerful visual cue, i.e., the contour cue.

The second scheme, S2, utilizes the spatial proximity cue and contour information by defining the pairwise similarities based on the orientation energy. The multiclass spectral clustering (MSC) algorithm is used to obtain the partitioning, which provides a discrete solution that is close to the global optimum of our objective function (i.e., normalized cut). This scheme results in perceptually plausible results. It also outperforms the benchmark (i.e.,
Wishart classifier). However, fragmentation of non-convex regions may occur when S2 is used as a one-step approach. This scheme also does not solve the classification problem since both adjacent and non-adjacent fields of the same content are not labeled as a single class at this stage.

The third scheme, S3, is an attempt to address the potential issues that arise when S2 is used as a one-step approach. In the first step of S3, S2 is employed for obtaining an over-segmentation that ensures enough detail. A variant of S1 is then applied to the outcome, which allows merging fragmented segments and adjacent fields of the same content. Since this second step considers only adjacent segments to have potentially non-zero similarity, non-adjacent fields of the same content are still identified with different labels. Results show improvements over S2 in complex test regions. However the information content of the polarimetric data is still partially utilized, which was also the case in S1 and S2.

The fourth scheme, S4, utilizes the full extent of the polarimetric information in the data set through the coherency matrix representation. This scheme is used to analyze the behavior of the polarimetric cue (i.e, similarity of coherency matrices) along with the spatial proximity cue. Also, an attempt has been made to automate the selection of the scaling parameter in the affinity function. Visual assessment of the results however suggests that the approach used for estimation of the scaling parameter is not robust.

The fifth scheme, S5, combines the contour information, polarimetric cue, and spatial proximity, which can also be seen as an extension of S2 with the addition of the polarimetric cue, or of S4 with the addition of the contour cue. Even though polarimetric cue is included, this scheme still suffers from the same drawbacks of using S2 as a one-step approach. The results also confirm the findings of S4 on the automation of the scaling parameter selection.

The proposed method in this thesis is a two-step approach to the POL-SAR classification problem. First groups of pixels with similar properties are obtained, then these segments are grouped into classes. Both the con-
tour cue that is based on the edge information in the intensity data, and the polarimetric cue that is based on the similarity of coherency matrices are utilized. The partitioning problem is solved using the multiclass spectral clustering algorithm. This method is a two-step technique – similar to S3 – that employs S2 for over-segmentation, followed by a variant of the scheme S4. The similarity of coherency matrices was measured using the Symmetric Revised Wishart distance between segments, and the value of the scaling parameter was chosen adaptively using \textit{local scaling} as it was the case in S1. This methodology is conceptually very similar to S3 since over-segmentation ensures a high level of detail, and merging segments in a separate step addresses the issues that arise in S2. However, the grouping performed in the second step of the proposed method is not limited to the adjacent segments. As a result, the classification task is completed. This method also offers a more effective way of utilizing the complete polarimetric information available in the coherency matrix. It was tested on two data sets from two airborne sensors (i.e., Convair-580, AIRSAR) operating at two frequency bands (i.e., C- and L-bands). The results for two data sets were obtained using the same set of parameter values and compared to the Wishart classifier in terms of overall classification accuracy and Kappa coefficient. This approach offered quantitative and qualitative improvements over the Wishart classifier and the schemes we have discussed before. A direct comparison with our previous work \cite{51} shows that the proposed method has better performance than S1. Also, an indirect comparison can be done with Anfinsen \textit{et al.} \cite{52} since their methodology achieved the same level of accuracy as the Wishart classifier for the same subset of the Flevoland data set.

Finally, Chapter \cite{6} summarizes our findings and draws conclusions regarding the effectiveness of the tested schemes and the proposed method, outlines the research contributions of this work, and suggests future work in this area.
Chapter 2

Spectral Graph Partitioning

Spectral graph partitioning is a promising methodology to solve grouping problems in computer vision, where the performance level of humans is the desired outcome. This methodology has been used to incorporate multiple cues that represent the perceptual organization ability of the human vision system. It has been shown to achieve perceptually plausible results, which motivates its use for POLSAR data analysis.

The grouping problem is mapped onto a graph partitioning problem using an appropriate affinity function, which is used to form a pairwise similarity matrix. The information about the data points is encoded in this matrix. The eigenvalue decomposition of this matrix can be used to obtain an efficient solution for the global optimization of the objective function (i.e., normalized cut). An efficient solution refers to solving the problem without using brute force (i.e., trying all combinations of candidate partitions). It does not necessarily mean obtaining a solution that is computationally inexpensive. Since the technique involves pairwise affinity calculations, the similarity matrix is of size $N \times N$ where $N$ is the number of data points, and the eigenvalue analysis of such a matrix is $O(N^3)$. There are available techniques that circumvent the computational aspects of this problem \[53\], \[47\].

In Section 2.1, the graph partitioning concept is explained and two grouping algorithms (i.e., spectral clustering and multiclass spectral clustering) are discussed. The techniques that can be used to reduce the computational complexity of SGP are explained in Section 2.2.
2.1 Graph Partitioning with Normalized Cuts

Both clustering and image segmentation can be formulated as a graph partitioning problem. The formulation involves representing a set of points in an arbitrary feature space using an undirected graph $G = \{V, E\}$, where $V$ and $E$ represent the nodes and the edges (i.e., connections) respectively. Each node on the graph corresponds to a data point in a feature space and the edge between two nodes, $u$ and $v$, is associated with a weight, $\omega(u, v)$, that indicates the similarity of that pair. In general, $G$ is a fully connected graph (i.e., each node is connected to all the other nodes). For bi-partitioning such a graph, it is intuitive to minimize the similarity of candidate partitions (i.e., cut). The similarity of candidate partitions $V_1$ and $V_2$ is quantified by the sum of the weights associated with the edges between the nodes in $V_1$ and the nodes in $V_2$. In (2.1) the summation is taken over the total number of edges between the graphs $V_1$ and $V_2$.

$$\text{cut}(V_1, V_2) = \sum_{u \in V_1, v \in V_2} \omega(u, v)$$  \hspace{1cm} (2.1)

However, minimizing the cut objective function favors isolated nodes as partitions. This bias can be avoided by minimizing the normalized cut ($\text{ncut}$):

$$\text{ncut}(V_1, V_2) = \frac{\text{cut}(V_1, V_2)}{\text{assoc}(V_1, V)} + \frac{\text{cut}(V_1, V_2)}{\text{assoc}(V_2, V)}$$  \hspace{1cm} (2.2)

where $\text{assoc}(V_1, V)$ is the sum of the weights between the nodes in $V_1$ and all the nodes in the graph, $V$, i.e.

$$\text{assoc}(V_1, V) = \sum_{u \in V_1, t \in V} \omega(u, t)$$  \hspace{1cm} (2.3)

Shi and Malik have shown that an optimal partitioning can be obtained by minimizing the ncut cost function [15]. They have also shown that this minimization problem is equivalent to find $y$ that minimizes the following:

$$\frac{y^T(D - W)y}{y^TDy}$$  \hspace{1cm} (2.4)
2.1. Graph Partitioning with Normalized Cuts

where \( y = \{a, b\}^N \) is a binary indicator vector specifying the group identity for each data point (i.e., \( y_i = a \) if node \( i \) belongs to \( V_1 \) and \( y_j = b \) if node \( j \) belongs to \( V_2 \)). \( N \) is the number of nodes in the graph, \( W \) is the \( N \times N \) matrix whose entries are the weights \( \omega(i, j) \), and \( D \) is a diagonal matrix, whose diagonal elements are the sum of the rows of \( W \) (2.5):

\[
D_{ii} = \sum_{j=1}^{N} W_{ij}
\]

(2.5)

Note that the expression in (2.4) is the Rayleigh quotient, and if the condition on \( y \) is relaxed so that it can take any real value, the solution can be obtained by solving the generalized eigenvalue system,

\[
(D - W)y = \lambda D y
\]

(2.6)

where \( D - W \) is known as the graph Laplacian. Using \( z = D^{-\frac{1}{2}}y \), (2.6) can be rewritten as:

\[
D^{-\frac{1}{2}}(D - W)D^{-\frac{1}{2}}z = \lambda z
\]

(2.7)

Thus, the minimization of the \( \text{ncut} \) cost function is approximated by solving the generalized eigenvalue system for the normalized graph Laplacian:

\[
\mathcal{L} = D^{-\frac{1}{2}}(D - W)D^{-\frac{1}{2}}
\]

(2.8)

The eigenvector that corresponds to the second smallest eigenvalue is the real valued solution for (2.4). Therefore, this eigenvector can be used to bi-partition the graph [45]. Shi and Malik have also used this approach recursively to obtain more than two partitions.

Also note that the second smallest eigenvalue of \( \mathcal{L} \) corresponds to the second largest eigenvalue of \( I - \mathcal{L} \) or \( D^{-\frac{1}{2}}W D^{-\frac{1}{2}} \), where \( I \) is the identity matrix. This relation will be useful Section 2.1.1.
2.1. Graph Partitioning with Normalized Cuts

2.1.1 Spectral Clustering Algorithm

Based on the spectral graph partitioning approach described in the previous section, the spectral clustering algorithm given by Ng et al. [54] provides k-way partitioning. A slightly different notation is used by Ng, where the affinity matrix $A$ is employed instead of the similarity matrix $W$. Therefore, $D_A$ given in (2.11) will replace $D$ defined in (2.5). Also $I - L$ ($I$ is the identity matrix) is replaced with the normalized affinity matrix $L$ defined in (2.10). Thus, the eigenvectors that correspond to the largest eigenvalues are used instead of the smallest ones. For a set of points, $S = \{s_1, ..., s_N\}$ in $R^\ell$, and number of clusters, $k$, the algorithm involves the following steps:

1. Form the affinity matrix $A \in R^{N \times N}$ whose entries are defined by:

$$A_{ij} = \begin{cases} \exp\left\{ -\frac{d^2(s_i, s_j)}{2\sigma^2} \right\}, & \text{if } i \neq j \\ 0, & \text{if } i = j \end{cases} \quad (2.9)$$

where $\sigma$ is the scaling parameter (a.k.a., kernel bandwidth) and $d(s_i, s_j)$ is the Euclidian distance and represents the dissimilarity between points $s_i$ and $s_j$.

2. Construct the normalized affinity matrix, $L$ using (2.10):

$$L = D_A^{-\frac{1}{2}} A D_A^{-\frac{1}{2}} \quad (2.10)$$

where $D_A$ is a diagonal matrix and its elements are defined as:

$$D_{Aii} = \sum_{j=1}^{N} A_{ij} \quad (2.11)$$

3. Form matrix $X = [x_1 x_2 ... x_n ... x_k] \in R^{N \times k}$, whose columns $x_n$ are the eigenvectors that correspond to the $k$ largest eigenvalues of matrix $L$.

4. Normalize the rows of $X$ and form matrix $Y$ using: $Y_{ij} = \frac{X_{ij}}{\sqrt{\sum_j X_{ij}^2}}$

5. Cluster rows of $Y$ using the k-means algorithm
Since pairwise similarities are used to determine the groups, it becomes possible to recover complicated manifold structures in the feature space, which can not be achieved by central grouping techniques (e.g., k-means or EM) that require each group member to be close to a prototype (i.e., the cluster center). A synthetic data set that contains concentric circles in a two-dimensional feature space is often used to demonstrate the ability of spectral clustering to overcome the limitation of central grouping techniques and correctly solve the perceptual grouping problem (i.e., group points as two circles like humans do). However, the value of the scaling parameter ($\sigma$) plays an important role in obtaining successful results.

One approach to find an optimum value for the scaling parameter is to perform a grid search in the space of the hyperparameter and evaluate a measure of success (e.g., classification accuracy). This procedure however, requires supervision.

The potential problems due to manual parameter tuning can be circumvented by using local scaling as suggested in [55]. This procedure automates the parameter selection and provides good results by adaptively choosing a scaling parameter $\sigma_i$ for each point $s_i$. The affinity function for $i \neq j$ given in (2.9) is modified as follows:

$$\hat{A}_{ij} = \exp\left(\frac{-d^2(s_i, s_j)}{2 \sigma_i \sigma_j}\right) \quad (2.12)$$

and $\sigma_i$, the local scaling parameter for point $s_i$, is given by

$$\sigma_i = \text{median}_{s_n \in N(s_i)} \{d(s_i, s_n)\} \quad (2.13)$$

where $N(s_i)$ is the set of $n_{LS}$ nearest neighbors of $s_i$ in the feature space.

**A proof of concept example**

To demonstrate that the spectral clustering algorithm can solve the perceptual grouping problem, a synthetic data set that contains three concentric circles in two-dimensional feature space is used. This example is shown in Figure 2.1 where the K-means algorithm fails to group the points correctly.
(i.e., as perceived by humans). However, spectral clustering provides a way to achieve a result that is similar to what humans perceive. This example is not a representative of POLSAR data. However, it aims to illustrate that spectral clustering can achieve correct grouping even for a rather unusual distribution and when the information about that distribution is unknown.

The results obtained for two different values of the scaling parameter \( \sigma \), show that this parameter plays an important role in obtaining successful results.

![Figure 2.1: Three-circle data: K-means vs. Spectral Clustering (\( \sigma^2 = 0.1 \) and \( \sigma^2 = 0.05 \))](image)

Successful results were shown in [55] without manually setting a value for the scale parameter \( \sigma \). However, the number of nearest neighbors used in the estimation procedure, \( n_{LS} \), was arbitrarily set to 5 (i.e., \( n_{LS} = 5 \)). Therefore, we investigated the effect of the neighborhood size, \( n_{LS} \). The results in Figure 2.2 suggest that setting \( n_{LS} \), which defines a local scaling parameter automatically, has an effect similar to changing the scaling parameter manually. A large value of \( n_{LS} \) results in a high scaling parameter estimate, causing confusion among the groups.

### 2.1.2 Multiclass Spectral Clustering Algorithm

Spectral methods are appealing because their global-optima in the relaxed continuous domain are obtained by eigen-decomposition. However, getting a discrete solution from eigenvectors often requires solving another clustering problem, albeit in a lower-dimensional space. That is, eigenvectors are treated as geometrical coordinates of a point set. A clustering heuristic
2.1. Graph Partitioning with Normalized Cuts

Figure 2.2: Three-circle data: Spectral Clustering with local scaling $n_{LS} = 5, 50, 250$

such as K-means is subsequently employed on the new point set to retrieve partitions \[54\]. Yu has shown that there is a principled way to recover a discrete optimum \[56\]. This is based on a fact that the continuous optima consist not only of the eigenvectors, but of a whole family spanned by the eigenvectors through orthonormal transforms. The goal is to find the right orthonormal transform that leads to a discretization.

Yu’s method has the following steps:

1. Solve a relaxed continuous optimization problem. The global optima are given by some eigenvectors subject to arbitrary orthonormal transforms.

2. Iteratively solve for a discrete solution that is closest to the continuous optima using an alternating optimization procedure. The algorithm alternates the following: (1) the continuous optimum closest to a discrete solution is located by computing the best orthonormal transform, and (2) the discrete solution closest to a continuous one is located by non-maximum suppression. Such iterations monotonously decrease the distance between a discrete solution and the continuous optima. After convergence, an approximation to the globally optimal partitioning is obtained.
2.2 Techniques for Reduced Computational Complexity

The spectral graph partitioning framework involves solving the eigenvalue problem for a similarity matrix, of size $N \times N$, where $N$ is the number of points (or pixels for image segmentation problems). In the case of a fully connected graph, all the entries of this matrix are potentially non-zero, and the time complexity of solving the eigenvalue problem is $O(N^3)$. The computational cost of this dense solution quickly becomes prohibitive for images of useful size. The approaches summarized in Sections 2.2.1 to 2.2.4 address this problem by reducing the time and space complexity at the expense of obtaining approximate solutions.

2.2.1 Iterative Methods

Spectral clustering algorithm requires only the first few eigenvectors. Therefore, an iterative method (e.g., Lanczos [53]) can be used to obtain the solution with the complexity of $O(N^2)$ in the case of a fully connected graph.

2.2.2 Sparse Representation

Another approach to obtain a fast solution is to use a sparse representation instead of the dense $N \times N$ matrix. For example, one can explicitly form the similarity matrix, sort the entries and zero out the values that are very small. This approach can be quite accurate, but it requires $O(N^2)$ affinity calculations in the first place. The alternative is to zero out random entries in the matrix, although the effect of this is not well known.

In the context of image segmentation, one could also choose to have zero weights (i.e., affinities) for pixel pairs that are not within a neighborhood. This neighborhood can be specified either by a window mask or a radius measured from the pixel of interest. As a result, the number of connections from each pixel will be limited to a small value, and the similarity matrix will be sparse. Since the resulting graph will have very few connections in comparison to the fully connected case the complexity is reduced to $O(N)$. 

25
For these representations, the eigensolvers that exploit the sparsity structure of the matrix will provide reduced complexity. The iterative method Lanczos can be then used on the sparse matrix with reduced complexity.

### 2.2.3 The Nyström Extension Method

The Nyström extension is a technique that finds numerical approximations for eigenvalue problems. It extrapolates the eigenvectors that were computed using a small subset of sample points to the full size problem. A random selection of the subset has been found to work effectively in [47].

In the context of solving the eigenvalue problem for a pair-wise similarity matrix, $L$, of size $N \times N$, this turns out to be very useful when $N$ is large. If the matrix $L$ is partitioned into blocks as follows:

$$L = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}$$

where $A$ is the $n \times n$ similarity matrix formed using $n$ randomly sampled points; $B$ represents the $n \times m$ matrix of affinities between $n$ sample points and $m$ remaining points; and $C$ is the $m \times m$ matrix that contains the weights between all the remaining points. When $n \ll N$, $A$ will be of manageable size. Therefore it can easily be decomposed into its eigenvalues and eigenvectors, such that $A = U \Lambda U^T$.

