SUBPIXEL IMAGE CO-REGISTRATION USING A NOVEL DIVERGENCE MEASURE

by

Wit Tadeusz Wisniewski

Copyright © Wit Tadeusz Wisniewski 2006

A Dissertation Submitted to the Faculty of
ELECTRICAL AND COMPUTER ENGINEERING
In Partial Fulfillment of the Requirements
For the Degree of
DOCTOR OF PHILOSOPHY
In the Graduate College
THE UNIVERSITY OF ARIZONA

2006
As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Wit T. Wisniewski entitled

Subpixel Image Co-Registration using a Novel Divergence Measure

and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy

Date: November 29th, 2005
Dr. Robert A. Schowengerdt

Date: November 29th, 2005
Dr. Michael Marcellin

Date: November 29th, 2005
Dr. Robin Strickland

Date: November 29th, 2005
Dr. Kurtis Thome

Date: November 29th, 2005
Dr. M. Susan Moran

Final approval and acceptance of this dissertation is contingent upon the candidate’s submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Date: November 29th, 2005
Dissertation Director: Robert A. Schowengerdt
STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of the requirements for an advanced degree at the University of Arizona and is deposited in the University Library to be made available to borrowers under the rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgement of the source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the copyright holder.

Signed: Wit Tadeusz Wisniewski
ACKNOWLEDGEMENTS

I wish to thank my advisor for all his efforts, making this a more complete and better quality contribution to research. I likewise thank my committee members for their additions and corrections.


I thank various students from whom I have learned, or who provided moral support. From the Digital Image Analysis Lab I wish to thank Tai Hong, Jeff Mercier, Daniel Filiberti, Zhijun He, Corey Smeaton, Qianyi Xu, and Frank Rojas.

My family made this work possible with their support and great patience, but not too much patience.

IKONOS imagery used in this dissertation contains material © Space Imaging LLC, all rights reserved.

This research was partially supported by a National Geospatial intelligence Agency (NGA) NURI award to the University of Arizona, #NMA201-01-1-2006, "Combined Spatial and Spectral Processing of Multisource Data to Generate High Resolution Thematic Maps."
DEDICATION

I wish to dedicate this work to my father, Wieslaw Z. Wisniewski, who showed me that the greatest wonders of the world are usually not yet discovered.
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF FIGURES</td>
<td>10</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>14</td>
</tr>
<tr>
<td>CHAPTER 1 INTRODUCTION</td>
<td>17</td>
</tr>
<tr>
<td>1.1 Introduction</td>
<td>17</td>
</tr>
<tr>
<td>1.2 Overview</td>
<td>21</td>
</tr>
<tr>
<td>1.2.1 Application Needs</td>
<td>21</td>
</tr>
<tr>
<td>1.2.1.1 Super-resolution</td>
<td>22</td>
</tr>
<tr>
<td>1.2.1.2 Fusion and Multi-band Processing</td>
<td>22</td>
</tr>
<tr>
<td>1.2.1.3 Change Detection</td>
<td>23</td>
</tr>
<tr>
<td>1.2.1.4 Presentation</td>
<td>24</td>
</tr>
<tr>
<td>1.2.1.5 Stereo Analysis</td>
<td>24</td>
</tr>
<tr>
<td>1.2.2 Overview of Automatic Image Co-registration Methods</td>
<td>25</td>
</tr>
<tr>
<td>1.2.2.1 Correlation Coefficient Methods</td>
<td>25</td>
</tr>
<tr>
<td>1.2.2.2 Phase Correlation</td>
<td>27</td>
</tr>
<tr>
<td>1.2.2.3 Wavelets</td>
<td>29</td>
</tr>
<tr>
<td>1.2.2.4 Optical Flow</td>
<td>30</td>
</tr>
<tr>
<td>1.2.2.5 Object Feature Detection</td>
<td>30</td>
</tr>
<tr>
<td>1.2.2.6 Subtraction and Other Simple Arithmetic Operations</td>
<td>31</td>
</tr>
<tr>
<td>1.2.2.7 Invariance Methods</td>
<td>32</td>
</tr>
<tr>
<td>1.2.2.8 Pattern Recognition</td>
<td>33</td>
</tr>
<tr>
<td>1.2.2.9 Search versus Direct Methods</td>
<td>34</td>
</tr>
<tr>
<td>1.2.2.10 Using Image Subsets</td>
<td>34</td>
</tr>
<tr>
<td>1.2.3 Overview of IT-based Co-registration</td>
<td>35</td>
</tr>
<tr>
<td>1.3 Thesis Organization</td>
<td>36</td>
</tr>
<tr>
<td>1.3.1 Background</td>
<td>36</td>
</tr>
<tr>
<td>1.3.2 Approach</td>
<td>36</td>
</tr>
<tr>
<td>1.3.3 Description of Research</td>
<td>37</td>
</tr>
<tr>
<td>1.3.4 Description of Results</td>
<td>37</td>
</tr>
<tr>
<td>1.3.5 Appendixes</td>
<td>38</td>
</tr>
<tr>
<td>CHAPTER 2 BACKGROUND</td>
<td>39</td>
</tr>
<tr>
<td>2.1 Imaging Model</td>
<td>39</td>
</tr>
<tr>
<td>2.2 Probabilistic Image Description</td>
<td>42</td>
</tr>
<tr>
<td>2.3 Divergence</td>
<td>45</td>
</tr>
<tr>
<td>2.4 Divergence Relation to Entropic Information Measures</td>
<td>48</td>
</tr>
<tr>
<td>2.4.1 Entropy</td>
<td>49</td>
</tr>
<tr>
<td>2.4.2 Fisher Information</td>
<td>52</td>
</tr>
<tr>
<td>2.5 Image Co-alignment Estimation</td>
<td>56</td>
</tr>
<tr>
<td>2.5.1 Co-alignment Effect on JMS</td>
<td>56</td>
</tr>
<tr>
<td>2.6 Probabilistic and Statistical Dependence</td>
<td>63</td>
</tr>
<tr>
<td>CHAPTER 3 APPROACH</td>
<td>66</td>
</tr>
<tr>
<td>3.1 Approach</td>
<td>66</td>
</tr>
</tbody>
</table>
TABLE OF CONTENTS - Continued

3.2 Theoretical Design ................................................................. 66
  3.2.1 Fisher Information ....................................................... 67
  3.2.2 Divergence Estimator Development .................................... 71
    3.2.2.1 Entropic Divergence ............................................. 71
    3.2.2.2 Fisher Information in Divergence Form ......................... 72
  3.2.3 Supporting Functions ................................................... 74
    3.2.3.1 Density Estimation ............................................. 74
    3.2.3.2 Derivatives .................................................... 75
    3.2.3.3 JMS Smoothing ............................................... 80
    3.2.3.4 Density Preservation and Restoration ....................... 80
  3.2.4 Application - Co-registration Algorithm ............................... 84
    3.2.4.1 Methodology .................................................. 84
    3.2.4.2 Evaluation .................................................... 86
  3.2.5 Experimental Considerations ......................................... 87
    3.2.5.1 Choice of Method and Parameters .............................. 87
    3.2.5.2 Search Strategy .............................................. 89

CHAPTER 4 DESCRIPTION OF RESEARCH ........................................... 90
  4.1 Algorithm Implementation .................................................. 92
    4.1.1 Image Generation/Preparation .................................... 92
    4.1.2 Fisher Information Estimation .................................. 94
    4.1.3 Comparison: IT Estimators ....................................... 95
    4.1.4 Spatial Transformation Looping .................................. 95
    4.1.5 Density Correction for Resampling (Applied Only in Phase 2) ... 96
    4.1.6 Detection of Alignment ........................................... 96
    4.1.7 Analysis of Results .............................................. 97
      4.1.7.1 Visual Analysis ............................................ 97
      4.1.7.2 Statistical Analysis ....................................... 99
  4.2 Test Methods ................................................................. 100
    4.2.1 Real Images ....................................................... 100
    4.2.2 Degraded Real Images (Monte Carlo) ............................. 102
    4.2.3 Synthetic Images (Monte Carlo) ................................ 103
    4.2.4 Statistical Analysis for Monte Carlo Runs ...................... 103
  4.3 Control Parameters .......................................................... 106
    4.3.1 Fisher Information Estimation .................................. 106
      4.3.1.1 Choice of Method .......................................... 106
      4.3.1.2 Derivative Bandwidth (Sigma) .............................. 107
    4.3.2 Entropy Estimation .............................................. 107
    4.3.3 JMS Density Estimation .......................................... 108
      4.3.3.1 Binsize, Range ............................................. 108
      4.3.3.2 Crop Size .................................................. 109
      4.3.3.3 Initial Offset (x, y) ..................................... 109
4.3.3.4 Search Range (x, y, shift, scale, rotation) ........................................ 110
4.3.3.5 Step Size ................................................................................….. 111
4.3.4 Degraded Real Image Generation ......................................................... 111
  4.3.4.1 Output Size ................................................................................ 111
  4.3.4.2 Downsampling Factor ............................................................... 112
  4.3.4.3 Window Center ......................................................................... 112
  4.3.4.4 Fixed/Random Relative Sampling Phase .................................... 112
  4.3.4.5 Gaussian Blurring Standard Deviation ....................................... 113
  4.3.4.6 SNR ....................................................................................... 113
4.3.5 Synthetic Image Generation ................................................................. 113
  4.3.5.1 Object Scale (Blurring Factor for Object Generation) ............... 114
  4.3.5.2 Number of Type I Classes ........................................................ 114
  4.3.5.3 Number of Type II Classes ........................................................ 114
  4.3.5.4 Number of Type III Classes ...................................................... 115
  4.3.5.5 Gaussian Blur (PSF Standard Deviation) .................................... 115
  4.3.5.6 SNR ....................................................................................... 116
4.3.6 Monte Carlo .................................................................................... 116
  4.3.6.1 Number of Runs (N) ................................................................. 116
4.4 Datasets .............................................................................................. 117
  4.4.1 Permanent Data ........................................................................... 117
  4.4.2 Volatile Synthetic Data ................................................................ 117
CHAPTER 5 DESCRIPTION OF RESULTS ......................................................... 119
5.1 Initial Tests on Fisher Information .......................................................... 119
  5.1.1 Correlated Image - i.i.d. Pixels ...................................................... 120
  5.1.2 Uncorrelated Image - Real Imagery .............................................. 121
  5.1.3 Correlated Image - Rotation, Scale and Translation Estimation. .... 123
5.2 Marginal Density Correction Tests .......................................................... 127
  5.2.1 Synthetic Test Images .................................................................. 128
5.3 Fisher Information Divergence Forms ..................................................... 129
  5.3.1 Real Images ................................................................................ 129
    5.3.1.1 TM Band3/Band4 Uncorrelated Agricultural ......................... 130
    5.3.1.2 TM Band3/Band4 Multi-Temporal Agricultural Scene ........... 134
    5.3.1.3 TM Band 3/Band 6 Agricultural ............................................ 135
    5.3.1.4 ERS SAR/TM Band 4 Commercial Phoenix ......................... 137
  5.3.2 Degraded Real Imagery - Monte Carlo .......................................... 149
  5.3.3 Synthetic Images - Correlated i.i.d. .............................................. 158
    5.3.3.1 Verification with Single Synthetic White Noise Image ............ 158
  5.3.4 Synthetic Images - Monte Carlo on Random Scenes. .................... 163
    5.3.4.1 Zero SNR behavior .............................................................. 166
    5.3.4.2 Two Type I Classes .............................................................. 168
    5.3.4.3 Two Type II Classes .............................................................. 173
TABLE OF CONTENTS - Continued

5.3.4.4 Two Type III Classes. ................................................. 180
5.3.4.5 Object Class Combination -
11 Type I; 3 Type II; and 5 Type III Classes .......................... 183

CHAPTER 6 SUMMARY AND CONCLUSIONS .................................. 192
6.1 Future Improvements ....................................................... 198
   6.1.1 JMS Density Estimation ............................................ 198
   6.1.2 Fusion of Results Presented ...................................... 199
   6.1.3 Further Experimentation with New Fisher Information Variations . 199
   6.1.4 Hierarchical Resolution Processing .............................. 200

APPENDIX A GENERATION OF SYNTHETIC TEST IMAGES .................. 201
   A.1 Rationale ................................................................. 201
   A.2 Procedure ............................................................... 203
   A.3 Examples ................................................................. 208
      A.3.1 Two Type I Classes .............................................. 210
      A.3.2 Two Type II Classes ............................................. 211
      A.3.3 Two Type III Classes ........................................... 211
      A.3.4 One Type I, Two Type II, and One Type III Class .......... 214
      A.3.5 As A.3.4 with PSF Blurring ................................... 216
      A.3.6 All Elements of Synthesis Simultaneously ................... 217

APPENDIX B GENERATION OF DEGRADED REAL IMAGE PAIRS .............. 218
   B.1 Rationale ................................................................. 218
   B.2 Implementation ......................................................... 219
   B.3 Example Utilizing IKONOS Data ..................................... 222

APPENDIX C PROOF: EQUALITY OF FISHER INFORMATION FORMS .......... 226

APPENDIX D DENSITY ESTIMATION AND HISTOGRAM MATCHING
FOR CORRECTION OF INTERPOLATION ARTIFACTS .......................... 229
   D.1 Overview ................................................................. 229
   D.2 Purpose ................................................................. 230
   D.3 Theory ................................................................. 231
      D.3.1 Density Estimation .............................................. 231
      D.3.2 Anti-Aliasing ..................................................... 233
      D.3.3 Density Matching .............................................. 235
   D.4 Implementation for this Dissertation .............................. 237
      D.4.1 Density Estimation .............................................. 237
      D.4.2 Anti-aliasing ..................................................... 239
      D.4.3 Density Matching .............................................. 241
   D.5 Examples ............................................................... 242
      D.5.1 White Noise Image .............................................. 242
      D.5.2 Ikonos Image ..................................................... 243
   D.6 Conclusion ............................................................. 244
REFERENCES ................................................................. 246
LIST OF FIGURES

Figure 2-1. Layout of imaging model components. ........................................... 40
Figure 2-2. Comparison of IT measures. .......................................................... 54
Figure 2-3. Example of JMS structure with mis-alignment progressing to co-alignment. ................................................................. 58
Figure 2-4. JMS alignment effect for a white noise image. ................................. 60
Figure 2-5. Uncorrelated, multi-spectral image pair. ....................................... 61
Figure 2-6. Image chips that are statistically independent when co-aligned and become dependent upon mis-alignment. ......................... 64
Figure 3-1. First order difference frequency response. ..................................... 77
Figure 3-2. Spatial location of first order differences with respect to the data point locations from which they were computed. ......................... 78
Figure 3-3. Modified central first order difference: Magnitude of frequency response. ................................................................. 79
Figure 3-4. Resampling scheme to minimize interpolation artifacts. .................. 84
Figure 4-1. Process flow for real images in Phase 1 and Phase 3. ....................... 101
Figure 4-2. Process flow for degraded real images (Phase 3). ............................ 102
Figure 4-3. Synthetic image flow diagram for Phase 2. .................................... 104
Figure 4-4. Synthetic image flow diagram for Phase 3. .................................... 105
Figure 5-1. Response to white noise image shifted with respect to itself. ............. 120
Figure 5-2. Uncorrelated image set from Landsat5 TM agricultural scene. ........... 121
Figure 5-3. Alignment estimation of uncorrelated images using Fisher Information, Entropy, correlation coefficient, and phase correlation. ......................... 122
Figure 5-4. ERS SAR test image. Phoenix, AZ urban area. .............................. 123
Figure 5-5. Responses to translation and rotation. .......................................... 124
Figure 5-6. Responses to translation and scale. .............................................. 125
Figure 5-7. Responses to scaling and rotation. .............................................. 126
Figure 5-8. Responses to Density corrected data. .......................................... 128
LIST OF FIGURES - Continued

Figure 5-9.  Divergence responses to uncorrelated TM agricultural scene. ........ 131
Figure 5-10. Responses to Figure 5-9 red band data co-aligned with itself. .......... 132
Figure 5-11. Responses to Figure 5-9 TM5 near infrared band data
co-aligned with itself. .................................................. 133
Figure 5-12. TM Change detection scene acquired over 5 month interval. ......... 134
Figure 5-13. Divergence responses to multi-temporal
and multi-spectral TM agricultural scene. ................................ 136
Figure 5-14. Divergence responses to agricultural scene in TM red
and thermal bands. .......................................................... 138
Figure 5-15. Responses to Figure 5-14 red band data co-registered with itself. ...... 139
Figure 5-16. Responses to Figure 5-14 thermal infrared band data
co-registered with itself. .................................................. 140
Figure 5-17. Responses to ERS SAR image shifted against Landsat 5 TM band 4 (NIR)
of urban/commercial Phoenix area. ..................................... 141
Figure 5-18. ERS SAR image co-registered with itself. ................................. 143
Figure 5-19. Responses to TM band 4 image co-aligned with itself. ................. 144
Figure 5-20. Figure 5-19 results repeated with spatial transformation
realized through bicubic interpolation. .................................. 145
Figure 5-21. Sharpening effect of finer JMS estimation. ................................. 146
Figure 5-22. Responses to bandwidth-reduced SAR co-registered with
TM5 band 4. ................................................................. 148
Figure 5-23. Test images subset from IKONOS scene. ................................. 152
Figure 5-24. Responses to degraded IKONOS image at SNR = infinity. .......... 153
Figure 5-25. Cumulative error magnitude distribution at SNR = infinity. ......... 154
Figure 5-26. Error magnitude density plot for SNR = infinity. ...................... 155
Figure 5-27. Typical responses to degraded IKONOS image at SNR = 0.5. ....... 156
Figure 5-28. Cumulative error magnitude distribution at pixel SNR = 0.5. ....... 157
Figure 5-29. Error magnitude density plot for pixel SNR = 0.5. .................... 158
LIST OF FIGURES - Continued

Figure 5-30. Performance of IT alignment estimators vs. seven noise levels. ....... 159
Figure 5-31. Responses of the six divergence-based algorithms
to a white noise image shifted with respect to itself. ................. 160
Figure 5-32. Demonstration of 0.00001 pixel co-registration precision. ....... 162
Figure 5-33. Examples of synthetic image pair, as generated for Section 5.3.4.2. .. 165
Figure 5-34. IT measure example responses at SNR = 0. ....................... 169
Figure 5-35. IT alignment estimate error magnitude density for SNR = 0. ......... 170
Figure 5-36. IT alignment estimate error magnitude cumulative distribution
for SNR = 0. ......................................................... 171
Figure 5-37. Error magnitude density for Type I objects with no noise added. .... 174
Figure 5-38. Cumulative error magnitude. Type I objects, SNR = 0.143. ........ 175
Figure 5-39. Error magnitude density for run in Figure 5-38 - Type I objects. .... 176
Figure 5-40. Type I object IT alignment responses at SNR = 0.5. ............... 177
Figure 5-41. Error magnitude density for Type I objects at SNR = 0.5 ........... 178
Figure 5-42. Cumulative error magnitude for Type I objects at SNR = 0.5 ........ 179
Figure 5-43. IT alignment responses for Type II objects. ......................... 181
Figure 5-44. Cumulative error magnitude for Monte Carlo run ................. 182
using Type II objects. .............................................. 182
Figure 5-45. Error magnitude densities for Type III objects and no added noise. .. 184
Figure 5-46. Error magnitude densities for Type III objects. Pixel SNR = 0.5. .... 185
Figure 5-47. Representative IT measure responses for Type III objects.
Pixel SNR = 0.5. .................................................. 186
Figure 5-48. Error magnitude density for combined object type mixture
and no added noise. ............................................. 188
Figure 5-49. IT measure surfaces for combined object type mixture
at pixel SNR = 0.5. ............................................. 190
Figure 5-50. Error magnitude density for combined object type mixture
at pixel SNR = 0.6. .............................................. 191
LIST OF FIGURES - Continued

Figure A-1. Image synthesis flow diagram. ................................. 202
Figure A-2. Objects in class map - four level segmentation. .......... 203
Figure A-3. Synthetic image pair: Two Type I classes. ................. 210
Figure A-4. Two Type II classes. ........................................ 212
Figure A-5. Two Type III classes. ........................................ 213
Figure A-6. Mixed class synthetic image pair. ........................... 214
Figure A-7. Mixed classes at an SNR of 7.0. ............................ 215
Figure A-8. Mixed class image pair with blurring and noise. ........ 216
Figure A-9. Mixed class image pair with blurring and noise. ........ 217
Figure B-1. Degraded image generation flow diagram. ............... 220
Figure B-2. Original IKONOS image of Maricopa agricultural fields. .. 222
Figure B-3. Correlation map for IKONOS bands 3/4 image pair. ....... 223
Figure B-4. Ikonos data prepared for downsampling and degrading. .... 224
Figure B-5. Example: Downsampled and degraded IKONOS image set. ... 225
Figure D-1. Statistical matching/restoration flow diagram. .......... 237
Figure D-2. Diagram: Density function estimation based on
    cumulative data distribution. ...................................... 238
Figure D-3. Density estimation of random data. ......................... 240
Figure D-4. Density matching example: white noise image. .......... 243
Figure D-5. Density matching example: IKONOS image. ............... 245
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-1</td>
<td>Summary of experimental attributes in each phase.</td>
<td>91</td>
</tr>
<tr>
<td>4-2</td>
<td>Responses of IT measures to co-alignment.</td>
<td>98</td>
</tr>
<tr>
<td>4-3</td>
<td>Permanent datasets.</td>
<td>118</td>
</tr>
<tr>
<td>5-1</td>
<td>Success rate for SNR = 0 case.</td>
<td>167</td>
</tr>
<tr>
<td>5-2</td>
<td>Success rates for Type I object synthetic images.</td>
<td>173</td>
</tr>
<tr>
<td>5-3</td>
<td>Success rates for Type II objects.</td>
<td>180</td>
</tr>
<tr>
<td>5-4</td>
<td>Success rates for Class III objects.</td>
<td>183</td>
</tr>
<tr>
<td>5-5</td>
<td>Success rate table for complex object type mixture.</td>
<td>189</td>
</tr>
<tr>
<td>A-1</td>
<td>Image Synthesis Nomenclature.</td>
<td>206</td>
</tr>
<tr>
<td>A-2</td>
<td>Summary of Control Parameters.</td>
<td>207</td>
</tr>
</tbody>
</table>
ABSTRACT

**Keywords**: Subpixel Co-Registration; Fisher Information; Divergence Measures; Mutual Information; Multi-modal Imaging.