Let $\bar{U}$ denote the approximate eigenvectors of $L$ that are given by the Nyström extension: $\bar{U} = \begin{bmatrix} U & B^T U \Lambda^{-1} \end{bmatrix}^T$. Then the approximation of $L$, denoted by $\bar{L}$, will be:

$$\bar{L} = \bar{U} \Lambda \bar{U}^T = \begin{bmatrix} A & B \\ B^T & B^T A^{-1} B \end{bmatrix}$$

Equations (2.14) and (2.15) show that the Nyström extension implicitly approximates the $C$ block with the matrix $B^T A^{-1} B$. The accuracy of this approximation can be quantified by $\|C - B^T A^{-1} B\|$. Since the columns of $\bar{U}$ are not orthogonal, the following steps should be performed to orthogonalize the eigenvectors:
2.2. Techniques for Reduced Computational Complexity

1. Let $A^\frac{1}{2}$ denote the square root of matrix $A$.
   Define $S = A + A^\frac{1}{2}BB^T A^\frac{1}{2}$;

2. Diagonalize $S$, such that $S = U_S \Lambda_S U_S^T$;

3. Form $V$ using: $V = \begin{bmatrix} A & B \end{bmatrix}^T A^\frac{1}{2} U_S \Lambda_S^\frac{1}{2} S$.

It can be shown that $\hat{L} = V \Lambda_S V^T$. Therefore the columns of $V$ are the orthogonal eigenvectors of $\hat{L}$ and they approximate the eigenvectors of $L$.

The Nyström extension is useful for obtaining approximate solutions for large spectral clustering problems. To do that, it is necessary to compute $\hat{d}$, the sum of the rows of $\hat{L}$. This is possible without explicitly evaluating the $B^T A^{-1} B$ block. Having $\hat{d}$ in hand, the required blocks of $\hat{D}^{-\frac{1}{2}} \hat{L} \hat{D}^{-\frac{1}{2}}$ are obtained and the orthogonalized approximate eigenvectors are computed as described above.

2.2.4 Other Techniques for $O(N^2)$ Problems

$O(N^2)$ complexity exists in many machine learning techniques (e.g., Gaussian processes, radial basis networks, and support vector machines). It also arises in the nearest neighbor search and when the max operator is used. All of these involve either the Sum-Product or the Max-Product problem. Techniques to speed-up the solution for these problems have been studied in the last few years. Lang [57] defines the problem as follows:

For a given set of points $X = \{x_i\}, \ i = 1, ..., N$, called source points, there is a set of points $Y = \{y_j\}, \ j = 1, ..., M$, called target points at which we need to measure a quantity of interest (i.e., influence of a source point on a target point). So, the Sum-Product problem is:

$$f_j = \sum_{i=1}^{N} w_i K_i(||x_i - y_j||), \ j = 1, ..., M \quad (2.16)$$

where $f_j$ is the influence at point $y_j$, $w_i$ is the weight, and $K_i$ is the kernel function for point $x_i$. This problem is called weighted kernel density estimation. In our case, there is a single kernel function, sources and targets are the same set of points and $M = N$. Therefore the complexity is $O(N^2)$.  

27
2.3. Implementation and Runtime Analysis of the Algorithms

Fast N-body Methods

To speed-up the Sum-Product problem, there are several methods available in the literature (e.g., Fast Multipole Method (FMM), Fast Gaussian Transform (FGT), Improved FGT, and Dual-Tree methods). A brief introduction to these methods and their performance comparison can be found in [57]. Among these approaches Dual-Tree techniques stand out, since they are easily applicable to different types of kernels and they use adaptive space partitioning.

Gray and Moore [58] introduced Dual-Tree recursion, where the tree is expanded in such a way that only areas that contain useful information are explored. The algorithm works as follows: given a node from each of the \( X \) and \( Y \) trees (in our case they are identical), the upper and lower bounds on the influence of all points in the \( X \) node upon all points in the \( Y \) node are computed. The influence bounds are tightened by expanding the nodes in \( X \) and \( Y \) into their children. Since the children nodes are more compact, the distance bounds get tighter, which leads to tighter influence bounds. As a result, the difference between the upper and lower influence bounds becomes smaller and the error is reduced. The tree expansion stops when a specified error bound, \( \epsilon \), is achieved. At this step, if the node pair contains \( N_X \) and \( N_Y \) points, the interaction calculation, with a known error bound, is achieved by a single operation. This is a significant reduction compared to the naïve approach which requires \( N_X \times N_Y \) operations. The dual tree recursion algorithm briefly explained above has been used for several applications by Dr. Nando de Freitas and his colleagues (Dept. of Computer Science, UBC) [57, 59].

2.3 Implementation and Runtime Analysis of the Algorithms

Spectral graph partitioning methodologies that will be discussed in detail in Chapter 4 were implemented in Matlab and C. Some subroutines were written in C and complied into Matlab executables (MEX files). The imple-
mentation was performed by Kaan Ersahin, i.e., the author of this thesis. The Wishart classifier uses Matlab code that was initially written by Bernd Scheuchl and was later modified by Kaan Ersahin. Scheuchl and Ersahin are the co-authors of [44] and were members of the UBC Radar Remote Sensing Group.

The proposed method for POLSAR classification, which will be discussed in Section 4.2, uses the spatial proximity cue via sparse representation of the pairwise similarity matrix. As a result computational cost and memory requirement are both reduced. Both the proposed method and the Wishart classifier were tested on the Flevoland subset data (200 × 320 pixels) presented in Section 5.1. Tests were performed on an IBM workstation running Windows XP, Pentium IV - 3.6 GHz, 2 GB RAM. The proposed method uses block processing to reduce memory requirement, a total of 16 blocks (50x80 pixels each) were used for segmentation (20 sec/block), including pre-processing data into blocks and classification step based on segments resulting from first step, the total runtime for the entire subset was 355 seconds. The Wishart classifier on the other hand, is iterative and on average took 30 sec per iteration for the entire subset. Stopping criteria used was the percentage of change in pixel labels. 10 iterations achieved about 5% and 30 iterations were necessary to achieve 1%. As a result, depending on the convergence requirement it took between 300 to 1500 seconds to process. These tests, which were performed using MATLAB’s Profiler utility, suggest that the runtime for the two algorithms are in the same order of magnitude. Although the implementations were not optimized, these results are consistent with and complementary to the theoretical complexity associated with these techniques, i.e., \(O(N)\).

2.4 Summary

In this chapter, the spectral graph partitioning methodology and two grouping algorithms based on minimizing the normalized cut objective function have been discussed. The first algorithm is the spectral clustering algorithm of Ng et al. [54]. An adaptive selection procedure for the value of the
2.4. Summary

scaling parameter (i.e., local scaling) suggested by Zelnik-Manor [55] is also presented. A proof of concept example using the “three-circle” data set is provided for both fixed values of the scaling parameter and using the local scaling procedure. The second technique is the multiclass spectral clustering algorithm of Yu [60], where a principled way is suggested to recover a discrete optimum as opposed to a clustering heuristic such as K-means used by Ng et al. [54].

To reduce the computational complexity of the spectral graph partitioning technique, four different approaches are discussed: iterative methods, sparse representations, the Nyström extension method, and fast N-body methods. In this work an iterative technique, i.e., Lanczos, is used in the MATLAB package that performs the eigenvalue decomposition of large, sparse matrices, which reduces the complexity from $O(N^3)$ to $O(N^2)$. The spatial proximity cue – one of the four cues that are explained in Chapter 3 – results in a sparse representation of the similarity matrix, which reduces the complexity to $O(N)$. The Nyström extension method explained in this chapter is also used in one of the schemes explained in Chapter 4. Fast N-body methods, were studied as a potential approach for computational complexity reduction.
Chapter 3

Cues for Polarimetric SAR Data

In Chapter 1, the discussion on background technology in polarimetric radar data analysis, and an emerging technique in computer vision that can utilize some ideas from perceptual organization, was used to motivate using the spectral graph partitioning for segmentation and classification of polarimetric SAR data. In Chapter 2, spectral graph partitioning methodology was formally introduced and two algorithms (i.e., spectral clustering and multiclass spectral clustering) were explained. Also, different approaches for reducing the computational complexity were discussed. In the process of developing new schemes for POLSAR data classification, the performance level of humans is aspired to through the use of four cues: patch-based similarity (i.e., intensity cue), contour cue, spatial proximity cue, and similarity of coherency matrices (i.e., the polarimetric cue). In this chapter, these cues are explained in detail, some of which are customized for POLSAR data, and used as building blocks of the tested schemes and the proposed method that will be introduced in Chapter 4 and Appendix B.

3.1 Patch-based Similarity: Intensity Cue

Spectral graph partitioning framework requires us to define pairwise similarities. However some special properties (e.g., presence of speckle) of SAR image data do not allow one to use pixel-to-pixel comparisons. Instead, a patch (i.e., a group of pixels) is more likely to correctly represent the underlying information and is more robust to variations due to speckle.

To compare two patches statistically and measure the similarity between
them we used the $\chi^2$ distance between the intensity distributions (i.e., histograms) of those two patches [51]. The patches were formed using the edge-aligned window\(^9\) mask associated with the pixel of interest. These masks are determined using the procedure described in Lee et al. [17].

Below is a more formal description of how to calculate the affinity matrices for each of the input channels (i.e., $|HH|^2$, $|HV|^2$, and $|VV|^2$).

1. For a pixel of interest $s_i$, the histogram $h_i$ is computed from the pixels that fall in the patch defined by an edge-aligned window mask. This window mask is determined using the procedure described in Lee et al. [17] for polarimetric SAR speckle filtering.

2. Using the chi-squared distance defined in (3.1) calculate the dissimilarity between the histograms $h_i$ and $h_j$ that are associated with pixels $s_i$ and $s_j$, respectively:

$$
\chi^2(h_i, h_j) = \frac{1}{2} \sum_{m=1}^{M} \frac{[h_i(m) - h_j(m)]^2}{h_i(m) + h_j(m)}
$$

(3.1)

3. As per the definition of the local scaling parameter (2.13), $\sigma_i$ is modified as follows:

$$
\sigma_i = \text{median}_{h_n \in N(h_i)} \left\{ \chi^2(h_i, h_n) \right\}
$$

(3.2)

where $N(h_i)$ is the set of $n_{LS}$ nearest neighbors of $h_i$.

4. Form a similarity matrix for each input channel independently using the following affinity function:

$$
W_{ij} = \exp \left( -\frac{\chi^2(h_i, h_j)}{2 \sigma_i \sigma_j} \right)
$$

(3.3)

\(^9\)Edge-aligned windowing scheme is used to exclude non-relevant parts of the content in a square window which will possibly contain a boundary and is not homogeneous.
3.2 Contour-based Similarity Cue

Contour detection is a useful tool for segmentation of image data, where the attributes of the data points in the feature space such as brightness or color do not contain sufficient information to allow separable clusters to be obtained.

Most contour detection methods rely on an edge detection step followed by a process that links broken pieces of edges. Edge detection has the drawback that decisions are made locally i.e., without considering the information from the entire image. In Leung and Malik’s approach the contour information is computed locally, but decisions are made via a global optimization technique where the information from the entire image is considered.

Information about the strength of a contour or an edge is obtained through orientation energy. Let $F_1(x,y)$ be the second derivative of an elongated Gaussian kernel and $F_2(x,y)$ be its Hilbert transform, i.e.:

$$F_1(x,y) = \frac{d^2}{dy^2} \left( \frac{1}{C} \exp\left\{ \frac{y^2}{\sigma^2} \right\} \exp\left\{ \frac{x^2}{\lambda^2 \sigma^2} \right\} \right)$$

$$F_2(x,y) = \text{Hilbert} \{ F_1(x,y) \}$$

where $C$ is a constant, $\sigma$ is the scale and $\lambda$ is the elongation of the kernel. The orientation energy at angle $0^\circ$ is defined as:

$$OE_{0^\circ} = \{ I * F_1(x,y) \}^2 + \{ I * F_2(x,y) \}^2$$

where $I$ is the image data and $*$ is the convolution operator.

The filter output $OE_{0^\circ}$ has maximum response for horizontal contours. Rotated copies of the second derivative of an elongated Gaussian and its Hilbert transform are able to pick up edge contrast at various orientations. These kernels are shown in Figure 3.1.

The orientation energy, $OE(x,y)$ of pixel $(x,y)$ is the maximum value of the orientation energy values calculated at multiple orientations, $\phi$:

$$OE(x,y) = \max_{\phi} OE_{\phi}(x,y)$$

33
To calculate the orientation energy of the image $I$ at a specific angle, e.g., $0^\circ$, a convolution is performed with odd-symmetric and even-symmetric filters, obtained using the difference of an offset Gaussian (DOOG) kernel (For details see [61]). The filter parameters used are: filter size, number of orientations ($ori$), elongation ($\lambda$), and scale ($\sigma$).

Since the filters are elongated, the definition of the orientation energy given by (3.7) allows the integration of information along the edge. As a result, the extended contours (even those have low contrast) stand out with respect to high-contrast contours that are short.

Based on the contour information, the similarity of two pixels can be defined. To compute this similarity, first we draw a line to join the two pixels, then calculate the orientation energy values for all the pixels that lay on this line, and finally take the maximum value among those orientation energy values to obtain the distance between the pair. The similarity is inversely proportional to the pairwise distance. Intuitively, if there is an extended contour (such as a boundary between fields of different crops), that is crossing between a pair of pixels, leaving one pixel on either side, these two pixels should belong to different partitions.

Figure 3.2 illustrates this concept, where $s_1$, $s_2$, and $s_3$ have similar intensity values. Typically in region-based segmentation, intensity is the only information source. Since these pixels have almost the same intensity values, their pairwise similarities are high. However, there is an extended contour separating $s_3$ from $s_1$ and $s_2$. Thus, we expect $s_1$ to be much more strongly related to $s_2$ than to $s_3$. This intuition carries over to the
3.2. Contour-based Similarity Cue

Figure 3.2: Orientation energy as a measure of contour information. Adapted from [48].

definition of similarity based on contour information: If the maximum value of orientation energy on the line joining two pixels is high – suggesting the presence of an extended contour – then the similarity of that pair should be low.

Formally, the dissimilarity of two pixels based on contour information, \( d_c(s_i, s_j) \), is defined as:

\[
d_c(s_i, s_j) = \text{OE}(\hat{x}) \quad (3.8)
\]

\[
\hat{x} = \arg \max_{x \in l} \text{OE}(x) \quad (3.9)
\]

where \( l \) is the line joining \( s_i \) and \( s_j \); and \( \hat{x} \) is the location where the orientation energy is maximum along \( l \).

Using this distance measure, the pairwise similarity is obtained as:

\[
W_{ij}^C = \exp \left\{ -\frac{d_c^2(s_i, s_j)}{2 \sigma_c^2} \right\} \quad (3.10)
\]

where \( \sigma_c \) is the scaling parameter for the kernel. The value of this parameter is needed in the similarity matrix computation. As suggested in [56], this value can be determined by multiplying the maximum value of the orientation energy in the region of interest with a constant \( (ev) \) representing the edge variance, whose value is fixed for the entire image:

\[
\sigma_c = ev \cdot \max(\text{OE}) \quad (3.11)
\]
3.3 Similarity of Coherency Matrices: Polarimetric Cue

The coherency matrix is one of the most concise ways of representing polarimetric information. The assumed statistical distribution of the multi-look processed coherency matrix leads to the maximum likelihood classification scheme based on the Wishart distribution (a.k.a. Wishart classifier).

Both the patch-based similarity cue and the contour cue, as outlined in the previous two sections, were designed to work on individual polarimetric channels, and were limited to the intensity information. In a classification scheme based on the spectral graph partitioning approach, utilizing the complete polarimetric information, through the coherency matrix, is desired. Therefore, the measurement of distances between coherency matrices is essential. However, the Wishart distribution results in a non-symmetric distance measure, i.e., \( d(A, B) \neq d(B, A) \), and thus cannot be used directly in the spectral clustering algorithm, which requires the similarity matrix \( W \) to be symmetric. Two well-known matrix distance measures are: (i) Frobenius distance, and (ii) Euclidean distance (or \( \ell_2 \) metric). However, these cannot be related to the Wishart distribution.

To solve the spectral clustering problem based on pairwise distances between coherency matrices, Anfinsen et al. [52] suggested the use of two distance measures: the Bartlett distance and the Symmetric Revised Wishart distance. A brief explanation of these two distance measures is provided in Sections 3.3.1 and 3.3.2.

We define the polarimetric cue based on the Symmetric Revised Wishart distance. Polarimetric cue is used in the second step of the proposed scheme which will be presented in Section 4.2.

3.3.1 Bartlett Distance

Conradsen et al. [62] considered two sample coherency matrices \( A \sim W(n, \Sigma_A) \) and \( B \sim W(n, \Sigma_B) \), generated from unknown and potentially different Wishart densities, and the likelihood ratio hypothesis test of \( H_0 : \Sigma_A = \Sigma_B \)
3.3. Similarity of Coherency Matrices: Polarimetric Cue

versus $H_1: \Sigma_A \neq \Sigma_B$. Under the null hypothesis, the distributions of $A$ and $B$ are equal, and it follows from the properties of the Wishart density that $A + B \sim W(2n, \Sigma)$, where $\Sigma = \Sigma_A = \Sigma_B$. Using the likelihood functions under $H_0$ and $H_1$, the likelihood ratio test statistic becomes (See [62] for details):

$$Q_1 = 2^{2nq} \frac{|A|^n |B|^n}{|A + B|^{2n}}$$  \hspace{1cm} (3.12)

Using this test statistic, the following distance measure is defined:

$$d_B(A, B) = -\frac{\ln Q_1}{n}$$ \hspace{1cm} (3.13)

$$= \ln \left( \frac{|A + B|^2}{|A||B|} \right) - 2q \ln 2$$ \hspace{1cm} (3.14)

Without the constant term, this measure is called the Bartlett distance Kersten et al. [32].

Using a Gaussian kernel, an affinity function can be defined to form the corresponding similarity matrix, $W^B$, given by:

$$W^B_{ij} = \exp \left( -\frac{d_B^2(T_i, T_j)}{2 \sigma^2} \right)$$ \hspace{1cm} (3.15)

where $T_i$ and $T_j$ represent the coherency matrices for instances $i$ and $j$, respectively.

3.3.2 Symmetric Revised Wishart Distance

Using the two sample coherency matrices defined in Section 3.3.1, a second distance measure can be derived from the likelihood ratio test, if we assume the same hypotheses $H_0$ and $H_1$, given $\Sigma_B$ is known. The test statistic then becomes:

$$Q_2 = \frac{|A|^n}{|B|^n} \exp \left\{ -n(tr \{ B^{-1} A \} - q) \right\}$$ \hspace{1cm} (3.16)

Using this test statistic, the following distance measure, the Revised Wishart distance [32], is defined as:

$$d_{RW}(A, B) = -\frac{\ln Q_2}{n}$$ \hspace{1cm} (3.17)

37
3.4. Spatial Proximity Cue

\[ d_{RW} = \ln \frac{|B|}{|A|} + \text{tr}\{B^{-1}A\} - q \]  \hspace{1cm} (3.18)

Although, \( d_{RW} \) is not symmetric, a symmetric measure can be obtained by:

\[ d_{SRW}(A, B) = \frac{1}{2} (d_{RW}(A, B) + d_{RW}(B, A)) \]  \hspace{1cm} (3.19)

\[ = \frac{1}{2} \text{tr}(AB^{-1} + BA^{-1}) - q \]  \hspace{1cm} (3.20)

The Symmetric Revised Wishart distance is also a semi-metric, since the triangle inequality (i.e., \( d(A, C) \leq d(A, B) + d(B, C) \)) is not satisfied. Using a Gaussian kernel, an affinity function can be defined to form the corresponding similarity matrix, \( W^{SRW} \), given in (3.21).

\[ W^{SRW}_{ij} = \exp \left( \frac{-d^2_{SRW}(T_i, T_j)}{2\sigma^2} \right) \]  \hspace{1cm} (3.21)

where \( T_i \) and \( T_j \) represent the coherency matrices for instances \( i \) and \( j \), respectively.

3.4 Spatial Proximity Cue

Spatial proximity cue can be represented through two affinity functions, i.e., ramp function and step function. Ramp function allows the similarity to be weighted according to the distance between pixels in the image plane. On the other hand step function works as a logic operator and is used to gate the other cues.

3.4.1 Ramp Function

A similarity matrix that represents the spatial proximity of pixel pairs in the image domain can be defined using the following:

\[ W^P_{ij} = \begin{cases} 1 - \frac{||l_i - l_j||^2}{r}, & \text{if } ||l_i - l_j||^2 < r \\ 0, & \text{if } ||l_i - l_j||^2 \geq r \end{cases} \]  \hspace{1cm} (3.22)
where \( l_i \) represents the location of pixel \( i \) in the image plane and \( r \) is the radius of a circular neighborhood outside which the pixel pairs will have zero affinities.

### 3.4.2 Step Function

Spatial proximity information can also be represented by limiting the number of connections that a node can have (i.e., the number of non-zero values in a single row of the similarity matrix). The spatial proximity cue is only meaningful when used along with another cue, and the combined similarity matrix is obtained by element-wise multiplication as will be discussed in Chapter 4.

Limiting the number of connections for each pixel due to the spatial proximity cue will result in less computation. This approach is essentially the same as that using a sparse representation for the similarity matrix as discussed in Section 2.2.2. In the current scenario, the neighborhood used to represent spatial proximity is a square window centered at the pixel of interest. The size of the window is \((2d + 1) \times (2d + 1)\) where \( d \) is the maximum distance that two pixels may have in one dimension (i.e., pairs that are separated more than \( d \) in either \( x \) or \( y \) directions will have no connection – zero affinity). Note that definition of \( d \) is different from the definition of radius \( r \), given in Section 3.4.1.

### 3.5 Summary

In this chapter, four different cues that utilize various aspects of the Polarmetric SAR data are explained. These are: the patch-based similarity (i.e., intensity cue), the contour cue, the similarity of coherency matrices (i.e., polarimetric cue), and the spatial proximity cue. Some of these cues have been inspired by perceptual organization and are customized for POLSAR data. Others were simply adopted from computer vision applications. These cues are used as the building blocks of the tested schemes and the proposed method that will be introduced in Chapter 4.
Chapter 4

The Proposed Method

In this chapter, we introduce the segmentation/classification method that we propose for operational POLSAR data, and explain the experimental schemes that we investigated on the way to defining the final recommended scheme. A roadmap, which represents the research progression over several years, is used to show the connection between the various schemes investigated. These schemes were used to assess the utility of either individual cues that were presented in Chapter 3 or their combinations. The analysis of experimental results from such schemes has shaped the proposed method. The results and discussion corresponding to these schemes and the proposed method will be presented in Chapter 5 and Appendix B.

4.1 Preliminary Research

In this section, we explain the research methodology that used various schemes and cues that were tested along the way to developing the proposed scheme. A roadmap is drawn to explain the research progression over several years.

Our work on the background technology included some pre-processing steps, such as, data decoding, multi-look processing, and speckle reduction. These were studied, implemented and tested to be used as building blocks of the presented schemes.

The potential of utilizing the scattering mechanism information for discrimination between crop types was also studied through both $H/A/\alpha$ and Freeman-Durden decompositions. Both approaches were found to provide good separation between vegetation and bare fields. However, they were not useful for discrimination between crop types in our case, since most of the
fields have dense vegetation, especially in the late season, which causes a dominant volume scattering component. If the grouping is performed according to the dominant scattering mechanism as suggested in [30], many crop types are grouped in a single superclass called “vegetated field”. Therefore, scattering mechanism information, which may be useful in general, was not found to be helpful for classification of agricultural fields in the Westham Island data set.

The original spectral clustering algorithm developed by Ng et al. [54] was modified to handle three input channels and it was tested for POLSAR data classification at the pixel level as a preliminary assessment of the proposed approach.

Since SAR data at pixel level has high variation even after speckle reduction, pixel-based similarity is replaced with a novel definition of patch-based similarity, and spatial proximity cue is also used. The resulting classification scheme (i.e., S1) utilizes both the intensity cue (i.e., patch-based similarity) and the spatial proximity cue, through an adaptation of the spectral clustering (SC) algorithm to POLSAR data. The limitations of this scheme are: (1) computationally expensive, i.e., for every pairwise similarity calculation, \( \chi^2 \)-distance calculation between patch histograms is performed. Although Nyström method is used to reduce the required number of pairwise similarity computations, each pairwise computation is still memory expensive; (2) Only the intensity information of each polarimetric channel is utilized, not the full polarimetric information; (3) A powerful visual cue in human vision, i.e., the contour cue, is not utilized.

To address the first and the third limitations discussed above, a segmentation scheme that utilizes the contour cue and the spatial proximity cue is introduced. This scheme (i.e., S2) utilizes the contour information by defining the pairwise similarities based on the orientation energy. And spatial proximity cue is integrated in the form of sparse representation, therefore helps to reduce the computational complexity. The partitioning is obtained using the multiclass spectral clustering (MSC) algorithm, which iteratively provides a discrete solution. In [60], Yu has shown that the solution after convergence is close to the global optimum. S2 does not solve the classi-
4.1. Preliminary Research

fication problem since non-adjacent fields with the same content are not labeled as a single class at this stage. There are also other limitations due to the definition of the orientation energy (e.g., fields that are non-convex in shape may be fragmented and adjacent fields separated with a pronounced boundary even if have the same content will be assigned different labels). In addition to these, this scheme still does not utilize the full polarimetric information.

Scheme S2, as it was presented in [63], did not use speckle filtering as a pre-processing step – it only used the three intensity channels from the polarimetric data. For consistency with the published material the results presented in Section 5.3.1 reflect this approach. However, further testing with speckle filtering was performed and visual comparisons were presented to complement the analysis. The following cases were discussed: (1) Assessing the utility of co-polarized correlation coefficient as an additional input (for a fixed value of $n_c$, number of partitions); (2) Assessing the utility of a variety of inputs when solved for different number of partitions, and analyze the effect of $n_c$ (number of partitions). In each case, results were presented with and without speckle filtering and results helped to shape the first step of the proposed method. These test results and discussion can be found in Section 5.3.2 and Section 5.3.3.

As an attempt to address some of the limitations of the first two schemes, i.e., S1 and S2, a two-step approach was suggested (i.e., S3). In the first step, the contour and proximity cues are employed to obtain an over-segmentation that ensures enough detail. Three intensity channels as well as co-polarized correlation coefficient was used with speckle filtering. In the second step, patch-based similarity based on $\chi^2$-distance between the adjacent segments is used to merge the fragmented segments and adjacent fields of the same content. Although, S3 conceptually solves some of the limitations and combined the strengths of S1 and S2, it still lacks the capability to utilize the information available in the fully polarimetric data.

The most common approach to utilize the full extent of the polarimetric information in SAR data is through the coherency matrix representation. However, the statistical analysis of such data leads to the Wishart distri-
4.1. Preliminary Research

bution and results in a non-symmetric distance measure, which is not suitable for the spectral graph partitioning approach that requires a symmetric pairwise similarity matrix. Anfinsen et al. [52] suggested the Symmetric Revised Wishart distance and Bartlett distance to classify POLSAR data at the pixel level using spectral clustering. Due to the computational complexity involved, they have resorted to solve the grouping problem for only a subset of pixels and used the results to initialize the Wishart classifier. Based on Anfinsen’s symmetric distance measures, in Section 3.3 we have introduced the polarimetric cue (i.e., similarity of coherency matrices) as an additional information source that could potentially be used with other cues.

Two test schemes were introduced to have a better understanding of the behavior of the polarimetric cue along with others. First test scheme, called S4, used polarimetric cue and the spatial proximity cue. And the second test scheme, i.e., S5, combined the polarimetric cue with the contour cue and the spatial proximity cue. These test schemes were only intended to visually assess the utility of such schemes as the weighting of the cues were changed through the selection of the scaling parameter. These test schemes are not suggested as alternative solutions to the classification problem, but they have contributed to shape the final form of the proposed technique in this thesis. The procedure followed for these test schemes are given in Appendix B along with the corresponding results and discussion.

All of the discussed schemes represent the research progression over several years and are presented to mark the road map that leads to the proposed method in this thesis. The design of the final method was influenced by the analysis of results obtained by the preceding schemes that are presented in Chapter 5 and Appendix B of this thesis.

The proposed method is a two-step approach to the POLSAR classification problem, conceptually similar to S3. First, groups of pixels with similar properties, i.e., segments, are obtained using the contour cue based on the edge information in the intensity data and spatial proximity cue. Then, these segments are grouped into classes using the polarimetric cue based on the similarity of mean coherency matrices of the segments, where the pair-
4.1. Preliminary Research

Figure 4.1: The progression of research ideas over several years, where vertical axis represents time. Each cue is shown with an ellipse, and rectangles represent tested schemes and the proposed method.

wise similarity is measured using the Symmetric Revised Wishart distance. In both steps the partitioning problem is solved using the multiclass spectral clustering (MSC) algorithm.

The following three subsections present Schemes S1, S2 and S3. The proposed method is presented in Section 4.2 and Schemes S4 and S5 are discussed in Appendix B.

Figure 4.1 shows the progression of research ideas over several years, where vertical axis represents time. Each cue is shown with an ellipse, and the rectangles represent tested schemes and the proposed method. The relation between tested schemes and the proposed method is summarized in this figure.

4.1.1 Scheme S1: Patch-based Similarity & Spatial Proximity

This section describes the method we proposed for classification of POLSAR data using spectral graph partitioning based on patch-based similarity and spatial proximity cues [51].
4.1. Preliminary Research

Figure 4.2 shows the block diagram for Scheme S1. Polariometric SAR data first goes through the pre-processing steps (i.e., multi-look processing and speckle filtering). Then, three channel powers are used as input to the patch-based similarity cue, followed by the spatial proximity cue. Finally, spectral clustering algorithm of Ng et al. \cite{54} is used to obtain the classification results.

This scheme adapts the spectral clustering algorithm of Ng et al.\cite{54} so as to accommodate POLSAR data. In a spectral graph partitioning framework classification of polarimetric SAR data is performed through modifications that were made to accommodate the intensity cue and the spatial proximity cue using the following steps:

1. Perform multi-look processing on single look complex (SLC) data (e.g., for Convair-580 data, use 10 azimuth looks and obtain square pixels, approximately 4m × 4m);

2. Apply the polarimetric SAR speckle filter of Lee et. al\cite{17} (e.g., window size = 7);

3. Apply the spectral clustering algorithm (Section 2.1.1) with the following modifications:

   (a) Form a similarity matrix for each data channel (i.e., $|HH|^2$, $|HV|^2$ and $|VV|^2$) independently using the procedure described in Section 3.1 using (3.1), (3.2) and (3.3).

   (b) Define a similarity matrix that represents the proximity of pixel...
pairs in the image domain using the following scheme:

\[ W^P_{ij} = \begin{cases} 1 - \frac{||l_i - l_j||_2}{r} & \text{if } ||l_i - l_j||_2 < r \\ 0 & \text{if } ||l_i - l_j||_2 \geq r \end{cases} \]  

(4.1)

where \( r \) is the maximum distance in the image plane that is allowed between pairs. (i.e., outside the neighborhood pairwise affinity will be zero); \( l_i \) and \( l_j \) represent the locations of pixels \( i \) and \( j \) on the image plane, respectively.

(c) Form the combined similarity matrix, \( W^{tot}_{ij} \), by element-wise multiplication of the affinity matrices as follows:

\[ W^{tot}_{ij} = W^P_{ij} \times \prod_{b=1}^{n_b} W^b_{ij} \]  

(4.2)

where \( n_b \) is the number of input channels.

(d) Perform Steps 2 to 5 of the spectral clustering algorithm as described in Section 2.1.1 (i.e., form matrix \( D_A \) and \( L \), calculate the first \( k \) eigenvectors of \( L \), form matrix \( X \), normalize its rows and perform clustering using the k-means algorithm).

4.1.2 Scheme S2: Contour-based Similarity & Spatial Proximity

This section describes the method we proposed for segmentation of POLSAR data using spectral graph partitioning based on contour information and spatial proximity [63].

Figure 4.3 shows the block diagram for Scheme S2. Polarimetric SAR data first goes through multi-look processing. Then, three channel powers are used as input to the contour-based similarity cue, followed by the spatial proximity cue. Finally, multiclass spectral clustering (MSC) algorithm of Yu [60] is used to obtain the segmentation results.

Instead of the patch-based similarity of intensities used in S1 that was explained in Section 4.1.1, this scheme utilizes the contour cue – a powerful
visual cue in perceptual organization. The following steps are performed:

1. Perform *multi-look processing* on the single look complex (SLC) data set;

2. Form a similarity matrix for each of the input data channels (i.e., $|HH|^2$, $|HV|^2$, $|VV|^2$) using (3.10). Limit the number of affinity calculations per pixel by choosing an appropriate neighborhood size. Outside this neighborhood the similarity due to spatial proximity will be “zero”. This will not only reduce the computational complexity, but also represent the spatial proximity cue. The proximity cue is incorporated using the sparse representation of the similarity matrix, which corresponds to use a step function to calculate the pairwise similarity matrix for the spatial proximity cue as discussed in Sections 2.2.2 and 3.4.2.

3. Form the combined similarity matrix, \( W_{\text{tot}} \), using:

\[
W_{ij}^{\text{tot}} = \prod_{b=1}^{n_b} W_{ij}^{C_b} \quad (4.3)
\]

where the subscript \( b \) is the index for input data channels, and \( n_b \) is the number of input channels;

4. Perform the multiclass spectral clustering (MSC) algorithm \([60]\) explained in Section 2.1.2 on the combined similarity matrix, \( W_{\text{tot}} \).
4.1. Preliminary Research

Figure 4.4: Block diagram for Scheme S3.

4.1.3 Scheme S3: Contour and Spatial Proximity Followed by Intensity Cue

This section describes the method we proposed for segmentation and classification of POLSAR data using spectral graph partitioning based on multiple cues: contour, spatial proximity, and patch-based similarity of intensity [64].

Using the segmentation technique presented in the previous section, which is based on contour information and spatial proximity, promising results are obtained that will be presented in Section 5.3.1. However, when S2 is used as a single-step solution, the definition of the pairwise distances that represent contour information through the orientation energy may lead to the following problems:

- Non-convex regions may result in multiple partitions (i.e., fragmented regions);
- Adjacent fields of same content (e.g., crop type) will likely be two separate partitions;
- Non-adjacent fields with the same crop type result in separate partitions.

The last two items can be addressed using an additional step to group the segments into classes. However, the first one requires further attention. For example, if the number of segments in a region of interest is either
known or fixed based on a desired level of detail, and a field is fragmented
(i.e., represented by more than one segment), some of the other fields – pos-
sibly with different content – will be merged into a single segment. This
can not be corrected at a later stage. Therefore, we propose to start with
an over-segmentation step using the contour-based technique presented in
Section 4.1.2, which can be used to ensure enough detail without forcing dif-
ferent classes to be merged. A second step can later group the segments that
are similar in content using a modified version of the patch-based similarity
approach presented in Section 4.1.1 In this scheme, each partition resulting
from the first step is represented as a node in a coarser graph, where only the
adjacent partitions are connected (i.e., have non-zero affinity). Figure 4.4
shows the block diagram for Scheme S3. This scheme follows the procedure
below:

1. Perform multi-look processing on the single look complex (SLC) data;
2. Apply the polarimetric SAR speckle filter of Lee et. al \[17\] (e.g., win-
dow size = 7);
3. Form a similarity matrix for each of the channel powers (i.e., $|HH|^2$, $|HV|^2$, $|VV|^2$) using (3.10). To represent spatial proximity, limit the
number of non-zero affinities for each pixel by choosing a neighborhood
as discussed in Section 2.2.2;
4. Form the combined similarity matrix, $W_{tot}$, using (4.5);
5. Perform the multiclass spectral clustering (MSC) algorithm \[60\] on
$W_{tot}$ to obtain an over-segmented result. $nc$ is the number of par-
titions for over-segmentation. This variable will be evaluated in the
results chapter.;
6. Use the resulting partitions from Step 5 as input to a modified version
of the technique presented in Section 4.1.1 (i.e., patch-based simi-
arity):
   (a) Form a new similarity matrix for each data channel (i.e., $|HH|^2$, $|HV|^2$ and $|VV|^2$) independently using:
4.2 The Proposed Method: Contour and Spatial Proximity Followed by Polarimetric Cue

\[ W_{ij}^b = \exp \left( \frac{-\chi^2(h_i, h_j)}{2\sigma_i\sigma_j} \right) \]  

(4.4)

where \( \chi^2(h_i, h_j) \) represents the dissimilarity between the histograms \( h_i \) and \( h_j \) that are associated with the partitions \( p_i \) and \( p_j \), respectively. Only the adjacent partitions are considered for grouping and have non-zero affinities, which represents the spatial proximity cue.

(b) Use the MSC algorithm to obtain a new partitioning for \( k \) partitions, where \( k < nc \).