Sub-pixel image alignment estimation is desirable for co-registration of objects in multiple images to a common spatial reference and as alignment input to multi-image processing. Applications include super-resolution, image fusion, change detection, object tracking, object recognition, video motion tracking, and forensics.

Information theoretical measures are commonly used for co-registration in medical imaging. The published methods apply Shannon’s Entropy to the Joint Measurement Space (JMS) of two images. This work introduces into the same context a new set of statistical divergence measures derived from Fisher Information. The new methods described in this work are applicable to uncorrelated imagery and imagery that becomes statistically least dependent upon co-alignment. Both characteristics occur with multi-modal imagery and cause cross-correlation methods, as well as maximum dependence indicators, to fail. Fisher Information-based estimators, together as a set with an Entropic estimator, provide substantially independent information about alignment. This increases the statistical degrees of freedom, allowing for precision improvement and for reduced estimator failure rates compared to Entropic estimator performance alone.
The new Fisher Information methods are tested for performance on real remotely-sensed imagery that includes Landsat TM multispectral imagery and ESR SAR imagery, as well as randomly generated synthetic imagery. On real imagery, the co-registration cost function is qualitatively examined for features that reveal the correct point of alignment. The alignment estimates agree with manual alignment to within manual alignment precision. Alignment truth in synthetic imagery is used to quantitatively evaluate co-registration accuracy. The results from the new Fisher Information-based algorithms are compared to Entropy-based Mutual Information and correlation methods revealing equal or superior precision and lower failure rate at signal-to-noise ratios below one.
CHAPTER 1  INTRODUCTION

1.1  Introduction

The need for image co-registration arises anytime there is more than one imaging of a common scene, and the positions of objects in the multiple images must be known relative to a common reference frame. The object locations may be different if the scene has changed between acquisitions, the transmission medium has changed, the imager is different, or has moved. The goal may be to create a new image (or images) that have the object positions corrected to be the same in all images. This is desirable for subsequent processing that involves point or pixel-wise operations, and for color (multiband) display. Or it may be sufficient just to know the displacement(s) to proceed with further processing. Stereo processing, and super-resolution rely on object displacement estimates to recover depth information or sampling phase, respectively.

Numerous alignment estimators are in use; however, in the more difficult cases, where the objects appear different in the two images, many automatic methods fail. Multi-modal images, in which the images are formed through disparate physical processes, such as optical and microwave scattering, are typically difficult to align. Both multi-modal and
multi-spectral imaging can produce uncorrelated images of the same scene. Alignment methods that rely on a straight-line relation, or some other assumed radiometric relation between two images that is transformable to a straight-line, will fail.

This dissertation is intended to improve alignment estimation performance on the more difficult cases, and therefore continues work on information theoretic-based (IT-based) methods, as they have been most successful so far [1], [2].

The reader is directed to Brown [3] for a general taxonomic review of successful methods. She begins with: “Registration methods can be viewed as different combinations of choices for the following four components:

(1) a feature space,
(2) a search space,
(3) a search strategy, and
(4) a similarity metric.”

In this dissertation, the feature space is the set of pixel values. The goal is to introduce and evaluate similarity metrics in that feature space for multimodal images that are robust. The search space and strategy are dictated by the needs of the similarity metrics as applied to particular image pairs, i.e., their individual characteristics, and the distortion that mis-aligns them.
Applications for using the estimated alignment information such as distortion modelling, data fusion, and interpolation leading to co-registration are not elaborated upon in this dissertation.

This dissertation is motivated by the question: *how well can a given image pair be co-registered?* No universal bound has been found in literature that would set an absolute standard by which existing methods could be assessed, and would set goals for future methods. For some specific existing methods, the Cramer-Rao Bound may be applied, to gradient-based co-registration [4].

To estimate the alignment of images, we extract information about their relative alignment based upon imaged objects. The information that is available from the physical sensing of the objects is in a continuous random field that inherently has an infinite amount of information [5]. Some of this information is lost in the physical imaging process, sampling in the spatial domain and quantization in the radiometric domain. Noise further reduces available information in the resulting images. Until discrete quantization takes place, the information content in the imaging chain is infinite.

All methods for alignment information are in fact information measures. This dissertation focuses specifically on formulas explicitly named IT measures to determine if such universal methods may best extract the needed information.
With a couple of hundred papers written on use of Entropic measures for alignment estimation, *are IT methods exhausted?* Fisher Information is different in form from the plethora of other measures, that are similar to, derived from, or generalized from Shannon’s Entropy. Literature searches do not indicate any direct application of Fisher Information in co-registration, therefore it was decided to learn how to apply it and evaluate it as a potential measure.

The philosophy taken in this dissertation is that successful methods that are robust to differences in sensor physics, and require little or no prior information about the data are more likely to be based on sound principles. If the universality of the methods can not be demonstrated through theory, then they should be verified on large sets of diverse data.

Image data are treated here as stochastic data - that is, each pixel value is a realization of a random process specific to that pixel.
1.2 Overview

1.2.1 Application Needs

The following applications have a common requirement, that object positions are to be known or corrected with an error of less than one pixel spacing. The actual requirements are dependent on both the details of the method and its implementation, and on the nature of the data. If values of adjacent pixels are substantially dependent, then they carry redundant information. Misregistration then has less impact than if adjacent pixels are non-redundant. In literature, the technical requirements of these methods are disjoint from the applications. Most applications authors ignore alignment entirely, or assume that their data are perfectly co-aligned. Some papers report error statistics from spatial transformation modelling as a global RMS error, usually based on manually-selected Ground Control Points (GCPs), providing a measure of precision, but not accuracy. This is reported without reference to the alignment requirement of their methods.
1.2.1.1 Super-resolution

Super-resolution algorithms require acquisition of data such that the phase of scene information relative to the sampling grid be varied and known. The sampling phase must be estimated if the sensor was not micropositioned to obtain a known sampling phase for each image. The precision required can be expressed as a spatial translation of less than one pixel spacing of the higher resolution (reconstruction) grid [6].

1.2.1.2 Fusion and Multi-band Processing

Data processing that combines information from more than one source is usually intended to improve knowledge of the underlying objects in the scene. To properly interpret information about the objects in each data layer, the object locations must be precisely known with respect to each layer’s reference frame, and then be related to a common reference frame. An example dataset consists of multitemporal, multispectral landsat images as a component of the North American Landscape Characterization (NALC) project. The objective is +/- 0.5 pixel mis-alignment image-to-image [7].
1.2.1.3 Change Detection

Typical change detection consists of operations that compare information pixel-wise in layered data that has a common spatial reference frame. Mis-alignment of the layers results in some errors from mismatched locations.

Several authors have studied the implications of mis-alignment for change detection. Dai [8] explains how errors of over and underestimation of change detection occur and how correct responses are diminished. He also reports a sensitivity to mis-alignment of 10% error for 0.2 pixel spacing mis-alignment in two of his papers on the topic [8], [9]. This sensitivity is also reported by Townshend [10] for error in estimating NDVI for a variety of natural and agricultural areas with moderate vegetation density. For sparse desert vegetation he reports 10% errors at mis-alignments of 0.5 to 1.0 pixel spacings. A statistical framework for computing change detection error under given data is derived by Roy [11]. Corrective processing of detected change is presented by Bruzzone [12, [13] with adaptive decision making using Change Vector Analysis, and by Beauchemin [14] using adaptive median filtering to post-process detected change.
1.2.1.4  Presentation

Visual presentations performed in color or color visualizations are degraded if the three bands that are chosen for color display are not well aligned. The human eye readily sees color fringing if they are not. Townshend [10] states that 0.5 to 1.0 pixel mis-alignments are satisfactory for visually presented results. The author of this work believes that mis-alignments of less than 0.5 pixels are visible for some imagery.

1.2.1.5  Stereo Analysis

Recovery of depth (or height) from stereo pairs (images deliberately taken from different, but known locations) is done at view angles nearly perpendicular to the to-be-recovered dimension. This geometry helps maintain similarity between the images, but foreshortens the depth dimension. To resolve depths with resolution approaching that available on the horizontal object plane, sub-pixel estimates of object shifts caused by varying object heights are required.
1.2.2 Overview of Automatic Image Co-registration Methods

1.2.2.1 Correlation Coefficient Methods

The ubiquitous correlation methods [15], [16] rely on the assumption that the correlation coefficient between two images is maximal at co-alignment, and the correlation peak indicates location of best alignment. This is certainly true if the two images are identical. If they differ because of differential noise or some changes in the scene, the correlation peak at co-alignment is less distinct, and may be surpassed by additional and undesirable peaks. The sharpness of the peak is limited by scene auto-correlation and differential noise. If the images are acquired by different physical processes, the results are likely to be more difficult to interpret. Finally, if the two images are uncorrelated at co-alignment, correlation-based methods fail. Correlation coefficients are indications of the extent to which a straight line relation exists between two images. They are statistical measures of common information, limited to information present only in moments of second order.

\[
R(u, v) = \frac{\sum \sum g_A(i, j)g_B(i - u, j - v)}{\sqrt{\sum \sum g_A^2(i, j) \cdot \sum \sum g_B^2(i, j)}}
\]  

(1.1)
A great advantage of discrete cross-correlation is that computation of the correlation coefficient can be obtained from only two computations of an FFT and three inverse FFTs.

\[
R(u, v) = \frac{F^{-1}(F_g \cdot F^*_{g_B})}{\sqrt{[F^{-1}(F_g \cdot F^*_{g_A})]_{(0,0)} \cdot [F^{-1}(F^*_g \cdot F_{g_B})]_{(0,0)}}}
\]

\(F\) is the Fourier Transform operator, and \(g\) are two images \(A\), and \(B\). The denominator is a geometric mean of variances that may be computed directly instead of via Fourier Transforms. These computations are very fast relative to most other algorithms, and need to be done only once to obtain a peak that directly indicates where the images co-align.

The width of the correlation peak is on the order of twice the width of the objects that produced it. Most imagery has spatial spectral energy concentrated at low frequencies [16], resulting in a broad correlation peak. Noise features on this peak can define the maximum, thus the location of precise co-alignment is obscured. Sharpening the peak by high-pass filtering, whitening, and decorrelating the data has been suggested [15]. Sharpening of the peak is beneficial only if noise superimposed on the peak is not accentuated to such a degree that alignment precision is degraded.
1.2.2.2 Phase Correlation

This method is derived from the fact that translations in object position manifest themselves as a rotating unit vector in complex Fourier space, with frequency of rotation (rate of phase change) proportional to the shift. The rotating vector multiplies the unshifted Fourier Transform result. If the phases of the same objects are compared between two images the relative displacement is measured. To achieve this, the cross-spectrum is computed because the phases subtract in the computation. Whitening per Section 1.2.2.1 is simply accomplished with division of the cross-spectrum by its magnitude, leaving only phase information. This result is a rotating complex vector with unit magnitude who’s spatial frequency is proportional to the difference of object positions. Inverse Fourier transforming recovers this relative phase term frequency as an impulse at a position that directly indicates the relative translation between objects [17, [18]. The method is relatively fast as it requires three FFT computations (Eq. (1.3) defines $M$, the phase correlation function).

$$M(x, y) = F^{-1} \frac{Fg_A(x, y) \cdot F^*g_B(x, y)}{|Fg_A(x, y) \cdot F^*g_B(x, y)|}$$ (1.3)
$F$ is the Fourier Transforms operator that is applied to images $g$ defined on an $(x,y)$ space. The coordinates that maximize $M$ are the estimate of correct alignment. For two images that are identical, except for a translation, the method precisely indicates the correct shift. Components in the images that are not identical in both contribute to a relative Fourier phase, resulting in shift estimation error and possibly false peaks that compete with the correct one.

Several other issues that limit performance and applicability of Phase Correlation were previously noted by the author. The whitening enforces contribution of phase information from all spatial frequencies, even if the objects are not composed of some frequencies. Assuming some noise is present at all frequencies and images contain mostly low frequency energy, then most of the phase information may come from noise. To mitigate this problem, the spatial bandwidth of the whitened cross-spectrum may need to be limited, to encompass only the meaningful part of the spatial spectrum, i.e., whitening is not optimal. This may be done by simple, circularly symmetric low pass filtering, or more data specific filtering may be used such as adaptive Wiener filtering. The issue is finding a compromise between sharpness of response and noise rejection. Phase Correlation is susceptible to phase distortion from non-linear phase filtering or other processes. The author has noted that some interpolation schemes embed a phase distortion pattern on an image that resembles a beat frequency pattern. This imposes a periodic pattern on the phase of the
cross-spectrum. The pattern competes with the co-registration information, and may bias or throw off the results. JPEG compression has also been noted for imprinting phase anomalies in images.

1.2.2.3 Wavelets

Wavelet transforms have been mostly used as progressive resolution reduction filters. This allows a coarse and rapid search to be made first, over a large search range. Then finer resolution results can be obtained without the necessity for searching a large range. The spatial decomposition also provides a filtering opportunity to select ‘better’ spatial components [19]-[25].

Complex wavelets can be used to estimate alignment because they encode local shift information in their phase, much like Fourier Transforms do globally. For each resolution level, a contribution to the shift is estimated. These then must be reconciled to one result. Another way to directly apply wavelets is by creating a similarity metric that assesses wavelet component matches [26].
1.2.2.4 Optical Flow

Optical Flow is the migration of objects in a common image frame described pointwise throughout an image. It is not necessarily a specific algorithm but a point of view suited for arbitrary motion estimation of objects and/or the camera. Clearly, mis-alignment of imaged objects constitutes Optical Flow, and minimizing it leads to co-alignment. The pointwise approach allows for co-registration despite complex distortions. Rougon [27] demonstrate a co-registration, and others [28]-[32] show ways to Compute Optical flow.

1.2.2.5 Object Feature Detection

If scene objects or parts thereof can be delineated, then their spatial attributes can be estimated and compared between images to assess alignment. These salient features include but are not limited to edges/boundaries, centroids enclosed by boundaries, active contours [33], inflection points, and crossing lines. Corners are a much sought after feature [34], [35]. The choice of spatial feature type and individual feature should be done according to which ones are invariant to spatial distortions between the images. The features should also have a fixed position with respect to scene objects, and have to be compared to determine which ones in the first image are the same as ones in the second image (the “correspondence matching problem”) [36], and if they are indeed invariant.
1.2.2.6 Subtraction and Other Simple Arithmetic Operations

These methods are found in Maintz [37]. If two aligned images of the same scene have the same radiometric scale, then, on average, accumulated subtraction of them should be zero. The actual accumulated absolute value of the difference at each pixel position is not zero, but it is small if the images are similar. A mis-alignment throws off this similarity, thus substantially increasing the accumulated absolute value. A search for correct alignment may be done by performing the accumulation at different alignments and identifying a minimizing alignment.

This method is very fast, as it is based on pixel-wise subtraction and summing of absolute values. The method may be substantially accelerated if the processing is prematurely terminated due to a too-high sum up to that point in the sequential summing process - a threshold is set to reject a trial alignment if the accumulation already reaches a value that indicates bad alignment before summing is completed.

Pixel-wise division of two images, of the same scene from sensors with proportionally shaped responsivity curves, that may differ in overall contrast and have the same spatial response, should result in a constant image equal to the proportionality constant. Sensor noise or variation of objects in the scene results in some variation around this constant.
Mis-aligned images will produce a different value than this constant at the locations where the objects fail to overlie. Co-registration is accomplished by relative spatial adjustment that minimizes variation as indicated by the quotient image variance. [37].

Another simple method involves subtraction, however, the number of sign changes is counted rather than the actual difference. The number of sign changes is maximal upon co-alignment because they are presumed to be due only to random noise. Evaluating a sign change image for statistical dependence between the input images is another variation. If the sign changes are due to overlapping image features in mis-aligned images, then they are more pattern-like than if they are caused by random noise. [37].

1.2.2.7 Invariance Methods

These are augmentations of other methods to make them robust under combinations of mis-alignment causing spatial distortions. For example, most methods handle translation or rotation as long as only one is unknown, but fail if both are to be determined. This is because rotation transforms spatial coordinates not uniquely according to a rotation angle, but also around the center of rotation that may be displaced by translation. Methods that are designed only for translation are usually adapted to rotation by means of transformation to polar coordinates [23], [38], [39], through the Radon Transform [40], [41], re-mapping to polar coordinates [18], [42], [43], higher order statistics [44]-[46], and wavelet transforms [22], [23].
Some measures dealing with unknown scale, rotation, and translation are called invariants, as they produce the same value despite these distortions. Moment statistics that are invariants have been identified. These are combinations of normalized central moments that make them invariant to the aforementioned transformations. Invariant moments attributed to Hu are listed in [15], and applications of Zirneke Moment invariants are in [47]-[51].

1.2.2.8 Pattern Recognition

Pattern recognition involves locating the position of an image chip in an image, that is in a second image, and contains the same objects imaged in the chip. Defining chip to be a substantial portion of the whole images or utilizing many chips to span whole images allows pattern recognition to serve as a co-registration methodology. The chip may be most of the image, and centered allowing a small margin for variable overlap over the search range. The recognition mechanism may be based on shape recognition. The definition of ‘shape’ and an introduction to shape science is presented by Pavel [52]. For imagery in which line features are detectable, the Hough Transform can be used as a line detector [53], [54].
1.2.2.9 Search versus Direct Methods

The image and phase correlation methods presented in sections 1.2.2.1 and 1.2.2.2 indicate the correct alignment in one trial of computation and are therefore fast. Image correlation may be implemented as a search, but it is usually computed using FFTs such that the entire cross-correlation function is obtained at once. The other methods provide their indication of alignment upon completion of a wide and fine enough spatial search to distinguish the correct co-alignment indication.

1.2.2.10 Using Image Subsets

The above discussion avoided details about what kind of spatial distortion is being estimated. For rigid transformations the entire image area may be used to estimate the shift. Distortions described by affine transformations require at least four independent data points for correction. To obtain them, four, or more non-overlapping subsets of one image may be located in the other. A relative alignment estimate obtained for each provides information that can be used the same way as positions of ground control points (GCPs). The center locations of the co-registered subsets emulate GCP coordinates.

Higher order polynomial distortions require more matches. Arbitrary elastic distortions become difficult to estimate because a large number of estimates is needed. If they are to be independent they are to be disjoint, then the area, and thus sample size in number of
pixels diminishes due to subdivision, compromising estimation quality for each point. If
the estimates are allowed to be dependent, then the alignment estimator may be run in a
moving window, producing a map of the distortion on the image’s pixel grid.

1.2.3 Overview of IT-based Co-registration

In most cases, the joint statistics of two images attain a condition that is unique to co-
alignment. Identification of this condition requires appropriate statistical measures. The
high sensitivity of IT-based measures to slight changes in density functions explains why
they are currently of research interest.

The reader is referred to Pluim [55], for a comprehensive survey of Entropy-based meth-
ods. His application arena is only medical imaging, but that is where most development
has taken place. This dissertation demonstrates more universality of IT methods. In the
next chapter, IT-based co-registration is discussed in further detail.
1.3 Thesis Organization

1.3.1 Background

Theoretical background and tools are described in this chapter. The IT measures of Entropy, Fisher Information, and divergence are explained. The Joint Measurement Space (JMS) is described as a core statistical entity for condensing information from two images into a two dimensional (2D) space. Relations between alignment of two images and features in the JMS are explained, as is their measurement. Finally, the relation between statistical dependence and image alignment is discussed.

1.3.2 Approach

In the Approach chapter, the theoretical derivations of IT measures are made, and how they lead to the goal of this dissertation. Here, the new divergence measures that are the core contribution of this dissertation are formed. To realize the algorithms, supporting functions and methods are developed. Experiments designed to realize the theory are set forth.
1.3.3 Description of Research

Computational realizations of the algorithms and experimental methods are described in detail. Background is also given to enable interpretation of the results in terms of their experimental parameters. Details of each functional module for alignment estimation, as well as assessment of performance in qualitative and quantitative terms, are specified along with corresponding key control parameters. Datasets used for testing are also enumerated. This chapter also describes the development history of this work in terms of changing methods motivated by experimental results.

1.3.4 Description of Results

Most of the chapter concentrates on qualitative and quantitative assessment of experiments with real and synthetic datasets applied to the newly developed divergence measures. A chronological history of results showing evolution to the final IT measures is also presented. A variety of datasets are used for testing the alignment estimators, most of which are difficult or impossible to process with correlation-based methods. Testing is done on real imagery for which alignment truth is not known, reduced resolution real imagery for which alignment truth is known within some bounds, and randomly-generated synthetic imagery that is perfectly co-aligned.
1.3.5 Appendixes

Some details left out of the main text that need further explanation are described here. In Appendix A, the procedure for generating synthetic image pairs is described. Appendix B describes degradation of real imagery to serve as test image pairs. Appendix C is an equivalence proof of two Fisher Information forms that are not intuitively perceived as the same. Appendix D describes a methodology for estimating the global probability density of image pixels, and for restoration of that density after it has been altered by interpolation. The density estimator is more generally applicable, and is utilized in this dissertation to evaluate accuracy and precision of the alignment estimates.
CHAPTER 2 BACKGROUND

2.1 Imaging Model

The image model used for this research is intended to support a random scene assumption. It is presented as an aid for explaining the results obtained from real imagery, and is the basis for generating synthetic image pairs for testing methods. The main objective is the spatial representation of object boundaries and textures within them. The scenario is terrestrial, remotely-sensed imagery, however, the imaging of surface objects of most any kind at a standoff distance (large enough to suppress perspective changes between images) is supported by this model.

Geophysical and anthropogenic causes are the foundation of this model - they place the various objects in the scene. Measurements of these objects is assumed as the goal of acquisition. Imaging of the objects begins with some physical process, such as emission or scattering of radiation. Each object’s sensed characteristics are describable by some random model and are generally referred to as texture. Each object’s shape is treated as random, allowing for visually shapeless regions to be treated as objects. The objects may be one of a kind, or be replicated in the scene with the same textures but various shapes. This model defines every location in the scene as containing some object as there is no
background. Spatial overlaps of objects are not allowed as viewed from the sensor, so boundaries between objects are viewed as infinitely sharp. This is shown in a component diagram - Figure 2-1.