7. Re-label the results from Step 5 according to the grouping obtained in Step 6b.

4.2 The Proposed Method: Contour and Spatial Proximity Followed by Polarimetric Cue

This section describes the technique proposed for segmentation and classification of POLSAR data using spectral graph partitioning [65]. In this two-step approach, segmentation is performed using contour and spatial proximity cues, followed by classification of segments based on the similarity of coherency matrices. The novelty of the proposed scheme includes improvements in the segmentation step in comparison with S2 that was presented in Section 4.1.2, and classification based on pairwise similarity of the mean coherency matrices of the segments. This technique incorporates some of the visual cues from perceptual organization (i.e., contour and proximity) into polarimetric analysis, where the complete polarimetric information, in the form of a coherency matrix, is utilized through pairwise grouping.

Figure 4.5 shows the block diagram for the proposed method. The polarimetric SAR data first goes through the pre-processing steps (i.e., multi-look processing and speckle filtering). Then, three channel powers and the magnitude of the co-polarized correlation coefficient are used as input to the contour-based similarity cue. This is followed by the spatial proximity cue.
4.2. The Proposed Method: Contour and Spatial Proximity Followed by Polarimetric Cue

Multiclass spectral clustering (MSC) algorithm of Yu [60] is used to obtain the segmentation results. In the second step, the segmentation results from the previous step are used together with polarimetric coherency matrix (T) as input to the polarimetric cue, where pairwise similarity between the mean coherency matrices for the segments are measured using the symmetric revised Wishart distance. Finally, MSC algorithm is used to obtain the final classification result.

In this method, the following procedure is used for the segmentation step:

1. Perform multi-look processing on the single look complex (SLC) data set. (e.g., for Convair-580 data set, 10 azimuth looks were used);
2. Perform speckle filtering using Lee’s POLSAR filter [17];
3. Form a similarity matrix for each of the input channels (i.e., $|HH|^2$, $|HV|^2$, $|VV|^2$) and the magnitude of the co-polarized correlation coefficient (i.e., $\rho_{HHVV}$) using (3.10).
   Limit the number of affinity calculations per pixel by choosing an appropriate neighborhood size, which represents the spatial proximity cue. Outside this neighborhood the pairwise affinities are “zero”;
4. Form the combined similarity matrix, $W^{tot}$, using:

Figure 4.5: Block diagram for the proposed method.
4.2. The Proposed Method: Contour and Spatial Proximity Followed by Polarimetric Cue

\[ W_{ij}^{\text{tot}} = \prod_{b=1}^{n_b} W_{ij}^{C_b} \]  

(4.5)

where the subscript \( b \) is the index for input data channels, and \( n_b \) is the number of data channels;

5. Perform the MSC algorithm on \( W^{\text{tot}} \) as described in Section 2.1.2.

To calculate the orientation energy (OE) the image is convolved with odd-symmetric and even-symmetric filters that are obtained using the difference of an offset Gaussian (DOOG) kernel (for details see [61]). The parameters selected are: filter size, number of orientations (\( \text{ori} \)), elongation (\( \lambda \)) and scale (\( \sigma \)) of the filter. The range of values for elongation and scale parameters are determined by the size of the filter mask. For the number of orientations there are two values that were commonly used in the literature: 4 and 6 orientations. The orientation parameter represents a tradeoff between resolution (i.e., accuracy in finding the correct orientation of an edge) and smoothing. Smoothing is an important aspect since we would like to have continuous boundaries rather than rough edges that are precise in orientation. The effect of filter parameter settings on the segmentation results will be discussed in Section 5.5.

The scale parameter of the kernel (\( \sigma_c \)) needs to be determined for the similarity matrix computation of each input channel. As suggested in [56], and described in Section 3.2, the value of the scaling parameter is determined based on the maximum value of the orientation energy in the region of interest and a constant representing the edge variance (\( \text{ev} \)), whose value is fixed for the entire image, i.e., \( \sigma_c = \text{ev} \cdot \max(\text{OE}) \).

For large and complex images, working with relatively small blocks allows the estimation of a kernel bandwidth that is adaptive to the image content through the maximum value of OE for that block. The value used for the edge variance parameter was determined based both on recommendations made by Yu [56] for object segmentation, and our preliminary tests on Polarimetric SAR data [63]. Optimization of this parameter was not attempted in this work.
4.2. The Proposed Method: Contour and Spatial Proximity Followed by Polarimetric Cue

On the other hand, a single parameter is required to compute the spatial proximity cue: the size of the neighborhood \((d)\). Outside this neighborhood pairwise affinity calculations will not be performed for other cues (i.e., weights will be set to zero). We call this case *locally dense*, since all the weights in that neighborhood will be calculated and potentially have non-zero values. Large values of the parameter \(d\) are likely to cause less fragmentation of large regions. However this will require increased computational resources (i.e., memory and runtime). To speed-up the procedure and to make the processing feasible for large neighborhoods, one can apply random sampling, through the sampling rate \((sr)\) parameter. In this scenario, only a subset of the weights in the neighborhood will be calculated. As a result, the similarity matrix becomes more sparse then the *locally dense* case, and the computational cost is kept at a manageable level, while longer connections are allowed in the graph.

The number of classes \((k)\) is assumed to be known. Thus, it should be provided by the user. Analysis of the segmentation and classification results obtained for different numbers of partitions \((nc)\) is provided in Section 5.5.

In the classification step the pairwise similarities of the mean coherency matrices are calculated based on the Symmetrical Revised Wishart distance, \(d_{SRW}\) and local scaling is employed to automate the selection of scaling parameter.

The following procedure is used for the classification step:

1. Form the similarity matrix based on the mean coherency matrices \((T)\) for each region obtained as a result of the segmentation step. If the two regions to be compared have a different number of pixels, sub-sample the larger region to the size of the smaller one. This ensures that, the mean coherency matrices are calculated using the same number of samples (i.e., pixels) for both regions.

\[
W_{ij}^{SRW} = \exp \left( \frac{-d_{SRW}^2(T_i, T_j)}{2 \sigma_i \sigma_j} \right)
\]  

(4.6)
Local scaling parameters, $\sigma_i$ and $\sigma_j$ are estimated using:

$$\sigma_i = \text{median}_{s_n \in N(s_i)} d_{SRW}(T_i, T_n)$$  (4.7)

where $T_i$ and $T_n$ represent the coherency matrices for instances $i$ and $n$, respectively. $s_i$ represents a region resulting from the segmentation step, and $N(s_i)$ is the set of $N_{LS}$ nearest neighbors of $s_i$ based on the pairwise Symmetric Revised Wishart distances ($d_{SRW}$) which is calculated for all the region pairs obtained as a result of the segmentation step.

2. Perform the MSC algorithm [60] on $W^{SRW}$ as described in Section 2.1.2.
Chapter 5

Experimental Results, Analysis and Discussion

This chapter includes experimental results, analysis and discussion of several schemes used to assess the utility of different cues as well as the final form of the proposed method in this thesis. The data sets used for obtaining these results are introduced in Section 5.1. Sections 5.2–5.5 present the results of research performed over several years and summarized in schemes S1, S2, S3, and the final form of the proposed method. Results for the test schemes S4 and S5 are presented in Appendix B. The quantitative analysis was presented based on overall classification accuracy in Sections 5.2 and 5.3.1. In addition to the overall accuracy, Kappa coefficient is also used in the analysis of the results for the proposed method given in Section 5.5. The performance of the discussed schemes and the proposed methods were compared with the iterative Wishart classifier. Random initialization steps are used in both the last step of spectral clustering and the Nyström technique that is used to reduce the computational cost in S1. The Wishart classifier also uses random initialization of class labels. Therefore, Wishart and S1 results are analyzed through multiple runs. In this chapter, 10 runs were performed and the average performance was reported for both Wishart and S1. To complement the numerical results, quantitative analysis was performed through visual inspection of segmentation and classification results. Visual inspection is a viable technique since our aim was not only obtaining higher classification accuracies but also perceptually plausible results. This approach is used for S3 as well as the test schemes S4 and S5 given in Section 5.4 and Appendix B.
5.1 Data Sets

In this section, the data sets used in this work are briefly introduced. These data are acquired from two airborne radar systems: Convair-580 of Environment Canada and AIRSAR of NASA/JPL.

The C/X-SAR system, carried by the Convair-580 aircraft, is an airborne SAR developed by Canada Centre for Remote Sensing (CCRS). The C/X-SAR system offers in two different frequency bands, i.e., C- and X-bands. In the C-band, the system is capable of operating in the polarimetric mode, thus enabling data acquisition for all four polarimetric channels. Fully polarimetric C-band Convair-580 data used in this work was acquired over Westham Island in British Columbia, Canada.

AIRSAR is an airborne SAR that can collect fully polarimetric data at three radar wavelengths: C-, L-, and P-bands. AIRSAR data set used in this work is over an agricultural test site known as Flevoland in the Netherlands.

5.1.1 Westham Island, BC, Canada

This data set acquired by Convair-580 of Environment Canada on 30 September 2004, consists of C-band (5300 MHz) data in four different polarizations, i.e., HH, HV, VH and VV. The original raw data was processed and calibrated by CCRS and stored as single look complex (SLC) data for each channel separately. The standard format adopted by CCRS for storing SLC image data is known as PolGASP format, where each pixel corresponds to a complex number and is written in big-endian format. The PolGASP files are essentially binary files with the extension, .img and each file represents the data acquired at a specific polarization. These image data, stored in the PolGASP files, are in slant range, and have a resolution of approximately 4 m in range and 0.4 m in azimuth. The data have been multi-look processed by a $10 \times 1$ averaging window (i.e., 10-looks in the azimuth direction) in order to obtain square pixels.

This particular data acquisition was part of the West Coast Campaign in 2004 and it was planned for The University of British Columbia Radar Remote Sensing Group (UBC-RRSG) in coordination with MacDonald, Det-
5.1. Data Sets

twiler and Associates Ltd. (MDA) and Canada Center for Remote Sensing (CCRS). The source of the Westham Island data set is the Canadian Space Agency (CSA) with the cooperation of Environment Canada, Natural Resources Canada (CCRS), Defense Research and Development Canada - Ottawa (DRDC-O), and MacDonald Dettwiler and Associates Ltd. Geospatial Services (MDA GS, formerly known as Radarsat Int., RSI). The availability of data should be discussed with these authorities.

Four members of the UBC-RRSG were involved in the field work for collecting the ground truth information during the two field trips on 28th and 30th of September 2004. GPS measurements taken at the field boundaries using a hand-held device were used as ground control points (GCPs). Photographs of the fields taken at these GCPs, visual inspection of the content for some fields, and interviews with local farmers were used to obtain the ground truth information. However, detailed ground truth information was collected at limited number of locations, and the rest is considered unknown. Therefore, we have chosen to work with three different test regions.

In this work, a partial scene over the Westham Island, located in the South of Vancouver (BC, Canada), is used. This scene contains small-sized fields of a variety of agricultural crops including corn, potatoes, raspberry, strawberry, hay, bare soil, barley, wheat, pumpkin, turnip, red cabbage, broccoli, and grass.

Figure 5.1 shows the color composite of this data set. The geographic North in this image points to the left side of the scene, which covers 2.17 km in the N-S direction and 3.03 km in the E-W direction. Scene center co-ordinates are 49° 05’ 16” N and 123° 08’ 57” W. Near range is in the East side (top) and far range is in the West side (bottom) of the scene. The size of the image data after multi-look processing is 600 × 500 pixels (600 pixels in E-W direction, and 500 pixels in N-S direction).

The three different test regions used in this work are shown in Figure 5.2. Regions # 1 and # 2 have a relatively simple structure which allow us to analyze the results easily and draw conclusions. These two regions represent the first scenario where segmentation and classification maps coincide. On the other hand, Region # 3 is more challenging, as it includes many classes.
in a compact region and multiple segments have a single class label. More specifically the challenge is due to the following properties: (1) adjacent segments with same class type; (2) non-adjacent segments with same class type; (3) non-convex regions that results in fragmented segments.

Figure 5.2 also shows the reference segmentation and classification maps obtained manually via visual inspection of the image data and fieldwork where possible. Figures 5.2(b) and 5.2(e) show the enlargements for Test Regions #1 and #2. Corresponding manual segmentations are given in Figures 5.2(c) and 5.2(f). For all three regions, the Wishart classifier results are shown in Figure 5.2(d), (g), and (k), with an overall pixel accuracy of 81.6%, 74.1%, and 63.1%, respectively.

Figure 5.3 provides the ground truth information and some field photographs from the Test Region #3, also known as region of interest (ROI). The size of this test region is 132 × 112 pixels, which was used to obtain some of the results presented in this chapter.
5.1. Data Sets

Figure 5.1: Westham Island Scene acquired on September 30, 2004 - Color Composite HH-HV-VV. Near range is along the top border of the image, with North pointing to the left side of the scene. The scene covers 2.17 km in the N-S direction and 3.03 km in the E-W direction. After multi-look processing the image contains 500 pixels in N-S direction, and 600 pixels in E-W direction (pixels represent approximately 4 m × 4 m in slant range). Scene center co-ordinates: 49° 05’ 16” N; 123° 08’ 57” W.
5.1. Data Sets

Figure 5.2: Westham Island Scene acquired by the Convair-580 ©CSA 2004
(a) RGB color composite [HH-HV-VV] (600 × 500 pixels); (b,e) Test Region # 1 (40 × 70 pixels) and Test Region # 2 (70 × 150 pixels); (c,f) Manual segmentations; (d,g) Wishart classifier results; (h) Test Region # 3 (132 × 112 pixels); (i) Manual Segmentation; (j) Ground truth; (k) Wishart classifier result for 13 classes
Figure 5.3: Region of Interest (ROI) in the Westham Island scene (a.k.a. Test Region # 3), ground truth and some field photographs
5.1. Data Sets

5.1.2 Flevoland, The Netherlands

This L-band (1249 MHz) data set was acquired by AIRSAR of NASA/JPL in 1989 over the agricultural test site Flevoland, the Netherlands. This test site and data set have been widely used in the polarimetric radar imaging literature over the last two decades [28], [52], [66]. This particular data set is distributed as multi-look processed and publicly available through The Polarimetric SAR Data Processing and Educational Tool (PolSARPro) [67] by ESA. Figure 5.4 shows the L-band image data as a color composite obtained from HH, HV, and VV channels.

Figure 5.4: Flevoland L-band data: color composite (750 × 1024 pixels)

The Flevoland data set has several ground truth maps that are published in the literature with different levels of detail [28], [52], [66]. However this information is self-consistent in most cases. In this study, a digitized ground truth map that combines the available information was used. This ground
true map was provided by the authors of Anfinsen et al. [52]. Figure 5.5 presents a subset of the data set along with its ground truth map and legend. This subset was used to obtain the experimental results which will be presented in Section 5.5.

![Figure 5.5: Flevoland data subset (200 × 320 pixels): RGB color composite, ground truth map and the corresponding legend.](image-url)
5.2 Results and Analysis for Scheme S1

The results presented in this section\(^{10}\) were obtained using a subset of the Westham Island Scene shown in Figure 5.2(a). Since the ground truth information is available at a limited number of fields, a small region of interest (ROI - also known as Test Region # 3) shown in Figure 5.2(h) is used for analysis.

A number of results were obtained for different values of the patch size \(d = 11, 15\) and 21) as well as the neighborhood size for local scaling \(n_{LS} = 5\) and 25). Two different scenarios for spatial proximity were also considered. The first one is with spatial proximity cue, where \(r = 45\) is used to calculate \(W_{ij}^{P}\) as explained in (4.1). The second case is without the spatial proximity cue, where proximity cue does not affect the combined similarity matrix (i.e., \(W_{ij}^{P} = 1\)). This is equivalent to having a neighborhood radius that is larger than the image size. For this scenario, the value of \(r\) is denoted as N/A in Table 5.1.

Table 5.1: Overall classification accuracy (in %) obtained by S1 for patch size, \(d \in \{11, 15, 21\}\), and local scaling neighborhood size, \(n_{LS} \in \{5, 25\}\) for two scenarios: with and without spatial proximity cue

<table>
<thead>
<tr>
<th>(r)</th>
<th>(n_{LS})</th>
<th>(d = 11)</th>
<th>(d = 15)</th>
<th>(d = 21)</th>
</tr>
</thead>
<tbody>
<tr>
<td>N/A</td>
<td>5</td>
<td>71</td>
<td>69</td>
<td>72</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>71</td>
<td>71</td>
<td>73</td>
</tr>
<tr>
<td>45</td>
<td>5</td>
<td>75</td>
<td>72</td>
<td>74</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>75</td>
<td>75</td>
<td>73</td>
</tr>
</tbody>
</table>

The Nyström extension method given in Section 2.2.3 is used to obtain the presented results. Since this technique involves a random selection of the sample points, 10 runs were performed for each of the cases mentioned above. The classification accuracies given in Table 5.1 are the average values obtained from these multiple runs, where the accuracy at each run is defined as the percentage of the total number of correctly classified pixels.

\(^{10}\)These results are published in part in the proceedings of IEEE Int. Geoscience and Remote Sensing Symposium (IGARSS), Denver, CO, August 2006.
5.2. Results and Analysis for Scheme S1

Figure 5.6: S1 results: One set of results used to obtain the average classification accuracy results given in Table 5.1. For this figure $n_{LS} = 25$.

Figure 5.6 shows a set of results obtained using the proposed scheme and the Wishart classifier. These results suggest that $d = 11$ provides the most accurate solution and the performance is improved if proximity information is included as an additional cue. These findings are also confirmed by the results given in Table 5.1. Increasing the patch size results in the blue and red regions (the three stripes in the middle of the ROI) being merged, since they are rather small compared to the size of the patch. However, in all these cases the classification performance is better than the Wishart classifier, whose accuracy was found to be 63.1%.

Figure 5.7 shows the spectral clustering result along with the normalized eigenvectors of matrix $L$ that correspond to the first $k$ eigenvalues. These eigenvectors are used in the last step of the spectral clustering algorithm as the new coordinates of the data points. This figure attempts to show the capability of obtaining separable clusters in this new space spanned by these eigenvectors.
5.3 Results and Analysis for Scheme S2

This section presents the results obtained using Scheme S2 and the subsections will discuss experiments for analyzing different aspects of this scheme.

5.3.1 Utilizing Three Intensity Channels

The segmentation results presented in this section\textsuperscript{11} were obtained using the methodology proposed in Section 3.2. This methodology (i.e., S2) uses contour-based similarity and spatial proximity cues. The results include Test Regions #1 and #2 of Westham Island scene shown in Figure 5.2(a), as well as the Test Region #3.

Wishart classifier results that were obtained after Lee’s polarimetric speckle filtering\textsuperscript{17} are used as a benchmark, and are shown in Figures 5.2(d) and 5.2(g). Based on the manual segmentations used as reference, Wishart classifier results are not as homogeneous as desired. Since each

\textsuperscript{11}These results are published in part in the proceedings of IEEE Int. Geoscience and Remote Sensing Symposium (IGARSS), Barcelona, Spain, July 2007
5.3. Results and Analysis for Scheme S2

segment represents a different class in these two test regions, classification accuracy can be calculated based on the manual segmentation. For Test Regions # 1 and # 2, Wishart classifier achieves 81.6 % and 74.1 % overall accuracies, respectively.

Figures 5.8 and 5.9 present the data, its ground truth, and obtained segmentation result for Test Regions #1 and #2, respectively. In each figure, the first row shows the three input data channels, their RGB color composite, and the manual segmentation which is used as the ground truth. The second row shows the orientation energy maps for the first four plots in the first row, and the segmentation result obtained using the proposed scheme, S2.

Figure 5.8: S2 Results: Westham Island Test Region # 1 ($\lambda^2 = 9$, ori = 4)

In Test Region #1 shown in Figure 5.8 there are six different fields. The result obtained using the proposed scheme (S2) based on MSC algorithm is very similar to the manual segmentation. However, Test Region #2 shown in Figure 5.9 has eight fields and the proposed scheme is not as accurate as it was for Test Region # 1, but in general still agrees with the reference manual segmentation, and provides more homogeneous results than Wishart classifier. The overall classification accuracy values in terms of pixel count for Test Regions #1 and #2 are 94.3% and 94.7%, respectively. These two examples show that the proposed scheme is able to provide good segmentation results, both qualitatively and quantitatively. The results presented in Figures 5.8 and 5.9 were obtained using the following parameters: filter size $d = 21$; elongation $\lambda^2 = 9$; and number of orientations $ori = 4$.

In order to assess the effects of elongation and number of orientations on the results, different values for these parameters were tested. Corresponding
Table 5.2: Overall classification accuracy (in %) obtained by S2 for elongation, $\lambda^2 \in \{5, 9\}$, and number of orientations, $ori \in \{4, 6\}$ for Test Regions # 1 and # 2 of the Westham Island data

<table>
<thead>
<tr>
<th># of orientations (ori)</th>
<th>4</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>elongation ($\lambda^2$)</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>Test Region # 1</td>
<td>94.3</td>
<td>94.3</td>
</tr>
<tr>
<td>Test Region # 2</td>
<td>95.2</td>
<td>94.7</td>
</tr>
</tbody>
</table>

Figure 5.9: S2 Results: Westham Island Test Region # 2 ($\lambda^2 = 9$, ori = 4) results are given in Table 5.2, which suggest that these parameters do not significantly affect the level of overall accuracy.

5.3.2 Utilizing Co-polarized Correlation Coefficient

Co-polarized correlation coefficient is known to be important discriminator in polarimetric data analysis, especially for target detection and identification. In this section, a comparison between using three input channels: $|HH|^2$, $|HV|^2$, and $|VV|^2$ (i.e., $n_b = 3$) and including the magnitude of the
co-pol correlation coefficient, i.e., $\rho_{HHVV}$, as an additional input ($n_b = 4$), is given for all of the three test regions. Both scenarios were discussed for two cases: (1) with speckle filtering; and (2) without speckle filtering.

Figure 5.10 shows the results obtained for Test Region # 1 using the segmentation technique explained in Section 4.1.2. These results suggest that using speckle filtering and utilizing the co-pol correlation coefficient as an additional input channel provides the best solution. Both cases, where only the three input channels are used, were found inferior to the suggested case. If speckle filter is not used, the addition of the fourth input (i.e., $\rho_{HHVV}$) results in lower performance.

Figure 5.10: S2 Results: Region # 1. (a) with filtering, $n_b = 3$; (b) with filtering, $n_b = 4$; (c) without filtering, $n_b = 3$; (d) without filtering, $n_b = 4$

Similarly, Figures 5.11 and 5.12 show the results obtained for Test Regions # 2 and # 3, respectively. These results generally agree with the findings from Test Region # 1, with the exception that in Test Region # 2, three-channel input without speckle filtering also provides good results.

Figure 5.11: S2 Results: Region # 2. (a) with filtering, $n_b = 3$; (b) with filtering, $n_b = 4$; (c) without filtering, $n_b = 3$; (d) without filtering, $n_b = 4$

Figure 5.12: S2 Results: Region # 3. (a) with filtering, $n_b = 3$; (b) with filtering, $n_b = 4$; (c) without filtering, $n_b = 3$; (d) without filtering, $n_b = 4$
5.3. Results and Analysis for Scheme S2

Figure 5.12: S2 Results: Region # 3. (a) with filtering, \( n_b = 3 \); (b) with filtering, \( n_b = 4 \); (c) without filtering, \( n_b = 3 \); (d) without filtering, \( n_b = 4 \)

5.3.3 Analysis for Different Number of Partitions

In this section, the analysis for individual input channels (i.e., \(|HH|^2\), \(|HV|^2\), \(|VV|^2\), \(\rho_{HHVV}\), and \(\phi_{HHVV}\)) are given for a varying number of partitions. In addition to that some combinations were considered, such as three channels, i.e., \(|HH|^2\), \(|HV|^2\), and \(|VV|^2\) (\(n_b = 3\)) and four channels with addition of co-pol correlation coefficient, i.e., \(\rho_{HHVV}\) (\(n_b = 4\)). The experiments were repeated for two cases: (1) with speckle filtering; and (2) without speckle filtering. The number of partitions (i.e., \(nc\) parameter) was varied between the minimum (i.e., \(nc = 2\)) and a reasonable upper limit that is larger than the correct number of partitions and is determined based on visual inspection for each test region.

The purpose of this experiment was to assess the capability of individual input channels in terms of their segmentation performance, and the number of segments they may be able to segregate for each configuration. This analysis also reveals some information regarding the stability of the technique over a range of values for the number of partitions. Thus, it attempts to answer the question: How critical is it to choose the correct value for this parameter?. It is expected that individual inputs can only differentiate between some of the segments, but not all. Therefore each input will provide good results for a value of \(nc\) value that is smaller than the actual value (i.e., reference or ground truth).
## 5.3. Results and Analysis for Scheme S2

### Test Region # 1

Figure 5.13 shows the input channels for Test Region # 1. The first and second rows show the input data channels before and after speckle filtering, respectively. From visual inspection of this figure, it is possible to conclude that individual channels are only capable of discriminating between some of the fields – not all – and different inputs play a complementary role.

![Figure 5.13: Test Region # 1 - Top: before filtering; Bottom: after filtering](image)

<table>
<thead>
<tr>
<th>HH</th>
<th>HV</th>
<th>VV</th>
<th>ρ</th>
<th>φ</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="HH" /></td>
<td><img src="image" alt="HV" /></td>
<td><img src="image" alt="VV" /></td>
<td><img src="image" alt="ρ" /></td>
<td><img src="image" alt="φ" /></td>
</tr>
<tr>
<td><img src="image" alt="w/ Lee" /></td>
<td><img src="image" alt="w/ Lee" /></td>
<td><img src="image" alt="w/ Lee" /></td>
<td><img src="image" alt="w/ Lee" /></td>
<td><img src="image" alt="w/ Lee" /></td>
</tr>
</tbody>
</table>

Figure 5.13: Test Region # 1 - Top: before filtering; Bottom: after filtering

<table>
<thead>
<tr>
<th>HH</th>
<th>HV</th>
<th>VV</th>
<th>ρ</th>
<th>φ</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="HH" /></td>
<td><img src="image" alt="HV" /></td>
<td><img src="image" alt="VV" /></td>
<td><img src="image" alt="ρ" /></td>
<td><img src="image" alt="φ" /></td>
</tr>
<tr>
<td><img src="image" alt="w/ Lee" /></td>
<td><img src="image" alt="w/ Lee" /></td>
<td><img src="image" alt="w/ Lee" /></td>
<td><img src="image" alt="w/ Lee" /></td>
<td><img src="image" alt="w/ Lee" /></td>
</tr>
</tbody>
</table>

Figure 5.14: Partitioning for individual inputs for a range of nc - Test Region # 1. First row: HH; Second row: HV; Third row: VV; Fourth row: ρ_{HHVV}

![Figure 5.14: Partitioning for individual inputs](image)

Figure 5.14 shows the results of four input channels (before speckle filtering) with nc (# of partitions) ranging from 2 to 8, while the actual number...
of partitions in our reference was 6. The first two rows in Figure 5.14 suggest that using HH or HV as single inputs we can achieve reasonable results up to \( nc = 3 \) with similar outputs. The third row shows the VV result, in which the partitioning result is not as good even though another boundary was captured. Thus, the maximum number of partitions correctly obtained from VV is 2. On the other hand, \( \rho_{HHVV} \) shown in the fourth row is able to provide good results up to \( nc = 4 \).

![Images of partitioning results for different \( nc \) values]

Figure 5.15: Partitioning for combined inputs for a range of \( nc \) - Test Region #1. First row: \( n_b = 3 \); Second row: \( n_b = 4 \)

Figure 5.15 shows the results of combined inputs for the same range of \( nc \). The first row shows results for three input channels. Good outputs are obtained for four segments (i.e., \( nc = 4 \)). Also, note that \( nc = 5 \) and \( nc = 6 \) provide consistent results with an additional region on the lower side, which can be explained by the variation in the level of detail that users may be looking for. The second row shows the case where four input channels are used. Good results are obtained up to five segments. Similarly, partitioning for seven and eight segments provide more details and are consistent with \( nc = 5 \). When the scenario with speckle filtered inputs is considered, all four rows in Figure 5.16 suggest that using \( |HH|^2 \), \( |HV|^2 \), \( |VV|^2 \), or \( \rho_{HHVV} \) as single inputs good results up to \( nc = 4 \) were achieved. These cases present different boundaries and carry complementary information, and therefore should be used together.

Figure 5.17 shows the results of combined inputs for the same range of \( nc \). The first row shows the results when three input channels are used. Good results are obtained for five segments, and can be extended up to seven with
5.3. Results and Analysis for Scheme S2

Figure 5.16: Partitioning with individual inputs (after filtering) for a range of $nc$ - Test Region # 1. First row: HH; Second row: HV; Third row: VV; Fourth row: $\rho_{HHVV}$

an addition of an extra region on one of the boundaries. Again, the second row shows the case where four input channels are used. Good results are obtained up to six segments. Similarly, seven and eight partitions provide more details consistent with six, which is the correct value according to the manual segmentation. These observations suggest that over-segmentation can be used to obtain higher level of detail. In this example, the best result was obtained using four inputs after speckle filtering.

Figure 5.17: Test Region # 1 - Partitioning with combined inputs (after filtering) for a range of $nc$. First row: $n_b = 3$; second row: $n_b = 4$
5.3. Results and Analysis for Scheme S2

Test Region # 2

The input data channels for Test Region # 2 are shown in Figure 5.18. The first and second rows show the two cases: before and after speckle filtering. From visual inspection of this figure it can be concluded that individual channels are only capable of discriminating between a subset of the fields, and different inputs seem to carry complementary information. This observation is consistent with Test Region # 1.

|HH| 2 |HV| 2 |VV| 2 |ρ |HHVV |
|---|---|---|---|---|---|---|
|HH| 2 |HV| 2 |after Lee filter |
|VV| 2 |ρ |HHVV |

Figure 5.18: Test Region # 2 - Top: before filtering; Bottom: after filtering

Figure 5.19 shows the results of four individual input channels (before speckle filtering) with \( nc \) (number of partitions) ranging from 2 to 10. Analysis of the first row in Figure 5.19 suggests that using HH one can achieve good results with three and four partitions (i.e., \( nc = 3 \) and \( nc = 4 \)). According to HV results shown in the second row, good results can be obtained for five and six segments. On the other hand, VV provides good results from two to four partitions, and \( \rho_{HHVV} \) gives an acceptable result for \( nc = 4 \).
5.3. Results and Analysis for Scheme S2

Figure 5.19: Partitioning for individual inputs for a range of $nc$ - Test Region # 2. First row: HH; Second row: HV; Third row: VV; Fourth row: $\rho_{HHVV}$

Figure 5.20: Partitioning for combined inputs for a range of $nc$ - Test Region # 2. First row: $n_b = 3$; Second row: $n_b = 4$
5.3. Results and Analysis for Scheme S2

Figure 5.21: Partitioning with individual inputs (after filtering) Test Region # 2. First row: HH; Second row: HV; Third row: VV; Fourth row: $\rho_{HHVV}$

Figure 5.22: Test Region # 2 - Partitioning with combined inputs (after filtering) for a range of nc. First row: $n_b = 3$; second row: $n_b = 4$
5.3. Results and Analysis for Scheme S2

Figure 5.20 shows the results of combined inputs for the same range of $nc$. In the first row three input channels are used, and good results are obtained for two, three, seven and eight partitions. The second row shows the case where four input channels are used. Good results are obtained up to five segments. However, with the addition of $\rho_{HHV}$, results for seven and eight segments are no longer good, which suggests that the the co-pol correlation coefficient should not be used before speckle filtering.

Results of individual input channels after speckle filtering are shown in Figure 5.21. HH, HV and VV results are quite similar to the corresponding results before filtering. On the other hand, $\rho_{HHV}$ after speckle filtering provides better results and is robust between $nc = 2$ and $nc = 4$.

Figure 5.22 shows the results of combined inputs for the same range of $nc$. In the first row three power channels are used, and good results are obtained up to $nc = 6$. The second row shows the case where four input channels are used after speckle filtering. Good results are obtained for two, three, seven, and eight segments. Also, nine and ten segments provide more details where boundaries are consistent with $nc = 8$ which is the correct number of partitions according to our reference (i.e., manual) segmentation for this example. Using four inputs after speckle filtering again provides the best result both in terms of accuracy at the correct number of partitions and robustness due to consistency between different levels of detail.

Test Region # 3

Figure 5.23 shows the input channels for Test Region # 3. The first and the second rows correspond to the input data channels before speckle filtering and after speckle filtering, respectively. Similar to Figures 5.13 and 5.18 for Test Regions #1 and #2, Figure 5.23 also suggests that individual channels are only capable of discriminating between a subset of the fields, not all, and different inputs carry complementary information.

Figures 5.24 - 5.27 show the results of HH, HV, VV and $\rho_{HHV}$ before speckle filtering with $nc$ (number of partitions) ranging from 2 to 19. Figure 5.24 suggest that using HH as single input good results can be achieved
5.3. Results and Analysis for Scheme S2

![Images showing various inputs and after Lee filter for Scheme S2]

Figure 5.23: Test Region # 3 - Top: before filtering; Bottom: after filtering

for two segments. Although increasing the number of partitions allows one to find many correct boundaries, some incorrect ones are also introduced. A similar observation can be made for HV shown in Figure 5.25. In this case correct partitioning results are obtained up to \( nc = 3 \). VV and \( \rho_{HVV} \) results are shown in Figures 5.26 and 5.27, which suggest that small values of \( nc \) give correct results. Although increased value of \( nc \) provides most of the boundaries, some fields are incorrectly divided into smaller regions. An important point is that among the four inputs, \( \rho_{HHVV} \) is the only one that enables the determine the boundary between the two fields on the lower right corner of this region using \( nc = 9 \) and \( nc = 10 \) or when \( nc \geq 14 \).

Figures 5.28 and 5.29 show the results from combined inputs for the same range of \( nc \). If three input channels (i.e., HH, HV, and VV) are used, good results are obtained for \( nc = 2 \), and \( nc = 4 \). Also from six to nine segments all the boundaries except one (i.e., the boundary between the two fields on the lower right corner) are correct. A similar result is obtained with four input channels. However in this case \( nc = 3 \) also gives correct partitioning. Once \( nc \) reaches 18 and 19 the boundary with over-segmentation of the field to the left can be seen. This can be considered an improvement over the three-channel input case.
5.3. Results and Analysis for Scheme S2

Figure 5.24: Test Region # 3: partitioning with HH.

Figure 5.25: Test Region # 3: partitioning with HV.
5.3. Results and Analysis for Scheme S2

Figure 5.26: Test Region # 3: partitioning with VV.

Figure 5.27: Test Region # 3: partitioning with $\rho_{HHVV}$. 
5.3. Results and Analysis for Scheme S2

Figure 5.28: Test Region # 3: partitioning with HH, HV, and VV.

Figure 5.29: Test Region # 3: partitioning with HH, HV, VV, and $\rho_{HHVV}$. 

81
As shown in Figure 5.30 using HH after speckle filtering as a single input we can achieve a reasonable results for $nc = 5$. Here, a similar trend, compared to the non-filtered case is present. Results from HV are shown in Figure 5.31 and in this case correct partitioning results are obtained from $nc = 2$ to $nc = 4$. VV and $\rho_{HHV}$ results are shown in Figures 5.32 and 5.33. Among the single inputs, VV after speckle filtering provides the highest number of correct partitions with $nc = 6$, while $\rho_{HHV}$ can only go up to $nc = 3$. However, with increasing number of partitions (starting from $nc = 7$) $\rho_{HHV}$ enables the detection of the boundary at the lower right corner.

Figures 5.34 and 5.35 show the results from combined inputs after speckle filtering for the same range of $nc$. If three input channels (i.e., HH, HV, and VV) are used, good results are obtained from two to nine segments. Further increase in the number of partitions again introduces an incorrect boundary at the lower right corner, but also detects many other correct ones. In the case of four input channels (after speckle filtering), results obtained up to $nc = 13$ are good. Further increase results in the same incorrect boundary (lower right corner), however many other correct boundaries are also detected. Increase in the level of detail is consistent with the increase in the number of partitions. Finally, $nc = 19$ provides almost all of the boundaries present in the reference segmentation, with two additional ones: (1) the boundary between the fields of the same content (upper center), and (2) the incorrect boundary formed due to the non-convex geometry of that field (lower right).