![Diagram of imaging model components]

*Figure 2-1 Layout of imaging model components.*

An example is forest with clear-cuts. The causes for the forest are the combined geology and ecology that resulted in tree growth at that forest location; the causes for the clear-cuts are anthropogenic, i.e., the choice of where to cut down forest and the execution. An image of the above has two types of objects with two distinct textures.

The second part of this model describes the radiation transmission and image formation in the sensing process. Deterministic components of the underlying radiative transfer model are assumed constant, as they do not carry spatial information (such as position in the
images) about the objects in the scene, nor the image forming process. A deterministic medium may be arithmetically accounted for and a uniform medium is assumed. For example, path radiance and atmospheric transmission can be removed through subtraction and division, respectively, to obtain radiometrically correct texture with respect to the object surfaces. Changes in illumination intensity may also be corrected by division. Directional reflectance or emissivity are treated as properties of the objects.

The image is formed by optics possessing a point spread function (image forming optics PSF) onto a focal plane populated by square uniform detectors that fill the focal plane area without gaps between detectors. This complete spatial response is described by a total system PSF. The sampling due to the discrete detectors may introduce spatial frequency aliasing.

Description of imagery for which the methods of this dissertation are intended may be accomplished by idealization of an image pair through the following assumptions: (1) The spatial representation of objects is identical in both images. This requires system PSF footprints on the object plane to match; (2) Both images depict identical object boundaries in the scene; (3) The sensing processes for each image may be either different or the same. Two different real image forming systems, particularly ones that utilize different physical sensing means are unlikely to have identical PSFs. To attempt to meet assumption (1), some matching of PSFs is needed, i.e., sharpening of the broader PSF and/or blur-
ring of the sharper one. Resampling necessitated due to different pixel scaling between sensors may also impact spatial representation, and may be correctable only approximately.

The above idealizations are intended to foster development of alignment estimators, define image synthesis methods, constrain discussion regarding their properties, and limit the scope of imagery considered appropriate for processing. Alignment estimators designed under these assumptions may function even if the assumptions are not met, and testing should not be limited only to imagery that meets them.

2.2 Probabilistic Image Description

A probability measure exists that can entirely describe the result of a single image acquisition. It is a function of the sensing process from object properties to final data quantization. It has an N-variate joint density function for the N random variables that are to be realized as N pixel values:

\[ P_{\{X_1, X_2, \ldots, X_N\}}(x_1, x_2, \ldots, x_N) = p(\{X_1, X_2, \ldots, X_N\} = \{x_1, x_2, \ldots, x_N\}) \]  (2.1)
\( P \) is the probability law for vector of random variables \( \mathbf{X} \), that take on vector pixel values \( \mathbf{x} \) with probability \( p \). This density encompasses all information that can be obtained by realizing this image. Two dimensional information is spatially encoded through location coordinates of the N pixels on the pixel grid. In this probability description, the probability law neither implies nor precludes spatial relations of the random variables, hence the 2D coordinates are implicit. For two images considered as a joint pair, the probability law is:

\[
P_{\{A_1, A_2, \ldots, A_N; B_1, B_2, \ldots, B_N\}}(a_1, a_2, \ldots, a_N; b_1, b_2, \ldots, b_N) = \\
p(\{A_1, A_2, \ldots, A_N\} = \{a_1, a_2, \ldots, a_N\}; \{B_1, B_2, \ldots, B_N\} = \{b_1, b_2, \ldots, b_N\})
\]

which is 2N dimensional for a pair of images. In this case, there are two random vectors, \( A \) and \( B \) that have pixel indexes in a relative spatial correspondence between the images, i.e., the location of pixel with value \( a_i \) is the same as for pixel with value \( b_i \) on a common spatial reference frame. This is not just a product of two Eq. (2.1) probabilities if the images are in any way probabilistically related over N common pixels.

The 2N dimensionality is generally huge. Fortunately, real imagery does not have mutually independent pixels within each image, nor are the image pairs independent if a relation exists between them (such as portrayal of a common scene), making the effective dimensionality \(<2N\) due to redundancy of information. For real imagery, \( N \) is in the \( 10^6 \) to
$10^9$ range, therefore, even with a substantial reduction from 2N due to redundancy, the raw data may not have a realistically manageable dimensionality for practical statistical image processing.

Further dimensionality reduction is accomplished by various means. Local processing in moving or fixed neighborhoods allows reduction of both the dimensionality and the amount of data. Markovian assumptions on the spatial dependence in the data are one tool and justification for doing so. In this dissertation, only pairwise pixel statistics are considered globally for the two images, thus reducing the measurement space to 2D. This new global 2D random variable is an expectation of Eq. (2.2) conditioned on pixel positions having the same position in coordinate systems of their images. Pixel probabilities in the Joint Measurement Space (JMS) are defined upon two dimensions that represent pixel values, one from each image, at all common locations of the pixel pairs in a common reference frame. This global pixel value probability measure is shared by each pixel pair regardless of location in the image, resulting in a near complete loss of spatial information. The only remaining spatial consideration is the co-location of the pixels pairs in the two images.

Note that it was not assumed that pixel pair values are measurements of exactly the same points on objects, because objects may be mis-aligned relative to each other in a common reference frame. One of the goals of this dissertation is to examine how the global joint
pixel statistics are affected by commonality of objects in the image frame coordinates due to co-alignment compared with varying mis-alignment, and find measures for quantitative estimation of the alignment’s effects upon the joint global probability.

Probability measures exist for a given measurement scenario as abstract concepts. Realization of the random processes of remote sensing results in image acquisition - data, and the two dimensional probability space is replaced by actual values in the JMS in which image pairs are quantified. Given JMS data, the joint global probability density may be estimated.

### 2.3 Divergence

Statistical divergence is a means of comparing probability densities. It is defined by various measures that indicate by a single value how much probability densities differ. Such measures are not unique.

The most commonly used divergence is defined by Kullback and Leibler [56] (known as the K-L Divergence) and referred to here as $I_{KL}$.

\[
I_{KL}(p, q) = \sum_{a, b} p(a, b) \log \left( \frac{p(a, b)}{q(a, b)} \right)
\] (2.3)
It is shown here in two dimensional form operating on joint discrete densities \( p \) and \( q \) of random variables that take on values \( a \) and \( b \). Note that K-L Divergence is not symmetrical with respect to the two densities, i.e., the densities are not interchangeable.

This two dimensional form may be used to compare two global joint densities of pixel values in two images. Two images may have various degrees of probabilistic dependence that may be ascertained from the JMS by means of a statistical divergence measure. In this case, the density comparison is made instead between the density of the actual image pair and a density of hypothetical independent images. To obtain the independent image pair joint density, two approaches are possible. One approximation is to spatially randomize the pixel positions in one image so that, on average, there is no deterministic spatial correspondence of objects between the two images. The other way is to estimate the individual densities of the images (marginal densities of the joint density). If the images are independent, then the joint density would be identical to the product of the marginals. A synthetic joint density is then defined from the actual marginal densities of dependent images by performing the multiplication \( p(a) \cdot p(b) \). The actual joint density \( p(a, b) \) can be compared against this synthetic density by computing a divergence.
Estimating divergence between actual and hypothetically independent-image joint aggregate densities using the K-L measure is done as follows:

\[ I_S(A, B) = \sum_{a,b} p(a, b) \log \left( \frac{p(a, b)}{p(a) \cdot p(b)} \right) \]  \hspace{1cm} (2.4)

The measure will be zero if the images are independent, because the ratio

\[ \frac{p(a, b)}{p(a) \cdot p(b)} \]  \hspace{1cm} (2.5)

is then one so

\[ \log \left( \frac{p(a, b)}{p(a) \cdot p(b)} \right) \]  \hspace{1cm} (2.6)

becomes zero. This mutual information is symmetric with respect to the order of the density functions. The divergence in Eq. (2.4) is also identically the Mutual Information as defined in terms of Shannon Entropy, hence the subscript S in \( I_S \).
If the actual joint aggregate density is more granular than the hypothetical independent density, the term in Eq. (2.5) will vary +/- around 1.0. Convexity of the log function will make the K-L measure more negative.

K-L divergence is a particular case of a divergence measure family known as Csiszar Divergences [57], [58].

\[
I(A, B) = \sum_{a,b} p(a, b) g \left( \frac{p(a,b)}{q(a,b)} \right); \quad g \text{ convex}; \quad g(1) = 0
\] \hspace{1cm} (2.7)

In the next chapter, alternatives to the Csiszar Divergences will be presented as the main contribution of this work. They are forms of Fisher Information applied on two dimensional data - divergences inspired by Fisher Information.

2.4 Divergence Relation to Entropic Information Measures

Information measures are functionals operating on probability measures of a random process that indicate the potential information in realization of the random process.
2.4.1 Entropy

The original intent and subsequent applications of Shannon Entropy, Eq. (2.8), are information measurement in data communication systems.

\[
H(X) = -\sum_{x} p_X(x) \log(p_X(x))
\]  

(2.8)

\(p_X(x)\) is the discrete density function of random variable \(X\) that takes on values \(x\). The plethora of Entropy applications is explained by the units of Entropy which are bits, if a base 2 logarithm is used in Eq. (2.8). Entropy is the average amount of information expressed in the number of ‘yes’ or ‘no’ answers that is gained by realization of a random process. This is a ‘forward’ measurement or estimation problem in which the probability density is known or estimated and its Entropy is computed.

The second most common use of Entropy is in the reverse problem of density estimation from data. Discovery of a density function with more parameters than the number of data points requires some presumption on the unknown parameters. The Maximum Entropy Method (MEM) is a way of minimizing the degree of presumption as defined and measured by Entropy. Visually, the presumption is any roughness in the shape of the density
function that is not indicated by the data, making MEM a smoothing interpolation procedure. The unknowns in the density being estimated are adjusted in MEM to maximize the Entropy of the resulting density within constraints of the known data points [59], [60].

Entropy operates on ‘probability masses’ that are discrete in nature and nominally described - their order is irrelevant. Entropy can be intuitively viewed as a measure of diversity of probability masses, where diversity is defined as the number of different abscissa values with non-zero probability. Adding values with non-zero probability masses increases the Entropy because their diversity increases. It is less sensitive to the values of the probability masses because their variability is subdued by the logarithm function.

The Mutual Information measure presented before in Eq. (2.4) may be written in terms of joint Entropy of random variables $A$ and $B$, and their marginal Entropies [61].

$$I_{S}(A;B) = -H(A, B) + H(A) + H(B) \quad (2.9)$$

The individual images are fixed upon acquisition, so their entropy estimates from the data are constant, and do not vary if the image coordinates are transformed. If $H(A, B)$ is to be used for alignment detection, $H(A)$ and $H(B)$ are constant during a spatial search for best
correspondence between the images. $H(A, B)$ variation solely determines the variation of $I_S(A;B)$ that may indicate best alignment. The optimization of Entropy is then equivalent to optimization of Entropic Mutual Information, making computation of the marginal Entropies unnecessary.

In problems where at least one image is altered after acquisition, as is the case for applying spatial transformations and re-sampling to co-register the two images, the marginal Entropies are not constant. The same would be the case for pointing motion of one camera to align its output with a previous image. These scenarios are implicit in the real and synthetic image examples in this dissertation. The changing alignment of the images alters their content in the overlap where the images are treated as joint data. Resampling alters the data density because its results are taken from linear mixtures of numerous pixels, thus not preserving the original distributions. A lesser impact is caused by nearest neighbor resampling that always preserves an original pixel value when assigning output values. It, however, does not preserve the overall distribution because it may assign some values in duplicate and may not assign some others at all.