From segmentation point of view this result is very good and the two problems mentioned above for classification can be fixed at a later stage with approaches such as S3 given in Section 4.1.3 or the proposed method given in Section 4.2. Among the experiments with single and combined inputs used with and without speckle filtering, using four inputs after speckle filtering has consistently provided the best results in all three examples.
5.3. Results and Analysis for Scheme S2

Figure 5.30: Test Region #3: using HH after speckle filtering

Figure 5.31: Test Region #3: using HV after speckle filtering
5.3. Results and Analysis for Scheme S2

Figure 5.32: Test Region #3: using VV after speckle filtering

Figure 5.33: Test Region #3: using $\rho_{HHVV}$ after speckle filtering
5.3. Results and Analysis for Scheme S2

Figure 5.34: Test Region # 3: using HH, HV, and VV after speckle filtering

Figure 5.35: Test Region # 3: using HH, HV, VV, and $\rho_{HHV}$ after filtering
5.4 Results and Analysis for Scheme S3

The methodology presented in Section 4.1.3 called S3 was a first attempt to utilize over-segmentation in the classification process, and to avoid misclassification due to the fixed number of segments required by the algorithm. This method employs S2 (i.e., contour-based similarity and spatial proximity cues) to obtain an over-segmented result, followed by S1 (i.e., patch-based similarity based on $\chi^2$ distance) to merge the segments 12.

Figure 5.36: S3 Results for Test Region # 1 ($n_b = 3; nc=8; k=6$)

The approach explained in Section 4.1.2 called S2, was employed on the three unfiltered input channels. Figures 5.36 and 5.37 present the results obtained for Test Regions # 1 and #2 of the Westham Island scene. As discussed before in Section 5.3.1, S2 already achieved very good results on these two examples. Therefore the two-step approach: over-segmentation followed by merging did not improve the results, in fact reduced the accuracy in boundary localization, that is apparent by visual inspection.

Figure 5.37: S3 Results for Test Region # 2 ($n_b = 3; nc=10; k=8$)

12Results shown in this section were presented in part at Advanced Synthetic Aperture Radar (ASAR) Workshop, Vancouver, BC, Sept. 2007 [64]
5.4. Results and Analysis for Scheme S3

Figure 5.38: S3 Results for Test Region # 3 ($n_b = 3; n_c=18; k=13$) (a) RGB composite; (b) Wishart classifier result; (c) Ground truth and the three problems that S2 can not address; (d) S2 result; (e) Over-segmentation result after the first step of S3; (f) S3 result.

Figure 5.38 shows the results for Test Region # 3, that is more challenging due to the following properties: (1) adjacent segments with same class type; (2) non-adjacent segments with same class type; (3) non-convex regions that results in fragmented segments. These three cases are marked on the ground truth using an ellipse for (1), diamonds for (2) and circle for (3). The ground truth is shown in Figure 5.38(c).

Using the contour-based methodology (S2) as a one-step solution, shown in Figure 5.38(d), all of these problems can be observed. While some of these could potentially be solved using an additional classification step (e.g., the pumpkin fields or the grass/hay field boundary at the top marked with an ellipse), others can not be recovered because the resulting segmentation is simply not detailed enough in those areas. Due to the fixed number of segments in a region of interest, fragmented segments or adjacent fields of same content that are labeled different, forces other segments with different content to be merged. Therefore the detail is lost in the first step and recovery is not possible unless the number of initial segments ($n_c$) is chosen to be larger than the final number of segments or classes ($k$). This strategy proves to be useful in the case of Test Region # 3. As shown in Figures 5.38(e) and 5.38(f) over-segmentation allows more field boundaries to be determined in the expense of a few over-segmented fields. Fields that are
split either due to an existing boundary or the region being non-convex are both merged by S1. The grass/hay field at the top (marked by an ellipse) and the barley field on the lower right (marked by a circle) are both merged correctly by the second step. The qualitative results on Test Region # 3 show an apparent improvement over S2 and generally a good correlation with the ground truth, although no quantitative assessment was performed.

Soon after the implementation of S3, the idea of using the similarity of coherency matrices in the merging process, started to take shape as a potentially more effective approach. The similarity of coherency matrices were analyzed through S4 and S5, and S3 later evolved to the proposed method, which was described in Section 4.2 and is the most recent and promising scheme resulting from this work, as will be discussed in further detail in Section 5.5.

5.5 Results and Analysis for the Proposed Method

This section presents the experimental results of the proposed technique in Section 4.2 using two airborne data sets over agricultural fields. Westham Island data from Convair-580, a C-band airborne sensor, is used to demonstrate the technique and analyze its behavior for segmentation and classification for two scenarios: (1) two small regions where each segment represents a distinct class; and (2) a relatively large and more complex region where some classes are represented by multiple segments. The procedure was further tested with a second data set – the Flevoland data set acquired by AIRSAR in L-band. The performance was evaluated both in terms of overall classification accuracy (OA) and Kappa (K) coefficient and compared with the Wishart classifier performance. The overall accuracy is calculated as the ratio of the correctly classified pixels to the total number of pixels, given in percentage. It can also be obtained as the sum of the diagonal elements in the confusion matrix.
5.5. Results and Analysis for the Proposed Method

5.5.1 Convair-580: Westham Island Dataset

The results presented in this section are obtained using a subset of the Westham Island scene and the three test regions introduced in Section 5.1. These data are shown in Figure 5.2.

Segmentation results for the first two test regions are obtained using the following parameter sets: Number of orientations, \( ori \in \{4, 6\} \); \( \sigma \in \{1, 2\} \); \( \lambda^2 \in \{3, 5, 9\} \); and a filter size of \( 21 \times 21 \) pixels. The neighborhood that represents spatial proximity information is a \( (2d + 1) \times (2d + 1) \) window where \( d \in \{7, 11, 15\} \).

Tables 5.3 and 5.4 present the overall accuracy (OA) and Kappa coefficient (K) for Region # 1 and Region # 2, respectively. Results given in both tables suggest that the accuracy is not very sensitive to filter parameters. A similar conclusion can be drawn from Figures 5.39 and 5.40, which show the results for six different combinations of \( \sigma \) and \( \lambda^2 \), where \( d = 15 \). Overall accuracy and Kappa coefficient values corresponding to the results shown in Figures 5.39 and 5.40 are printed in bold typeface in Tables 5.3 and 5.4. A set of these results can be directly compared with the ones presented in [63] using the same parameter settings (i.e., \( ori = 6, \sigma = 1, \lambda^2 = 9 \)). While reported overall accuracy in [63] for Regions # 1 and # 2 were 89.6% and 94.4%, respectively, with the proposed technique, 90.4% and 92.9% were achieved. The variation in performance between these two very similar methods can be attributed to two differences. These are: pre-processing with the Lee speckle filter, and utilizing an additional input channel (i.e., magnitude of the co-pol correlation coefficient). While the new method provides a slightly better result in Region # 1 and slightly worse in Region # 2, these modifications provide a more general approach and are found to improve the overall performance on larger data sets, such as the Westham Island scene shown in Figure 5.45 and the Flevoland scene in Figure 5.46.
5.5. Results and Analysis for the Proposed Method

Table 5.3: Overall accuracy and Kappa coefficient obtained by the proposed method for Test Region # 1. $\lambda^2 \in \{3, 5, 9\}; \text{ori} \in \{4, 6\}; \sigma \in \{1, 2\}$

<table>
<thead>
<tr>
<th>ori</th>
<th>$\sigma$</th>
<th>$\lambda^2$</th>
<th>7</th>
<th>11</th>
<th>15</th>
<th>7</th>
<th>11</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>1</td>
<td>3</td>
<td>90.4</td>
<td>90.4</td>
<td>90.4</td>
<td>0.87</td>
<td>0.87</td>
<td>0.87</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>5</td>
<td>89.5</td>
<td>89.6</td>
<td>89.5</td>
<td>0.86</td>
<td>0.86</td>
<td>0.86</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>9</td>
<td>89.2</td>
<td>89.1</td>
<td>89.1</td>
<td>0.85</td>
<td>0.85</td>
<td>0.85</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>3</td>
<td>89.1</td>
<td>89.2</td>
<td>89.6</td>
<td>0.85</td>
<td>0.85</td>
<td>0.86</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>5</td>
<td>88.9</td>
<td>89.2</td>
<td>89.1</td>
<td>0.85</td>
<td>0.85</td>
<td>0.85</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>9</td>
<td>88.7</td>
<td>88.5</td>
<td>88.4</td>
<td>0.85</td>
<td>0.85</td>
<td>0.84</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>3</td>
<td>90.0</td>
<td>90.1</td>
<td>90.1</td>
<td>0.87</td>
<td>0.87</td>
<td>0.87</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>5</td>
<td>90.6</td>
<td>90.5</td>
<td>90.1</td>
<td>0.87</td>
<td>0.87</td>
<td>0.87</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>9</td>
<td>90.6</td>
<td>90.6</td>
<td>90.4</td>
<td>0.87</td>
<td>0.87</td>
<td>0.87</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>3</td>
<td>88.9</td>
<td>88.9</td>
<td>88.9</td>
<td>0.85</td>
<td>0.85</td>
<td>0.85</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>5</td>
<td>89.4</td>
<td>89.1</td>
<td>89.2</td>
<td>0.86</td>
<td>0.85</td>
<td>0.85</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>9</td>
<td>86.3</td>
<td>86.3</td>
<td>86.4</td>
<td>0.82</td>
<td>0.82</td>
<td>0.82</td>
</tr>
</tbody>
</table>

Table 5.4: Overall accuracy and Kappa coefficient obtained by the proposed method for Test Region # 2. $\lambda^2 \in \{3, 5, 9\}; \text{ori} \in \{4, 6\}; \sigma \in \{1, 2\}$

<table>
<thead>
<tr>
<th>ori</th>
<th>$\sigma$</th>
<th>$\lambda^2$</th>
<th>7</th>
<th>11</th>
<th>15</th>
<th>7</th>
<th>11</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>1</td>
<td>3</td>
<td>94.4</td>
<td>93.7</td>
<td>93.6</td>
<td>0.93</td>
<td>0.92</td>
<td>0.92</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>5</td>
<td>94.2</td>
<td>93.9</td>
<td>93.6</td>
<td>0.93</td>
<td>0.93</td>
<td>0.92</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>9</td>
<td>94.2</td>
<td>94.0</td>
<td>94.0</td>
<td>0.93</td>
<td>0.93</td>
<td>0.93</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>3</td>
<td>94.3</td>
<td>94.4</td>
<td>94.3</td>
<td>0.93</td>
<td>0.93</td>
<td>0.93</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>5</td>
<td>94.5</td>
<td>94.4</td>
<td>94.3</td>
<td>0.93</td>
<td>0.93</td>
<td>0.93</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>9</td>
<td>94.2</td>
<td>94.0</td>
<td>93.7</td>
<td>0.93</td>
<td>0.93</td>
<td>0.93</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>3</td>
<td>94.3</td>
<td>94.0</td>
<td>93.5</td>
<td>0.93</td>
<td>0.93</td>
<td>0.92</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>5</td>
<td>94.1</td>
<td>93.4</td>
<td>93.3</td>
<td>0.93</td>
<td>0.92</td>
<td>0.92</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>9</td>
<td>93.4</td>
<td>93.3</td>
<td>92.9</td>
<td>0.92</td>
<td>0.92</td>
<td>0.91</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>3</td>
<td>94.3</td>
<td>94.0</td>
<td>93.9</td>
<td>0.93</td>
<td>0.93</td>
<td>0.93</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>5</td>
<td>93.8</td>
<td>93.4</td>
<td>93.1</td>
<td>0.93</td>
<td>0.92</td>
<td>0.92</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>9</td>
<td>93.2</td>
<td>93.0</td>
<td>93.0</td>
<td>0.92</td>
<td>0.92</td>
<td>0.92</td>
</tr>
</tbody>
</table>
5.5. Results and Analysis for the Proposed Method

Figure 5.39: Segmentation and classification results for Test Region # 1 ($nc = 6, d = 15, ori = 6$) for different values of $\sigma$ and $\lambda^2$ parameters. Left to right: $[\sigma, \lambda^2]$ are $[1, 3]; [1, 5]; [1, 9]; [2, 3]; [2, 5]; [2, 9]$

Figure 5.40: Segmentation and classification results for Test Region # 2 ($nc = 8, d = 15, ori = 6$) for different values of $\sigma$ and $\lambda^2$ parameters. Left to right: $[\sigma, \lambda^2]$ are $[1, 3]; [1, 5]; [1, 9]; [2, 3]; [2, 5]; [2, 9]$

Segmentation results for the third test region are obtained using the same parameter set except for the neighborhood size, where $d \in \{7, 11, 15, 45\}$. Based on the results obtained for Test Regions # 1 and # 2, different values of $\sigma$ and $\lambda^2$ did not have significant impact on the segmentation results. Therefore in the remaining part of this section, only one case ($ori = 6, \sigma = 2, and \lambda^2 = 5$) is discussed for Region # 3.
5.5. Results and Analysis for the Proposed Method

Figure 5.41: Segmentation of Test Region # 3 for different number of partitions (ori = 6, σ = 2, \( \lambda^2 = 5 \), ev = 0.2, \( d = 15 \), sr = 1.0)

Figure 5.41 shows segmentation results of Region # 3 for different number of segments (nc), where ev = 0.2 and \( d = 15 \). In the spatial proximity window, all the weights were calculated (i.e., sampling rate, sr = 1.0). Although the segmentation technique is not hierarchical, and each time the global optimization problem is solved under a new constraint (i.e., minimizing the ncut for a given nc), increasing the number segments provide results that are consistent with the previous ones. Since an automated procedure is not provided to correctly estimate the number of segments in an arbitrary region of interest, having this consistency is important, and allows the user to choose the level of detail they require.

In Figure 5.42, three sets of results are presented in each row, where each row has four results that are obtained using different number of partitions (shown in the columns). The first row presents the results for locally dense case, where no sampling was used (sr = 1), and the neighborhood size for spatial proximity cue was 15 (\( d = 15 \)). The results shown in the second row were obtained using a sampling rate of 10\% (i.e., sr = 0.1), for the same neighborhood size (\( d = 15 \)). Visual comparison of these two rows suggests that differences in the resulting segmentation are minor. However, a large amount of speed-up is obtained in our implementation using MATLAB. This speed-up is due to both less computation and memory usage. More
5.5. Results and Analysis for the Proposed Method

importantly, this sampling scheme allows us to increase the neighborhood size, given limited resources. Thus, less fragmentation will occur. The third row of Figure 5.42 shows the results for a larger neighborhood size \((d = 45)\) with a sampling rate of 10\% \((sr = 0.1)\). The significant difference between second and third row is that relatively large segments on the lower right corner and upper center (i.e., barley and hay/grass fields) are not fragmented for a larger neighborhood size in the spatial proximity cue.

<table>
<thead>
<tr>
<th>Parameter Values</th>
<th>Segmentation Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>(d = 15)</td>
<td>(sr = 1)</td>
</tr>
<tr>
<td>(d = 15)</td>
<td>(sr = 0.1)</td>
</tr>
<tr>
<td>(d = 45)</td>
<td>(sr = 0.1)</td>
</tr>
</tbody>
</table>

Figure 5.42: Proposed Method Results: Segmentation of Test Region \# 3 where \(ev = 0.2, ori = 6, \sigma = 2, \lambda^2 = 5\). Columns: number of partitions, \(nc \in \{13, 16, 19, 22\}\). Rows: \((d; sr) \in \{(15; 1.0), (15; 0.1), (45; 0.1)\}\).

Figure 5.43 shows the corresponding classification results for 13 classes \((k = 13)\) obtained using different number of partitions, \(nc \in \{16, 19, 22\}\). In this figure, for all the cases, the two pumpkin fields are correctly recognized.
5.5. Results and Analysis for the Proposed Method

as a single class, shown in orange. In the first two rows where \( d = 15 \), the over-segmented barley field was successfully merged into a single class, shown in dark green. Two fields of potatoes were also labeled correctly in most cases, shown in light brown. However, in many cases two pieces of the road are not labeled as a single class.

Figure 5.43: Classification of Test Region # 3 using Symmetric Revised Wishart distance where \( ori = 6, \sigma = 2, \lambda^2 = 5, \ N_{LS} = 5. \) Columns: number of partitions, \( nc \in \{16, 19, 22\}. \) Rows: \( [d; sr] \in \{(15; 1.0), (15; 0.1), (45; 0.1)\}. \)

The graph shown in Figure 5.44 presents the overall accuracy vs. number of segments (i.e., partitions) given for the three cases that correspond to the rows of Figure 5.43. All three cases provide higher overall accuracies than the
5.5. Results and Analysis for the Proposed Method

Wishart classifier result (i.e., 63.1%). However, the best overall performance is achieved with \( d = 45 \) and \( sr = 0.1 \), where \( OA \) values range between 69\% – 81\% for \( nc \in \{14, \ldots, 22\} \). A direct comparison of results in terms of overall accuracy can be made with the results presented in [51], where \( OA \) values between 69\% – 75\% were reported. As a result, we can conclude that the proposed technique outperforms both the Wishart classifier and the technique proposed by the authors in [51]. An indirect comparison with [52] is also possible since they have reported no significant improvement in terms of classification accuracy with respect to the Wishart classifier.

![Figure 5.44](image)

Figure 5.44: Overall Accuracy vs. Number of partitions for Region # 3: \( nc \in \{14, 15, \ldots, 22\} \), \( ev = 0.2 \), \( ori = 6 \), \( \sigma = 2 \), \( \lambda^2 = 5 \), \( N_{LS} = 5 \), \( k = 13 \)

Figure 5.45 shows the Westham Island scene and the segmentation results obtained using the block processing scheme. Each block is processed independently, which allows parallel processing of blocks, reducing the processing time for the scene. The number of partitions per block were set to a fixed value (\( nsb = 15 \)), and 25 blocks of 120 \( \times \) 100 pixels were used to cover the entire scene. Since ground truth information over the whole scene is not available, classification results were not evaluated. Estimation of the number of partitions in each block is not attempted in this work.
5.5. Results and Analysis for the Proposed Method

Figure 5.45: (a) Westham Island data 600 × 500 pixels (b) Segmentation results obtained using block processing 5 × 5 blocks of 120 × 100 pixels ($d = 15$, $sr = 0.3$, $ev = 0.2$, $ori = 6$, $\sigma = 2$, $\lambda^2 = 5$)
5.5. Results and Analysis for the Proposed Method

5.5.2 AIRSAR: Flevoland Dataset

The second data set used to demonstrate the proposed technique and its comparison to Wishart classifier is the widely used agricultural data set over Flevoland, The Netherlands. This L-band data set was acquired by AIRSAR of NASA/JPL in 1989, and is distributed by ESA as multi-look processed. The Flevoland data set has several ground truth maps that are published in the literature \[66\], \[52\] with different levels of detail. However this information is self-consistent in most cases. In this study, a digitized ground truth map that combines the available information was used. It was provided by the authors of Anfinsen et al. \[52\].