One way of precisely obtaining Mutual Information under such changes, is by complete Entropic assessment of both the marginal densities and the joint density. An alternative way is to radiometrically correct the individual images so that minimal change in the marginal density occurs under minor spatial alteration of one or both images from resampling.
2.4.2 Fisher Information

Fisher Information is a lesser known information measure, that is commonly defined by two equivalent expressions.

\[
I_\theta(f(y|\theta)) = \sum_{\forall m} \int f(y|\theta) \left[ \frac{\partial \log f(y|\theta)}{\partial \theta_m} \right]^2 dy = -\sum_{\forall m} \int f(y|\theta) \frac{\partial^2 \log f(y|\theta)}{\partial \theta_m^2} dy
\]  \hspace{1cm} (2.10)

\(Y\) is a random variable vector that takes on values \(y\) with density \(f\). \(\theta_m\) is a single parameter in parameter vector \(\theta\). The equality between the two forms is proven in Appendix C.

This measure is named in honor of Ronald A. Fisher, who developed the formula as a way to determine variance of Gaussian distribution parameter estimators [62]. He asserted that a parameter estimator’s variance can not be less than the reciprocal of Fisher Information. Cramer and Rao further verified that the reciprocal of Fisher Information actually sets the lower bound on a parameter estimator for most distributions [63].

As far as the author knows, Fisher Information does not have another published forward application (where the joint distribution is known), unlike Entropy. B. Roy Frieden has pioneered a density estimation methodology of using Fisher Information in a reverse methodology that parallels MEM and is called Extreme Physical Information (EPI) [64], that is based on minimizing Fisher Information within constraints of data. His results indi-
cate that Fisher Information is a stronger measure of bias resulting from presumed information in a probability density function compared to Entropy. His results show density estimates from real data that are ‘smoother’ than those obtained from MEM, as demonstrated by Hawkins [65]. This is explained by Fisher Information’s much greater sensitivity to fine structure in the shape of a density. A smoother estimate is considered less presumptuous because there is less variation in the density shape that is not supported by data.

The parameters seen in the formula of Eq. (2.10) were originally intended to be that of a probability distribution; however Fisher Information with regard to any quantity may be computed and may be meaningful even if it is not with respect to a ‘named’ parameter. It may also be computed with respect to some combination or function of recognized parameters.
Fisher Information and Entropy may be compared by looking at their contributions to the respective expected value integral. In Figure 2-2, is shown the Differential Entropy defined as Eq. (2.11) for densities of continuous random variables

\[
H(X) = \int f(x) \cdot \log f(x) \, dx \quad (2.11)
\]

and the Fisher Information contribution function, given a Gaussian density. Note that the maximal contribution of Entropy is at the mode of the distribution, while the Fisher Information contribution is along the slopes. This is explained by Frieden [66], who derived a third form of Fisher Information that is closer in appearance to Differential Entropy.

\[
I(X) = \int f''(x) \cdot \log f(x) \, dx \quad (2.12)
\]
This illustrates that while the Entropy integrand is the log of the density function weighted by the density function, Fisher Information has instead a weighting by the second derivative of the density. In this example, the Fisher parameter is $x$, the argument of the density function, and differentiation is with respect to $x$.

In addition to the single parameter form of Fisher Information, a two parameter version that defines the Fisher Information Matrix can be found in literature [67], Eq. (2.13).

\[
I_{\theta_m, \theta_n}(f(y|\theta)) = \int f(y|\theta) \frac{\partial \log f(y|\theta)}{\partial \theta_m} \cdot \frac{\partial \log f(y|\theta)}{\partial \theta_n} dy
\]  

(2.13)

It is used to compare how much a density’s sensitivity to one parameter relates to sensitivity to another parameter. It can also be viewed as a measure of information in a density function in common with respect to two parameters. Note that it is the same expression found in the left hand side of Eq. (2.10), if the two parameters are one and the same.

In order to apply Fisher Information to data which is necessarily sampled in the JMS, a discrete approximation to the derivative is needed. The type and form of derivative approximator that will work best depends on the desirable part of the frequency content in
the JMS, as derivative approximators differ in the way they respond to higher frequencies. Further bandwidth control may be applied to the derivative approximation via filtering of JMS input.

The main result of this effort is development of a divergence method that is based upon or inspired by Fisher Information. It is presented in the next chapter.

### 2.5 Image Co-alignment Estimation

One purpose of co-alignment estimation is co-registration - to spatially transform two images of the same objects such that the objects will become co-aligned in a common reference frame. It may also be used as input to data fusion algorithms that require precise position information in the original images.

#### 2.5.1 Co-alignment Effect on JMS

To understand how co-alignment manifests itself in the JMS, let's look at the following examples. In Figure 2-3a, a pair of images is formed by similar sensors. The image contains digital number (DN) values due to three objects. Responses to each object are shown here placed in the JMS as a *locus* of non zero JMS values illustrated by an oval. When the objects are co-aligned, they perfectly overlap, and their JMS contributions are the same
from each sensor, on the 1:1 diagonal. There are only three loci of values as the objects do not intermix in the JMS. When mis-aligned, each object in an image overlaps to some degree with itself and remaining objects in the other image. Now, for a given misalignment, there are 9 combinations producing nine weaker loci in the JMS due to all combinations of matched and unmatched objects coinciding in various pixels. Subpixel shifts adjust fractional mixing of object signatures providing a continuum of possible locus positions. Visually, the difference appears in a scattergram (a visualization of the JMS) as a weak, diffuse texture if the objects are mis-aligned, and becomes a stronger, coarser texture at alignment.

Figure 2-3b shows the JMS effect in the case of multi-modal sensing, i.e., the responses of the sensors to the same objects are different. The fundamental effect of consolidating the loci is the same and the number of loci is the same, however overlap of the loci does not occur on the JMS diagonal, and there is a higher opportunity for some loci to overlap even if the images are mis-aligned. A seminal advantage of Information Theoretic analysis of the JMS is its low sensitivity to the locations of the loci. This is because the measures are sensitive to the magnitude, number, and shape of the loci but not necessarily their location. Measures such as linear correlation become weaker or fail entirely in detecting co-alignment if the loci do not fall to some degree on a straight line.
Figure 2-3 Example of JMS structure with mis-alignment progressing to co-alignment. The latter has fewer but more intense loci. (a) is a unimodal case; and (b) is bi-modal.
Arrows in Figure 2-3 indicate the path of loci positions as co-alignment is approached during a search. The rates at which these loci move as a function of mis-alignment depend on the size, shape, and sharpness of the imaged objects. For objects that are at most one pixel large in terms of sharpness and extent, the convergence to the diagonal occurs mostly on a sub-pixel scale and less abruptly for larger or blurred objects. The ovals depict the extents of the probability distributions of each object in the JMS. In reality, the distributions can have almost arbitrary shape, extent, and are generally different for each object as measured by each sensor.

An extreme example is shown in Figure 2-4. A white noise image (Gaussian distributed independent and identically distributed (i.i.d.)) is co-registered with itself. Here, the implied objects are all unique due to the independence of the pixel values, and are the size of pixel footprints, one object per pixel. The co-alignment case is a narrow line on the 1:1 diagonal. In reality, for a multimodal imaging case, the difference is much more subtle. It differs from the extreme and cartoon examples in that the objects are more numerous, and often cover areas of many pixel footprints. Sensing of the objects produces more numerous and more varied textures. PSF effects prevent pure signatures from being depicted in the JMS. The result is a very complex JMS picture. In Figure 2-5, a real image pair is
Figure 2-4  **JMS alignment effect for a white noise image.** (a) scattergram for i.i.d. pixel image mis-aligned with itself by 3.5 pixels in x, y; (b) i.i.d. pixel image; and (c) scattergram for perfectly alignment. JMS origin is in the upper left corner. The image scales are pixel coordinates and the JMS scales are bin numbers.
represented in image space and in JMS for both mis- and co-aligned cases. Even though

![Figure 2-5 Uncorrelated, multi-spectral image pair.](image)

(a) TM5 band3, and (b) band4. (c) JMS scattergram with 2 pixel x, y mis-alignment, and (d) consolidation of joint density displayed as brighter localized areas. (c) and (d) are presented on the same tonal scale for amplitude comparison.
the differences between the scattergrams are subtle, some consolidation of joint density from lower diffuse values to higher, more compact values is evident.

Distinction of the two cases by a quantitative operator allows detection of the correct co-alignment. Entropy, as a measure of diversity, has a higher value for the mis-aligned case. Approaches using Fisher Information are developed in this dissertation that respond to both the shape of the loci, and non-linearly to their strength, producing higher values for fewer, but more granular loci.

An undesirable effect occurs when overlaps between the objects cause mixed signatures to merge into single loci by chance. Referring to Figure 2-3, it is evident that at some sub-pixel shifts, some loci will coincide with others. This is a reduction of number and strengthening of loci even though co-alignment has not occurred. The effect causes ‘false peak’ indications from the information theoretic measures.

By reducing the number of objects, the number of opportunities for this coincidence is also reduced. Computing the JMS values for a spatial subset of the images may accomplish this at the expense of smaller sample size and thus poorer estimation of the joint probability density function, followed by poorer performance of algorithms operating on the JMS due to higher stochastic variation. If prior information can be used to select image subsets that are more amenable to alignment estimation, then performance may be improved [55].
2.6 Probabilistic and Statistical Dependence

If two images of the same object(s) are to be acquired, does it imply a causal relationship between the data manifested as probabilistic dependence?

Probabilistic dependence of two random variables or vectors occurs when they completely predict each other, i.e., a realization of one reveals all information in the other. Independence means no information about one random variable’s random outcome is gained through realization of the other. In between these extremes, realization of one random variable reveals some information about the other; how much depends on how this partial information is measured. Dependence measures may be used to detect co-alignment if dependence gives a unique indication of the co-aligned state (for instance, it is maximized upon co-alignment). A more detailed explanation of how some IT measures, including some Fisher Information measures used in this dissertation, are dependence measures, is given in [68].

If the image forming processes are correlated, as would be the case for unimodal data from similar sensors, the data will be dependent. In this limited case, it is expected that maximal overlap of the objects, the sources of the correlation, corresponds to maximal dependence.
In the more general case, in which correlation may be absent, there may still be causal relationships between the objects and the data. It is commonly assumed implicitly or explicitly [69]-[76] that statistical dependence between images is incurred by the common causality from objects in both images, and therefore, best alignment is indicated by application of a dependence measure, such as mutual information in Eq. (2.9). To show that this intuition may be false, consider the example in Figure 2-6:

![Image chips A and B](image)

**Figure 2-6** Image chips that are statistically independent when co-aligned and become dependent upon mis-alignment.

Image chips A and B depict two identical size images of the same co-aligned objects. In image A, one object is sensed as random noise, while the other is sensed as a constant. In image B, the reverse is true. A and B are independent in the JMS because each pixel’s change in one image can not predict the change in the other. If one image is shifted with respect to the other by some distance less than the size of the objects, then there will be
locations where constant values overlap in both images making those locations statistically dependent. Thus dependence is minimized (zeroed) in the co-aligned case. The above is presented as a counterexample to the premise that maximal dependence corresponds to best co-alignment. There may be other combinations of adjacent textures not illustrated here that also have minimum dependence at co-alignment.

Real images may have any mixture of objects, including ones that, when adjacent, may either drive the global dependence to a maximum or minimum at co-alignment. If both are present, the contributions to global dependence conflict, undermining robustness to image content of dependence measures intended for estimating the co-alignment location that is revealed by maximal dependence.
CHAPTER 3  APPROACH

3.1 Approach

It is hypothesized in this dissertation, that Fisher Information may be an alternative, and may in some cases outperform the commonly used Shannon Entropy. It is also hypothesized that use of both measures simultaneously in the co-registration problem may contribute independent information to the alignment estimate, allowing for more accurate or more robust co-registrations. Testing of methods derived in this chapter is intended to support both of these hypotheses.

3.2 Theoretical Design

The following is the basis for implementing Fisher Information and Entropic information measures directly on the 2D JMS.
3.2.1 Fisher Information

Fisher Information repeated from Eq. (2.10)

\[ I_\theta(f(y|\theta)) = \sum_{\forall m} \int f(y|\theta) \left[ \frac{\partial \log f(y|\theta)}{\partial \theta_m} \right]^2 dy = -\sum_{\forall m} \int f(y|\theta) \frac{\partial^2 \log f(y|\theta)}{\partial \theta_m^2} dy \]  

(3.1)

has two forms that are theoretically identical within minor constraints. The second derivative form was not implemented here because estimation of second derivatives is more problematic than estimation of first derivatives as explained in Section 3.2.3.2.

Parameter \( \theta \) may be chosen to be an independent variable of the density. If the derivative is taken with respect to each of the JMS abscissas, using the additivity of Fisher Information with respect to the Fisher Parameter, the form in Eq. (3.2) is obtained.

\[
\iint f(x, y) \left( \frac{\partial}{\partial x} \log f(x, y) \right)^2 dx dy + \iint f(x, y) \left( \frac{\partial}{\partial y} \log f(x, y) \right)^2 dx dy = \iint f(x, y) (\nabla \log f(x, y))^2 dx dy
\]  

(3.2)
This is Fisher Information with the JMS gradient as the Fisher Parameter. A discrete implementation of Eq. (3.2) is given in Eq. (3.3). This variation on Fisher Information will be called the Gradient Form.

\[
\hat{I}_G(A, B) = \sum_{\forall(a, b)} \hat{p}_{A,B}(a, b) \left\{ (D_A^1(\log \hat{p}_{A, B}(a, b)))^2 + (D_B^1(\log \hat{p}_{A, B}(a, b)))^2 \right\} \tag{3.3}
\]

\(\hat{I}_G(A, B)\) is a gradient implementation (approximating Eq. (3.2)) for image pair \{A, B\}. \(\hat{p}_{A,B}(a, b)\) is a JMS data density function quantized at points \((a, b)\) and \(\hat{p}_{A, B}(a, b)\) is its Gaussian smoothed version. \(D^1_2\) is a directional first derivative approximator on discrete two dimensional data. There are two derivatives computed, one with respect to each of the JMS abscissas.

Instead of summing measure contributions in the JMS isotropically, the derivative values may be multiplied as an implementation of Eq. (2.13), as shown in Eq. (3.4).

\[
\hat{I}_P(A, B) = \sum_{\forall(a, b)} \hat{p}_{A,B}(a, b) \{ D_A^1(\log \hat{p}_{A, B}(a, b)) \cdot D_B^1(\log \hat{p}_{A, B}(a, b)) \} \tag{3.4}
\]
This is intended as a way of obtaining Fisher Information summand contributions from JMS granularity features that occur only at intersections of distribution features in both images. This is in contrast to the gradient measure that is sensitive to any JMS perturbation, including that caused by statistical features exclusively in one image.

Further specialization of the Fisher Information function is achieved by treating separately the positive and negative results of this derivative product. Four possibilities are tested:

1) Allowing both contributions with opposite signs (Eq. (3.4) unaltered). This will be referred to as the Product Form. In this case, the Fisher Information function may produce positive or negative results;

2) Using only positive contributions (Eq. (3.5)),

\[
\hat{I}_{P_+}(A, B) = \sum_{\forall(a, b)} \hat{p}_{A, B}(a, b) \max\{D_A^1(\log \hat{p}_{A, B}(a, b)) \cdot D_B^1(\log \hat{p}_{A, B}(a, b)), 0\}
\]

where the sign of the derivative is the same in both JMS directions. Mixed sign responses from the derivatives are discarded in the sum. This will be referred to as the Positive Product Form;
(3) Allow only negative (mixed sign) contributions (Eq. (3.6)).

\[ \hat{I}_{P}(A, B) = -\sum_{\forall (a, b)} \hat{p}_{A,B}(a, b) \min \{ [D_A^{1}(\log \hat{\rho}_{A,B}(a, b)) \cdot D_B^{1}(\log \hat{\rho}_{A,B}(a, b))], 0 \} \]

The result is negated to make it non-negative for graphing purposes. It is named the Negative Product Form;

(4) To accumulate all responses from the derivatives, an absolute value of the product is computed (Eq. (3.7)). This is named the Absolute Product Form.

\[ \hat{I}_{P}(A, B) = \sum_{\forall (a, b)} \hat{p}_{A,B}(a, b) |D_A^{1}\log \hat{\rho}_{A,B}(a, b) \cdot D_B^{1}\log \hat{\rho}_{A,B}(a, b)| \] (3.7)

Only the Gradient Form is a Fisher Information; the product forms are derived from the Fisher Information Matrix and are therefore ‘Fisher Information-inspired functions’.
3.2.2 Divergence Estimator Development

The measures described in the previous section respond to changes in the marginal densities of each image. In this section, the same measures are put in divergence forms, mainly to focus their response on structure in the JMS, while reducing the response to the marginal densities.

3.2.2.1 Entropic Divergence

The entropy-based divergence from Eq. (2.4) is written in estimator form in Eq. (3.8)

Marginal densities are computed from the joint density and cross-multiplied together in the denominator to produce the hypothetical reference joint density that would be obtained if the images were independent, with the same individual densities as the actual images.

To estimate the Entropic divergence $I_{DS}$, the formulation in Eq. (3.8) is applied.

$$
I_{DS}(A, B) = \sum_{a, b} \hat{p}_{A, B}(a, b) \log \left( \frac{\hat{p}_{A, B}(a, b)}{\hat{p}_{A, B}(a) \cdot \hat{p}_{B}(b)} \right) 
$$

(3.8)

Probability masses $\hat{p}$ are JMS densities - estimates of joint and marginal global densities of images A and B. There are no parameters to set, making application to the JMS straightforward.
3.2.2.2 Fisher Information in Divergence Form

The concept of estimating information from a comparison of densities is applicable to Fisher Information. The author has not found such a formula in imaging or remote sensing literature. In the Eq. (2.10) excerpt cited below as Eq. (3.9),

\[
I_0(f(y|\theta)) = \sum \int f(y|\theta) \left[ \frac{\partial \log f(y|\theta)}{\partial \theta_m} \right]^2 dy
\]  \hspace{1cm} (3.9)

we can replace the joint density with a joint density ratio as the argument of the log function and explicitly write it in terms of the image A and B joint density, yielding Eq. (3.10).

\[
I_0(f((A, B)|\theta)) = \sum \int \int f(a, b) \left[ \frac{\partial \log f(a, b)}{\partial \theta_m} \frac{f(a, b)}{f(a) \cdot f(b)} \right]^2 dadb
\]  \hspace{1cm} (3.10)

The expectation is with respect to the joint random variables A and B that take on values a and b. Any information measure of this form should be called Fisher Information inspired, as Eq. (3.10) has never been called Fisher Information, as far as the author knows. The author proposes that it be named Fisher Divergence. The choosing of the Fisher Parameter parallels the approach taken for the non-divergence Fisher Information
measures. The four Fisher Information forms presented as a discrete implementation in Equations (3.3) - (3.7) are converted to divergence forms in Equations (3.11) - (3.15) that are in discrete form. The name of each one is modified to call it a divergence. Each of them is originally derived and defined in this dissertation.

\[ \hat{I}_{DG}(A, B) = \]  \hspace{1cm} (3.11) 
\[ \sum_{\forall (a, b)} p_{A, B}(a, b) \left[ D_A \left( \log \frac{\hat{p}_{A, B}(a, b)}{\hat{p}_A(a) \cdot \hat{p}_B(b)} \right) \right]^2 + \left[ D_B \left( \log \frac{\hat{p}_{A, B}(a, b)}{\hat{p}_A(a) \cdot \hat{p}_B(b)} \right) \right]^2 \]

\[ \hat{I}_{DP}(A, B) = \]  \hspace{1cm} (3.12) 
\[ \sum_{\forall (a, b)} p_{A, B}(a, b) \left[ D_A \left( \log \frac{\hat{p}_{A, B}(a, b)}{\hat{p}_A(a) \cdot \hat{p}_B(b)} \right) \cdot D_B \left( \log \frac{\hat{p}_{A, B}(a, b)}{\hat{p}_A(a) \cdot \hat{p}_B(b)} \right) \right] \]

\[ \hat{I}_{DP^+}(A, B) = \]  \hspace{1cm} (3.13) 
\[ \sum_{\forall (a, b)} p_{A, B}(a, b) \max \left[ D_A \left( \log \frac{\hat{p}_{A, B}(a, b)}{\hat{p}_A(a) \cdot \hat{p}_B(b)} \right) \cdot D_B \left( \log \frac{\hat{p}_{A, B}(a, b)}{\hat{p}_A(a) \cdot \hat{p}_B(b)} \right) \right], 0 \]

\[ \hat{I}_{DP^-}(A, B) = \]  \hspace{1cm} (3.14) 
\[ - \sum_{\forall (a, b)} \hat{p}_{A, B}(a, b) \min \left[ D_A \left( \log \frac{\hat{p}_{A, B}(a, b)}{\hat{p}_A(a) \cdot \hat{p}_B(b)} \right) \cdot D_B \left( \log \frac{\hat{p}_{A, B}(a, b)}{\hat{p}_A(a) \cdot \hat{p}_B(b)} \right) \right], 0 \]
3.2.3 Supporting Functions

3.2.3.1 Density Estimation

To obtain an estimate of the joint probability density function for the two image imaging process, the JMS density is measured. To represent it quantized according to the JMS independent variables, histogram binning and counting are done in two dimensions. To prevent possible clipping, the range of bin values should extend beyond the data range, forcing end-range bins to receive no counts. Both Entropy and Fisher Information involve computing the logarithm of the JMS density, but some histogram bins may be empty with a count of zero. This may be due to discrete data values that have a different support than the bin ranges, there may not be enough pixels in the images to fill all bins, or the imager response to a given scene may not include all possible bin values. An additive offset value is added to all JMS values to mitigate the log(0) singularity, and exaggerated sensitivity at JMS values of one or two. Further details of this parameter are given in section 4.3.3.1.