Results presented in this section are obtained on a subset (200 × 320) of this data set, shown in Figure 5.46. For this region the ground truth map consists of 9 different types of fields: Stem Beans, Potatoes, Lucerne, Winter wheat I, Winter wheat II, Bare Soil, Sugar Beat, Rapeseed (or Flax), and Grass. The rest of the pixels are unclassified according to the ground truth map. All the pixels were used to obtain the segmentation and classification results with both the proposed technique and the Wishart classifier. However, the class labels for the pixels that are unclassified in the ground truth map were not included in the accuracy calculations.

Figure 5.46 shows the data subset, ground truth map, Wishart classifier result, as well as segmentation and classification results obtained using the proposed technique. These results were obtained with parameters used for the Westham Island data set that was presented in the previous section (i.e., \(d = 15\), \(sr = 1.0\), \(ev = 0.2\), \(ori = 6\), \(\sigma = 2\), \(\lambda^2 = 5\)).

The block processing scheme is employed with blocks of size 50 × 80 pixels and a total of 16 blocks were used to cover the 200 × 320 region, where the number of partitions per block was 10 (\(nsb = 10\)). The resulting 160 segments were used in the classification step, to obtain 9 classes (\(k = 9\)). The local scaling parameters (\(\sigma_i\)) were estimated using \(N_{LS} = 20\).

Overall accuracy obtained with Wishart classifier for 9 classes was 74.1% with a Kappa coefficient, \(K\), of 0.69. The proposed technique, on the other hand, achieved 81.2% accuracy with a \(K = 0.77\).
5.5. Results and Analysis for the Proposed Method

Figure 5.46: First row: Flevoland data subset of size 200 × 320 pixels, and the ground truth map for 9 classes (k = 9). Second row: The Wishart classifier result masked for unknown labels and segmentation result using the proposed technique. (The parameters: Nblk = 16, nsb = 10, d = 15, sr = 1.0, ev = 0.2, ori = 6, σ = 2, λ² = 5). Third row: Classification result using the proposed method and masked for unknown labels (NLS = 20).
Chapter 6

Conclusion

6.1 Summary

This work developed a new technique for segmentation and classification of Polarimetric Synthetic Aperture Radar (POLSAR) data. This technique is based on a methodology that exploits the polarimetric attributes of pixels, and the visual aspect of the image data through computer vision techniques.

Different classification strategies for POLSAR data have been proposed in the last two decades based on conceptual developments and experimental observations. Most of the previous work in POLSAR classification have explored solely the polarimetric attributes of pixels. Many of them involve target decompositions based on physical models (e.g., Freeman-Durden decomposition) or mathematical models (e.g., eigenvalue decomposition, H-A-α analysis). The most widely used technique is based on maximum likelihood classification which uses the fact that the multi-look processed covariance matrix obeys the complex Wishart distribution. In this thesis, iterative Wishart classifier \[26\], an unsupervised technique, is used as a benchmark to evaluate the performance of the tested schemes and proposed method.

The incorporation of human perceptual methodologies into the developed classification scheme involved studying how humans handle the classification task, and applying computer vision techniques that aim to reach the performance level of humans. In this process, several cues were utilized, such as similarity, proximity and continuity. The pair-wise grouping technique of Spectral Graph Partitioning (SGP) was employed to combine these cues and generate partitions using the Multiclass Spectral Clustering (MSC) algorithm.

In order to reduce the computational cost of the process, different proce-
dures were employed, such as sparse representation of the similarity matrix and the Nyström method.

Two fully polarimetric data sets, both containing agricultural fields, were used to obtain the results presented in this work. The first was acquired at L-band over Flevoland (The Netherlands) by the AIRSAR system of NASA/JPL. The second data set was acquired at C-band over Westham Island (BC, Canada) by the Convair-580 system of Environment Canada.

The proposed method in this thesis was developed through a progression of experiments using SGP approach, during which different schemes were tested. These schemes and the most relevant conclusions are summarized below.

Scheme S1 was the first technique we proposed [51]. It utilizes both patch-based similarity and spatial proximity cues, through an adaptation of the spectral clustering algorithm developed by Ng et al. [54]. Different patch sizes were tested that consistently outperformed the Wishart classifier. However, this scheme had the drawback that for every pairwise similarity calculation, one has to perform the $\chi^2$-distance calculation between patch histograms. This procedure requires larger memory allocation than pairwise calculations between pixels. Therefore the computational cost of the procedure is a concern. Using only the intensity information from polarimetric data, and not utilizing a powerful visual cue, i.e., the contour cue, can also be noted as a drawback.

Scheme S2 utilizes the contour information by defining pairwise similarities using the orientation energy. Multiclass spectral clustering is used to obtain a discrete solution that is close to the global optimum of our objective function (i.e., normalized cut). This method provided perceptually plausible segmentations. The results closely agree with the manual segmentations and are more homogeneous than those obtained using the Wishart classifier. If used as a single-step solution, however this method has a drawback. Since the similarity between pixels is defined through the orientation energy (OE), i.e., an enhanced version of an edge strength map, the following issues may arise: (1) non-convex regions may result in fragmented segments; (2) both adjacent and non-adjacent fields whose content are same will be labeled as
6.1. Summary

different. In the unsupervised case where the number of classes or segments are unknown, these issues may result in loss of detail and accuracy.

Scheme S3 is an attempt to address the potential problems with S2 through a two-step approach. It employs S2 for over-segmentation to ensure enough detail. S1 is then applied on the resulting segments, which allows merging fragmented segments and combining adjacent fields that have the same content. S3 provides an apparent visual improvement over both the Wishart classifier and S2 for complicated scenarios such as Test Region # 3. Computational cost is less of a concern here then using S1 as a single-step solution, because in S3, grouping based on patch histograms is used after segmentation and the number of segments are much smaller than the number of pixels. However, the inability to utilize the complete polarimetric information is still a drawback.

Our next attempt involved utilizing the full extent of the polarimetric data, through the coherency matrix representation. S4 uses the polarimetric cue (i.e, similarity of coherency matrices) along with the spatial proximity cue. Two distance measures – Bartlett distance and Symmetric Revised Wishart distance – were used and were results discussed for a wide range of values for the scaling parameter. Visual assessment of the results showed poor agreement with the ground truth except for a narrow range of scaling parameter values, which will potentially vary between test regions and data sets. Therefore, the approach used for adaptive selection of the scaling parameter, which is based on the maximum value of pairwise distances computed within the spatial proximity neighborhood, is not considered to be robust. Scheme S5 combines three cues: contour, spatial proximity, and polarimetric cue. This can also be seen as an extension of S2 with the addition of the polarimetric cue, or of S4 with the addition of the contour cue. The scaling parameter values that performed well in the previous analysis (i.e., S4) were used to generate and discuss results for four different scenarios that involve speckle filtering, and utilizing an additional input channel for the contour cue (i.e., magnitude of the co-pol correlation coefficient). An appropriate scaling parameter value for the polarimetric cue allowed improvements over S2.
6.1. Summary

Finally, the proposed method in this thesis is a two-step approach to the classification problem. First groups of pixels with similar properties are obtained, then these segments are grouped into classes. Both the contour cue based on the edge information from brightness data, and the polarimetric cue based on the coherency matrix, are utilized through the multiclass spectral clustering algorithm, which offers a near-globally optimum solution. The proposed method is a two-step technique—similar to S3—that employs S2 for segmentation, followed by grouping based on the similarity of mean coherency matrices. The pairwise similarity was defined using the Symmetrical Revised Wishart distance between segments, and the value of the scaling parameter was determined using local scaling. Over-segmentation ensures a higher level of detail, and merging segments in a separate step allows addressing all of the issues that arise in S2, such as fragmented segments, adjacent or non-adjacent segments with the same content. This method is conceptually similar to S3, however, it offers a way of utilizing the complete information available in the polarimetric coherency matrix.

The proposed method was tested on two data sets from airborne sensors operating at C-band (Convair-580) and L-band (AIRSAR). Both data sets were processed using the same set of parameter values and compared to the Wishart classifier in terms of overall classification accuracy and Kappa coefficient. This scheme offered quantitative and qualitative improvements over the Wishart classifier. A direct comparison with S1 based on the results we had presented previously in [51] shows that the proposed method outperforms S1. Also, an indirect comparison can be done with the methodology of Anfinsen et al. [52] using the same subset of the Flevoland data set. Since they have reported that their performance of classification accuracy is at the same level with the Wishart classifier, our proposed technique also outperforms their approach. It was also shown that larger data sets can be processed in parallel using block processing. Thus, the computational load is reduced and the limitation on memory size is circumvented.
6.2 Research Contributions

Over the course of development of the techniques summarized in the previous section, this work has been presented at international conferences and accepted for publication as a journal article. The four main contributions of this thesis are as follows:

- For the first time, Spectral Graph Partitioning has been used for POLSAR data classification based on patch-based similarity (Published in the Proceedings of IEEE International Geoscience and Remote Sensing Symposium - IGARSS 2006 [51]). This scheme (i.e., S1) outperforms the benchmark we have used (i.e., Wishart classifier) by 8-12% and produces perceptually plausible results. The drawbacks are: (i) the computation of similarity between histograms is memory-expensive; (ii) polarimetric information is partially utilized in this first scheme; and (iii) a powerful visual cue (i.e., contour cue) is not yet included.

- Contour information has been utilized for the segmentation of POLSAR data (Published in the Proceedings of IEEE International Geoscience and Remote Sensing Symposium - IGARSS 2007 [63]). This method, i.e., S2, also outperforms the benchmark in this case by utilizing a powerful visual cue through the orientation energy. The spatial proximity cue reduced the computational cost through sparse representation of the pairwise similarity matrices. Perceptually plausible segmentation results (i.e., similar to the performance of the human vision system) were obtained for relatively simple scenarios. However, if this one-step approach is used in more complex scenarios, fragmentation of non-convex regions arise as a potential problem.

- An attempt towards a two-step approach to the classification problem has been made: segmentation using contour and proximity cues followed by classification using patch-based similarity of intensity (This work was presented at Advanced Synthetic Aperture Radar Workshop - ASAR 2007 [64]). This scheme (i.e., S3) addresses a potential problem that may arise in S2: fragmentation of non-convex regions. It
is also used to identify adjacent fields of the same content that were labeled differently in the first step. The drawbacks are: (i) it does not solve the classification problem for non-adjacent segments, and (ii) polarimetric information is still not fully utilized.

- A new classification algorithm is developed for POLSAR data. This method comprises of a segmentation step based on proximity and contour cues, and a classification step based on similarity of mean coherency matrices (This work has been accepted for publication in IEEE Transactions on Geoscience and Remote Sensing, April 2009 [65]). This method is conceptually similar to S3, and offers a way of utilizing the complete information available in the polarimetric coherency matrix. It also overcomes the problem of identifying non-adjacent segments of the same content as different, that had not been addressed in S3. The results have shown quantitative and qualitative improvements over the Wishart classifier, Anfinsen’s spectral clustering approach [52], as well as our previously suggested schemes, S1, S2 and S3 [51], [63], [64].

The spectral graph partitioning methodology used in the tested schemes and the proposed method is flexible in the definition of affinity functions and will likely allow further improvements through the addition of different features or data sources.

The segmentation and classification scheme developed in this work is suitable for applications where homogeneity within each field is desirable, such as mapping crops or other types of terrain. In some applications users may be interested in the variability with-in each field, e.g., mapping the spatial variation of yield due to differences in soil moisture, terrain slope and aspect, soil type or crop disease. These applications will still benefit from the developed schemes, however using different levels of detail may be required.
6.3 Future Work

The following future research directions are worth exploring:

- Validating results for data from space-borne sensors. Data from space-borne platforms typically have lower resolution and lower signal-to-noise ratio than the airborne data used in this work. Experiments should be performed with POLSAR data acquired by systems such as RADARSAT-2, TerraSAR-X, COSMO-Skymed and ALOS/PALSAR.

- Providing a wider range of experimental results and testing the effectiveness of the proposed schemes for scenes with other terrain and land cover types, e.g., sea-ice, forest.

- Including additional cues that utilize different representations of polarimetric information (e.g., scattering mechanisms).

- Including additional cues that represent other sources of information (e.g., optical data, auxiliary data).

Supervised techniques typically offer better performance than unsupervised approaches. It will be valuable to compare the performance of the proposed unsupervised method with the best of the currently available supervised schemes, such as:

- A supervised segmentation technique that utilizes the statistical distribution of the data and the interaction between adjacent pixels through the Markov Random Field model (i.e., Wishart MRF) \cite{37,38}. 


Bibliography


Chapter 6. Bibliography


Appendix A

Representations of Polarimetric SAR Data

Polarization of an electromagnetic wave is the locus of the electric field vector in the plane orthogonal to the propagation direction in time. For the general case, the shape of the locus is an ellipse, which is defined by two parameters: ellipticity angle ($\chi$) and orientation angle ($\psi$). Ellipticity angle takes values within $[-45^\circ, 45^\circ]$ and the sign of $\chi$ defines the direction of rotation. When $\chi = 0$ the ellipse reduces to a line and the wave is linearly polarized. On the other hand, $\psi$ specifies the relative orientation of the major axis with respect to the vertical axis, ranges from $-90^\circ$ to $90^\circ$. $\psi = 0$ indicates vertical ($V$) polarization, and $\pm 90^\circ$ corresponds to horizontal ($H$) polarization.

A.1 Scattering Matrix

The two linear polarizations, $H$ and $V$ form the orthogonal basis that can describe the complete backscattering properties of a target. This is possible using the scattering matrix representation, which shows the relation between the incident and the backscattered wave.

\[
\begin{bmatrix}
E_h^s \\
E_v^s
\end{bmatrix} =
\begin{bmatrix}
S_{hh} & S_{hv} \\
S_{vh} & S_{vv}
\end{bmatrix}
\begin{bmatrix}
E_h^i \\
E_v^i
\end{bmatrix}
\]

(A.1)


A.2 Covariance and Coherency Matrices

For most natural targets, the two cross-polarized components \( S_{hv} \) and \( S_{vh} \) can be assumed equal. Under this assumption, known as \textit{reciprocity}, the scattering matrix can be written in a vector form as follows:

\[
k_C = \begin{bmatrix} S_{hh} \\ \sqrt{2} S_{hv} \\ S_{vv} \end{bmatrix}
\]  
(A.2)

and the polarimetric \textbf{covariance matrix} is defined as:

\[
C = k_C k_C^+ = \begin{bmatrix} |S_{hh}|^2 & \sqrt{2} S_{hh} S_{hv}^* & S_{hh} S_{sv}^* \\ \sqrt{2} S_{hv} S_{hh}^* & 2 |S_{hv}|^2 & \sqrt{2} S_{hv} S_{sv}^* \\ S_{sv} S_{hh}^* & \sqrt{2} S_{sv} S_{hv}^* & |S_{sv}|^2 \end{bmatrix}
\]  
(A.3)

Similarly, the scattering matrix can also be written in the following vector form:

\[
k_T = \frac{1}{\sqrt{2}} \begin{bmatrix} S_{hh} + S_{sv} \\ S_{hh} - S_{sv} \end{bmatrix}
\]  
(A.4)

and the polarimetric \textbf{coherency matrix} is defined as:

\[
T = k_T k_T^+ = \ldots
\]

\[
\begin{bmatrix}
|S_{hh}|^2 + 2 \Re(S_{hh} S_{sv}^*) + |S_{sv}|^2 & |S_{hh}|^2 - 2 j \Im(S_{hh} S_{sv}^*) - |S_{sv}|^2 & 2 S_{hh} S_{sv}^* + 2 S_{sv} S_{sv}^* \\
|S_{hh}|^2 + 2 j \Im(S_{hh} S_{sv}^*) - |S_{sv}|^2 & |S_{hh}|^2 - 2 \Re(S_{hh} S_{sv}^*) + |S_{sv}|^2 & 2 S_{hh} S_{sv}^* - 2 S_{sv} S_{sv}^* \\
2 S_{hh} S_{sv}^* + 2 S_{sv} S_{sv}^* & 2 S_{hh} S_{sv}^* - 2 S_{sv} S_{sv}^* & 4 |S_{sv}|^2
\end{bmatrix}
\]  
(A.5)

Both the covariance matrix, \( C \), and the coherency matrix, \( T \), are positive semi-definite (p.s.d.) and Hermitian. They can be represented by nine independent real parameters. These parameters are the three real backscattering coefficients on the diagonal and three complex correlations – with real and imaginary components – on the upper or lower triangle.
Appendix B

Test Schemes Using the Polarimetric Cue

This appendix discusses two test schemes and experimental results that were used to assess the utility of polarimetric cue when combined with other cues.

B.1 Scheme S4: Polarimetric Cue & Spatial Proximity Cue

The test scheme introduced in this section utilizes the full extent of the polarimetric information in the data set through the coherency matrix. S4 is intended for analyzing the behavior of the polarimetric cue (i.e., similarity of coherency matrices) along with the spatial proximity cue. The scaling parameter for the polarimetric cue was determined adaptively based on the maximum value of the pairwise distances within the neighborhood specified by the proximity cue. Experimental results obtained using this scheme will be presented in Appendix B.1.1. The findings and discussion obtained from those experiments were used in the design of the final scheme proposed in Section 4.2.

The following steps are performed in S4:

1. Perform multi-look processing on single look complex (SLC) data set;
2. Apply the polarimetric SAR speckle filter of Lee et. al [17] (e.g., window size = 7);
3. Form a similarity matrix using the two distance measures defined in Section 3.3 Bartlett distance, $d_B$, (3.13) and Symmetric Revised
B.1. Scheme S4: Polarimetric Cue & Spatial Proximity Cue

Wishart distance, $d_{SRW}$, (3.19). The resulting affinity matrix is one of the following:

$$W_{ij}^B = \exp \left( \frac{-d_B^2(T_i, T_j)}{2\sigma_B^2} \right) \quad (B.1)$$

$$W_{ij}^{SRW} = \exp \left( \frac{-d_{SRW}^2(T_i, T_j)}{2\sigma_{SRW}^2} \right) \quad (B.2)$$

where $T_i$ and $T_j$ represent the coherency matrices for instances $i$ and $j$, respectively. The scaling parameter $\sigma_B$ (or $\sigma_{SRW}$) is selected adaptively using one of the following:

$$\sigma_B = s \left( \max_{i,j} \{d_B(T_i, T_j)\} \right) \quad (B.3)$$

$$\sigma_{SRW} = s \left( \max_{i,j} \{d_{SRW}(T_i, T_j)\} \right) \quad (B.4)$$

where $s$ is a constant.

The number of affinity calculations per pixel is limited to a small value by choosing a neighborhood, which corresponds to the spatial proximity cue. This results in a sparse representation of the similarity matrix. Therefore, all pairwise distances and affinities are not calculated.

4. Perform the multiclass spectral clustering algorithm [60] on $W^B$ or $W^{SRW}$ as described in Section 2.1.2.

B.1.1 Results and Analysis for Scheme S4

The partitioning scheme described in Appendix B.1 – based on the similarity of coherency matrices ($T$) and spatial proximity – is used to obtain the results presented in this section. These experiments were performed to get a better understanding of the two distance measures suggested by Anfinsen et al [52], who have used these distance measures to define pairwise similarity of pixels in the feature domain and solved the spectral clustering algorithm for a subset of pixels. However, our scheme incorporates the spatial proximity cue and solves the eigenvalue problem for a sparse matrix through the MSC algorithm.
B.1. Scheme S4: Polarimetric Cue & Spatial Proximity Cue

We also attempt to automate the selection of scaling parameter using \( \sigma = s(\max_{i,j} \{d(T_i, T_j)\}) \) as described in Appendix B.1. Using a wide range of values for the scaling parameter allows one to analyze an extreme case where one cue becomes dominant over the other. Large values of \( \sigma \) (i.e., \( s = 1 \)) result in a dominant spatial proximity cue with no contribution from the polarimetric cue.

Three test regions from Westham Island data are used to demonstrate the capability of the technique and the effect of scaling parameter on the results. The constant \( s \) was varied from 1 to 0.0001 for Bartlett distance, and from 1 to 0.001 for Symmetric Revised Wishart distance.

![Figure B.1: S4 Results with Bartlett distance - Test Region # 1 (d = 7) for \( s \in \{1, 0.1, 0.05, 0.01, 0.005, 0.001, 0.0005, 0.0001\} \)](image)

![Figure B.2: S4 Results with Bartlett distance - Test Region # 2 (d = 7) for \( s \in \{1, 0.1, 0.05, 0.01, 0.005, 0.001, 0.0005, 0.0001\} \)](image)

Figure B.1 shows the results of partitioning the Test Region # 1 into six segments (i.e., \( nc = 6 \)) using the polarimetric cue (based on the Bartlett distance) along with the spatial proximity cue. There are three main characteristics that are demonstrated with these results. Very low values of \( \sigma \) represents an unstable behavior and the resulting regions are not homogeneous. This is due to the fact that when a very small bandwidth is used for the Gaussian kernel, the distances are on the tail of the curve and are
B.1. Scheme S4: Polarimetric Cue & Spatial Proximity Cue

mapped to very small affinities. On the other hand, very high values of $\sigma$ maps the distances to an almost fixed and high value of affinity. Thus, every instance becomes as likely and the information is lost. Therefore, spatial proximity becomes the dominant cue, resulting in the symmetric partitioning based on pixel locations only, and the data content (i.e., polarimetric information) is not utilized. The third is the mid-range of $\sigma$ values, where the Gaussian kernel is able to map the distances to a wider range of affinities and allows the discrimination between the partitions.

Figure B.3: S4 Results with Bartlett distance - Test Region # 3 ($d = 7$) for $s \in \{1, 0.1, 0.05, 0.01, 0.005, 0.001, 0.0005, 0.0001\}$

Similarly, Figure [B.2] and Figure [B.3] show the results for Test Regions #2 and #3, respectively. Quite consistently over the three test regions, relatively good results are obtained at the mid-range (i.e., $s = 0.05 - 0.005$).

Figure B.4 shows the results for partitioning the Test Region # 1 based on the Symmetric Revised Wishart distance, and Figures [B.5] and [B.6] show the results for Test Regions #2 and #3, respectively. Similar characteristics are demonstrated with these results. However, with the Revised Wishart distance, mid-range and low-range values are shifted to higher values. Therefore, changing $s$ between 1 and 0.001 is enough to demonstrate all three trends. Again, quite consistently over the three test regions, good results are obtained at the mid-range centered at $s = 0.01$.

Figure B.4: S4 Results with Symmetric Revised Wishart distance - Test Region # 1 ($d = 7$) for $s \in \{1, 0.1, 0.05, 0.01, 0.005, 0.001\}$
B.2 Scheme S5: Polarimetric Cue & Contour Cue & Spatial Proximity Cue

This section describes the test scheme that combines three cues: contour, spatial proximity, and polarimetric cue (i.e., similarity of coherency matrices). Similar to the scheme presented in Appendix B.1, this scheme also utilizes the full extent of the polarimetric information in the data set through the coherency matrix representation, and uses the two measures suggested by Anfinsen et al. [52]: Symmetric Revised Wishart distance and Bartlett distance.

The scaling parameter for the polarimetric cue was determined adaptively based on the maximum value of the pairwise distances within the neighborhood specified by the proximity cue. Estimating the scaling parameter using this approach was not found to be robust, which will be discussed.
B.2. Scheme S5: Polarimetric Cue & Contour Cue & Spatial Proximity Cue

Using this scheme, the effects of applying speckle filtering, and utilizing an additional input channel for the contour cue (i.e., magnitude of the co-polar correlation coefficient) were also analyzed. Comparisons were made with previously introduced schemes, S2 and S4. These results will be analyzed in Appendix B.2.1.

The following are the steps performed in this scheme:

1. Perform multi-look processing on the single look complex (SLC) data set;

2. Apply the polarimetric SAR speckle filter of Lee et al. [17] (e.g., window size = 7);

3. Form a similarity matrix for each of the input data channels (i.e., $|HH|^2$, $|HV|^2$, $|VV|^2$) using (3.10). To represent spatial proximity, limit the number of non-zero affinities for each pixel by choosing an appropriate neighborhood.

4. Form a similarity matrix using one of the two distance measures defined in Section 3.3: Bartlett distance, $d_B$, (3.13) and Symmetric Revised Wishart distance, $d_{SRW}$, (3.19).

$$W_{ij}^B = \exp \left( -\frac{d_B^2(T_i, T_j)}{2\sigma_B^2} \right)$$ (B.5)

$$W_{ij}^{SRW} = \exp \left( -\frac{d_{SRW}^2(T_i, T_j)}{2\sigma_{SRW}^2} \right)$$ (B.6)

where $T_i$ and $T_j$ represent the coherency matrices for instances $i$ and $j$, respectively. The scaling parameter $\sigma_B$ (or $\sigma_{SRW}$) is determined adaptively based on a constant $s$ and the maximum value of $d_B$ (or $d_{SRW}$) within the neighborhood specified by the spatial proximity cue in Step 3.

$$\sigma_B = s (\max_{i,j} \{d_B(T_i, T_j)\})$$ (B.7)

$$\sigma_{SRW} = s (\max_{i,j} \{d_{SRW}(T_i, T_j)\})$$ (B.8)
5. Form the combined similarity matrix, \( W^{\text{tot}} \), using one of the following:

\[
W^{\text{tot}}_{ij} = W^{B}_{ij} \prod_{b=1}^{n_b} W^{C_b}_{ij} \quad (B.9)
\]

\[
W^{\text{tot}}_{ij} = W^{SRW}_{ij} \prod_{b=1}^{n_b} W^{C_b}_{ij} \quad (B.10)
\]

where the subscript \( b \) is the index for input data channels, and \( n_b \) is the number of data channels. \( W^{SRW} \) and \( W^B \) represent the affinity matrices calculated using the Symmetric Revised Wishart distance and the Bartlett distance, respectively.

6. Perform the multiclass spectral clustering (MSC) algorithm \cite{60} on \( W^{\text{tot}} \) using the procedure described Section 2.1.2.

### B.2.1 Results and Analysis for Scheme S5

Results obtained using the partitioning scheme described in Appendix B.2 that is based on the similarity of Coherency matrices (\( T \)), contour information, and spatial proximity are given in this section. Three test regions from Westham Island data are used to demonstrate the capability of the technique. Two distance measures suggested by Anfinsen et al.: Bartlett distance and Symmetric Revised Wishart distance were used. Based on the results presented in Appendix B.1.1, the scaling parameter, \( \sigma \), was only varied at the mid-range (i.e., \( s = 0.05 - 0.005 \)). Using both distance measures the following four sets of results were obtained:

1. Contour-based similarity on speckle filtered data for \( n_b = 3 \) (i.e., input channels are \( |HH|^2, |HV|^2, \) and \( |VV|^2 \));
2. Contour-based similarity on speckle filtered data for \( n_b = 4 \) (i.e., input channels are \( |HH|^2, |HV|^2, |VV|^2, \) and \( \rho_{HHVV} \));
3. Contour-based similarity before speckle filtering for \( n_b = 3 \);
4. Contour-based similarity before speckle filtering for \( n_b = 4 \).
B.2. Scheme S5: Polarimetric Cue & Contour Cue & Spatial Proximity Cue

S5 based on Bartlett Distance

Figure B.7 shows the results for partitioning the Test Region # 1 based on the Bartlett distance \((s = 0.05)\), and spatial proximity \((d = 7)\). Contour-based similarity on speckle filtered data, where \(n_b = 3\) (i.e., input channels are \(|HH|^2\), \(|HV|^2\), and \(|VV|^2\)) is shown in the Figure B.7(a). Figure B.7(b) presents the case where four inputs are used (i.e., \(n_b = 4\), input channels are: \(|HH|^2\), \(|HV|^2\), \(|VV|^2\), and \(\rho_{HHVV}\)). When this methodology was applied without speckle filtering, Figures B.7(c) and B.7(d) were obtained for three and four inputs, respectively.

If we recall the results obtained using contour information and spatial proximity in Section 5.3.2, where Figure 5.10 showed the corresponding four cases, we can easily explain the results presented in this section. In the first case – speckle filtered data with three input channels – (i.e., Figure 5.10(a)) contour and proximity information was not enough to correctly partition the test region. Two adjacent fields were merged and a new partition was formed at the boundary of two other fields. The corresponding case with the addition of polarimetric cue based on Bartlett distance \((s = 0.05)\) is shown in Figure B.7(a). In this case, Scheme S5 provides a very good partitioning result. For Figure 5.10(b), where contour cue already achieved sufficiently good results, addition of the polarimetric cue did not improve the outcome as shown in Figure B.7(b). For the third case, where three input channels are used without speckle filtering, the result of using the contour and proximity cues was not good since two adjacent fields were merged and another field was split into two. Similar to the first case, addition of the polarimetric
B.2. Scheme S5: Polarimetric Cue & Contour Cue & Spatial Proximity Cue

cue resulted in a better partitioning, shown in Figure B.7(c). On the other hand, the last case, shown in Figure B.7(d), although improved slightly, did not reach to an acceptable performance.

![Figure B.8: S5 Results with Bartlett distance - Test Region # 2 (d = 7, s = 0.05). (a) with filtering, \( n_b = 3 \); (b) with filtering, \( n_b = 4 \); (c) without filtering, \( n_b = 3 \); (d) without filtering, \( n_b = 4 \)]

Similarly, Figure B.8 shows the results for partitioning the Test Region # 2 based on the Bartlett distance \( (s = 0.05) \) and spatial proximity \( (d = 7) \) for all four scenarios. Again, if we recall the results obtained using contour information and spatial proximity in Section 5.3.2, where Figure 5.11 shows the corresponding four cases, we can easily explain the results discussed below.

In the first case, which was shown in Figure 5.11(a), the contour and proximity cues provided a good result, except one field boundary being misplaced. The corresponding case with the addition of polarimetric cue, i.e., Figure B.8(a), shows improved results. Figures 5.11(b) and 5.11(c) already provided sufficiently good results. Therefore, including the polarimetric cue did not improve the outcome as shown in Figures B.8(b) and B.8(c). On the other hand, the last case (i.e., four input channels without speckle filtering) shown in Figure B.8(d) has the worst performance among these four, and did not improve much with the addition of the polarimetric cue at \( s = 0.05 \).

Figure B.9 shows the results for Test Region # 3 using Bartlett distance, \( s = 0.05 \) and \( d = 7 \). The results obtained from S4 (polarimetric and proxim-
B.2. Scheme S5: Polarimetric Cue & Contour Cue & Spatial Proximity Cue

Figure B.9: S5 Results with Bartlett distance - Test Region # 3 ($d = 7$, $s = 0.05$). (a) with filtering, $n_b = 3$; (b) with filtering, $n_b = 4$; (c) without filtering, $n_b = 3$; (d) without filtering, $n_b = 4$

Figure B.10: S5 Results with Bartlett distance - Test Region # 3 ($d = 7$, $s = 0.05$). (a) with filtering, $n_b = 3$; (b) with filtering, $n_b = 4$; (c) without filtering, $n_b = 3$; (d) without filtering, $n_b = 4$

For this example, Figure B.3 suggested that S4 result was dominated by the spatial proximity cue and the polarimetric cue ($s = 0.05$) did not affect the outcome. Therefore S5 result is also expected not to provide any improvement over S2 (i.e., visually Figures 5.12 and B.9 are very similar).

Figure B.10 shows the S5 results for $s = 0.005$. As can be seen from Figure B.3, the proximity cue is not dominant as it was in the previous case (i.e., $s = 0.05$). With the polarimetric cue S5 provides improvement over S2 and Figure B.10(b) offers the best result among the four scenarios.
B.2. Scheme S5: Polarimetric Cue & Contour Cue & Spatial Proximity Cue

S5 based on Symmetric Revised Wishart Distance

Figure B.11 shows the results for partitioning the Test Region # 1 based on the Symmetric Revised Wishart distance and spatial proximity ($s = 0.05$; $d = 7$). Contour-based similarity using speckle filtered data, where three input channels (i.e., $|HH|^2$, $|HV|^2$, and $|VV|^2$) are used, is shown in Figure B.11(a). Figure B.11(b) is the case where $n_b = 4$ (i.e., input channels are $|HH|^2$, $|HV|^2$, $|VV|^2$, and $\rho_{HHVV}$). When this scheme was used without speckle filtering, Figure B.11(c) and Figure B.11(d) were obtained for three and four inputs, respectively.

If we recall the results obtained using contour information and spatial proximity in Section 5.3.2, where Figure 5.10 showed the corresponding four cases, we can easily explain the results presented in this section. In the first case – speckle filtered data with three input channels – (i.e., Figure 5.10(a)) contour and proximity information was not enough to correctly partition this test region. Two adjacent fields were merged and a new partition was formed along the boundary of two other fields. The corresponding case with the addition of polarimetric cue based on the Symmetric Revised Wishart distance ($s = 0.05$) is shown in Figure B.11(a). In this case, the Scheme S5 provides very good partitioning results. For Figure 5.10(b), where the contour cue already achieved sufficiently good results, addition of the polarimetric cue could not improve the result as shown in Figure B.11(b). For the third case ($n_b = 3$ without speckle filtering), the result of contour and proximity cues was not good since two adjacent fields were merged and another field was split into two. Similar to the first case, addition of the polarimetric
B.2. Scheme S5: Polarimetric Cue & Contour Cue & Spatial Proximity Cue

cue resulted in a better partitioning, shown in Figure B.11(c). On the other hand, the last case, shown in Figure B.11(d), although improved slightly, did not reach to an acceptable performance.

![Figure B.12: S5 Results with Symmetric Revised Wishart distance - Test Region # 2 (d = 7, s = 0.05). (a) with filtering, n_b = 3; (b) with filtering, n_b = 4; (c) without filtering, n_b = 3; (d) without filtering, n_b = 4](image)

Similarly, Figure B.12 shows the results for partitioning the Test Region # 2 based on the Symmetric Revised Wishart distance and spatial proximity (s = 0.05; d = 7) for all four scenarios. Again, if we recall the results obtained using contour information and spatial proximity in Section 5.3.2, where Figure 5.11 shows the corresponding four cases, we can easily explain the results presented in this section. In the first case, that was shown in Figure 5.11(a), contour and proximity cues provided a good result, except one field boundary being misplaced. The corresponding case with the addition of polarimetric cue, Figure B.12(a) shows improved results. Figures 5.11(b) and 5.11(c) already provided sufficiently good results. Therefore, including the polarimetric cue did not improve the outcome as shown in Figures B.12(b) and B.12(c). On the other hand, the last case (i.e., four input channels without speckle filtering) shown in Figure B.12(d) has the worst outcome among these four, and did not improve much with the addition of polarimetric cue at s = 0.05.

Figure B.13 shows the results for Test Region # 3 using Symmetric Revised Wishart distance, s = 0.05 and d = 7. The results obtained from
B.2. Scheme S5: Polarimetric Cue & Contour Cue & Spatial Proximity Cue

Figure B.13: S5 Results with Symmetric Revised Wishart distance - Test Region # 3 ($d = 7$, $s = 0.05$). (a) with filtering, $n_b = 3$; (b) with filtering, $n_b = 4$; (c) without filtering, $n_b = 3$; (d) without filtering, $n_b = 4$

S4 (polarimetric and spatial proximity cues) shown in Figure B.6, and S2 (contour and spatial proximity cues) shown in Figure 5.12 are useful to explain the results presented in this section. For this example, Figure B.6 suggested that S4 result was dominated by the spatial proximity cue and the polarimetric cue ($s = 0.05$) did not affect the outcome. Therefore, S5 result is also not expected to provide any improvement over S2.

Figure B.14: S5 Results with Symmetric Revised Wishart distance - Test Region # 3 ($d = 7$, $s = 0.05$). (a) with filtering, $n_b = 3$; (b) with filtering, $n_b = 4$; (c) without filtering, $n_b = 3$; (d) without filtering, $n_b = 4$

Figure B.14 shows the S5 results for $s = 0.005$. As can be seen from Figure B.3, the proximity cue is not dominant as it was in the previous case (i.e., $s = 0.05$). With the polarimetric cue S5 provides improvement over S2. A significant effect of including the polarimetric cue is that, all four scenarios now find the correct boundary between the two fields on the lower right corner. Using Scheme S2, this was only achieved by the second case, i.e., four inputs with speckle filter, shown in Figure B.3(b).