The proper offset mitigates the aforementioned numerical problems, but does not appreciably impact the Fisher Information responses to JMS locations where the majority of the probability is located.

\[
\hat{I}_{D|P}(A, B) = -\sum_{\forall(a,b)} \hat{p}_{A, B}(a, b) \text{abs} \left[ D_A \left( \log \frac{\hat{p}_{A, B}(a, b)}{\hat{p}_A(a)} \right) \cdot D_B \left( \log \frac{\hat{p}_{A, B}(a, b)}{\hat{p}_B(b)} \right) \right]
\]
This research is intended to evaluate measures that are sensitive only to the finer details in marginal and joint densities and not their mean and variance. The data scale is normalized to zero mean and unit variance to prevent these two parameters from affecting the results. An added benefit of the normalization is little need to adjust the bin ranges. If they are set to accommodate data with heavy-tailed densities (density functions with tails that diminish less steeply compared to a Gaussian density), then they will work for all data. Although the bins can be set individually to any step size, to make other computations that are sensitive to scale in the JMS simpler to implement, the bin centers are uniformly spaced, making bin width equal.

3.2.3.2 Derivatives

The Fisher Information formula allows obtaining the same result with either first or second derivatives of the measured density function. Density data in this project is estimated in discrete form, therefore its derivatives are to be estimated with a discrete approximation of continuous derivatives. The first order difference is a good approximation to a continuous first derivative at low frequencies, and it deteriorates with increasing frequency. The ratio of the actual response in the frequency domain to the ideal differentiator response is a Sine function with first zeros at +/- the sampling frequency. Such a comparison is valid only over a frequency range of +/- the Nyquist Frequency.
An approximation to the second derivative may be computed by two applications of first order differences sequentially. The approximation is very poor and is described by a Sinc-squared ratio of actual to desired responses, with the first zeros at +/- the Nyquist Frequency. [77] Better approximations to the second derivative may be obtained with finite impulse response (FIR) filters having a length of many tens of points. This is undesirable, as it would greatly slow down derivative computations. Drawbacks of both approaches are avoided by utilizing only the first derivative form of Fisher Information.

Two different discrete first derivative approximations were implemented in order to realize Fisher Information divergences defined in Section 3.2.2.2: (1) First order difference, accomplished by convolution with the horizontal and vertical kernels show in Eq. (3.16).

\[
\begin{bmatrix}
  1 & -1 \\
\end{bmatrix}
\begin{bmatrix}
  1 \\
-1
\end{bmatrix}
\] (3.16)
This function has a relatively high magnitude response bandwidth shown, in Figure 3-1 (a). It responds proportionately to frequency at low frequencies, then a partial attenuation occurs near the Nyquist frequency of 0.5 cycles per pixel. This can be compared to an ideal differentiator (Figure 3-1 (b)) that has a response proportional to frequency at all frequencies up to Nyquist.

**Figure 3-1 First order difference frequency response.** (a) transfer function magnitude of first order difference compared to (b) that of an ideal differentiator. Amplitude scale is relative to maximal response for each plot. Frequency scale is in cycles per pixel from zero to the sampling frequency.
A major drawback in using the first order difference is its even-size kernel. For the product of derivatives Fisher Information forms that require orthogonal directional derivative estimates, it is not possible to obtain slope results for the same point in the JMS (Figure 3-2); (2) The central first order difference $[1, 0, -1]$ does not possess the same problem - an odd size kernel allows both horizontal and vertical directional derivatives to be located at precisely the same location. This kernel inherently has a more restricted bandwidth response compared to the first order difference. The ratio of a true derivative to the first order central difference is the Sinc function, with the first zeros at +/- Nyquist frequency [77]. This lower bandwidth is of less consequence if the JMS is relatively diffuse, and its derivative is to be intentionally blurred as well. The actual kernels used are presented in Eq. (3.17).

$$
\begin{bmatrix}
0.25 & 0.5 & 0 & -0.5 & -0.25 \\
0.5 & 1.0 & 0 & -1.0 & -0.5 \\
0.25 & 0.5 & 0 & -0.5 & -0.25
\end{bmatrix}
\begin{bmatrix}
0.25 & 0.5 & 0.25 \\
0.5 & 1.0 & 0.5 \\
0 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
0.25 & -0.5 & -0.25 \\
-0.5 & -1.0 & -0.5 \\
-0.25 & -0.5 & -0.25
\end{bmatrix}
$$

Figure 3-2 Spatial location of first order differences with respect to the data point locations from which they were computed. Squares are geometric pixel areas; dots are locations of pixel values, and circles are locations of the computed
This is similar to a central first order difference, but modified to additionally restrict the differentiation bandwidth, and with triangular averaging included perpendicular to the differentiation direction. The frequency response for this kernel in the differentiation direction is shown in Figure 3-3. The response is nearly proportional to frequency up to about 0.1 of the sampling frequency, then peaks at around 0.17. The main difference seen when comparing to the response of the first order difference is that beyond the frequency of 0.17, the response diminishes and reaches zero at the Nyquist frequency instead of increasing monotonically.

Implementation and control of the derivative bandwidths is explained further in Chapter 4.

---

**Figure 3-3** Modified central first order difference: Magnitude of frequency response. in the differentiation direction. Amplitude scale is relative to maximal response. Frequency scale is in samples per pixel from zero to sampling frequency.
3.2.3.3 JMS Smoothing

Smoothing was applied to the JMS space only for inputs to the derivative operators, and was not applied to the density weighting in the expectation. It was accomplished by convolution with a bivariate, circularly symmetric Gaussian image, truncated to a square kernel of size 4.5 standard deviations, rounded up to an integer. The smoothing standard deviation was controlled by a $\sigma$ (‘Sigma’) parameter in the code.

3.2.3.4 Density Preservation and Restoration

In a spatial search for best image co-alignment, it is not possible to spatially transform one or both images and maintain their statistics. This is due to two factors: (1) on a pixel grid, resampling is required to estimate pixel values between the pixel center positions; (2) when an image is undergoing variable spatial transformation, such as during a spatial search to align it with another image, its content around its edges varies because different parts of the scene are included or excluded. With varying content, the global density of that image may vary. In each case, the statistics of the transformed image are not held constant, leading to variation in its density that will usually effect the joint density, which in turn will affect most measures based upon it. The impact manifests itself as an artifact pattern upon the search space that is dependent on the resampling rather than on the image
alignment, modifying the search results. These artifacts may alter locations of search extrema, and in severe cases, will result in false identification of correct alignment location.

The divergence measures described in Section 3.2.2 above, those that are divergences of the joint density relative to a hypothetical independent image density, have substantial built in immunity to changes in the marginal densities. This is because both the real and hypothetical joint densities change together during the spatial transformation. They are designed to null IT responses to marginal density change when the dependence of the two images is near non-existent - the worst scenario (as happens when signal to noise ratio is approaching zero). The operand of Eq. (3.10) repeated below

\[
\frac{p(a, b)}{p(a) \cdot p(b)}
\]  \hspace{1cm} (3.18)

has an equal numerator and denominator when the two images are independent. At higher dependence levels (associated with high signal to noise ratio), the denominator in Eq. (3.18) excessively compensates for marginal density changes in the joint density, but the alignment response of the IT measures is stronger than artifacts, making them less susceptible to artifacts encoded in a marginal density. The divergence methods are therefore suitable for both high and low noise levels.
Another way to prevent impact of varying marginal densities is to estimate the density of an image prior to transformation, apply the spatial transformation, and then correct its density to the estimated original. If this could be accomplished without altering spatial information, the problem would be entirely solved. However, due to the changes in image content because of spatial transformation, and the necessity for some randomization of values during the correction, some noise is introduced into the image and into the marginal and joint statistics, as explained in Appendix D.

The histogram matching algorithms that are commonly implemented in image processing software (such as ENVI and TclSadie) are inadequate for the aforementioned corrections because they adjust the density by varying the amplitude spacing of pixel values. Although the results often appear successful visually, IT measures, especially Fisher Information, readily react to the changes in the amplitude spacing leading to random errors due to JMS artifacts. Such histogram matching is not an alternative to the Appendix D method and is not used in this work.

Neither the internally correcting divergence measures, nor the density restorations completely eliminate the artifacts of resampling in the search space. For interpolation-based resampling, as is used in this dissertation, the resampling artifacts impact the alignment estimate the most when the sampling positions are near or on-pixel, even though on pixel resampling actually produces the least information measure error with respect to the un-resampled data. At on-pixel positions the resampled data is the original pixel values, but
slightly off-pixel resampling results in blending of several pixels and alters the resampled statistics. IT measures, being very sensitive to the statistics, change abruptly near the on-pixel sampling positions, and gradually over the span of on-pixel positions. As a function of shift, the effect is a spike at the on-pixel position, and a round bottom in between. To minimize sensitivity of the results to resampling position around the origin of the search range, both input images are resampled at the 2D mid-pixel position to obtain zero relative translation while simultaneously achieving minimized sensitivity to the artifacts. Relative translations other than zero are achieved through differential offsets from the mid-pixel sampling (Figure 3-4). This artifact avoidance scheme allows low artifact IT results over most of a ±1 pixel search in x and y, making it effective only for sub-pixel co-alignment.
3.2.4 Application - Co-registration Algorithm

3.2.4.1 Methodology

To apply the divergence measures designed above to the co-registration problem, two images of the same scene, that have a maximal mis-alignment of a few pixels anywhere in the image frame, are transformed relative to each other with a corrective transformation. This is carried out repeatedly over a search range of transformation parameters that allows finding the co-alignment. The information theoretic measures are used to indicate their
extreme values when the best alignment has occurred. To apply the alignment estimate to
the co-registration problem, the same corrective transformation and interpolation should
be applied as the ones used for estimating the alignment. An exception is use of alignment
estimates using a simple spatial transformation in small image subsets as virtual ground
control points for determining a more complex corrective transformation. The corrective
transformation and interpolation should be chosen prior to estimating the parameters.

The spatial transformation used for most of this work is Matlab’s two dimensional interpo-
lation function ‘interp2’ that utilizes look up tables. It is given sampling coordinates gen-
erated by ‘meshgrid’ according to a desired translation. This method allows only
translations. To test rotating, scaling, and skewing transformations, the ‘imtransform’
function was used. The chosen resampling method is ‘*linear’ which invokes a bilinear
interpolation. Nearest neighbor resampling can not be used as it does not estimate sub-
pixel information, and bi-cubic or cubic spline interpolation provides no advantage over
bilinear, i.e., the difference is not identifiable in the results.

Images are input in a pair, but may be from arbitrary sources and may have disparate spa-
tial characteristics. If the spatial representation assumption of Section 2.1 is to be met, a
restoration-like process of equalizing the spatial responses that create the two images may
be applied to one or both images.
To allow for spatial searches, and other spatial processing of the images, without problems from border effects, only a central portion of the images is processed by the information theoretic estimators.

3.2.4.2 Evaluation

For real imagery, it is usually not possible to compare estimated alignment with true co-alignment because the latter is not known. By visual evaluation of the Fisher Information and Entropy plots, one can judge the potential precision of locating the extreme value, but overall accuracy cannot be determined.

Synthetic imagery has known alignment, thus alleviating the lack of truth problem. To gather statistics on co-registration performance of an algorithm, many synthesized images were generated and put through the algorithm. There are instances where the algorithm failed entirely for some particular images, algorithm parameters, or image generation parameters, and such failures are tallied.

Synthetic imagery also has the advantages of parameter control, that allows testing algorithms under differing conditions for different runs while holding difficulty levels constant for each run. The disadvantages are a lack of realism in the objects depicted. Details on image synthesis, its methodology, and controls used in this dissertation are described in Appendix A.
3.2.5 Experimental Considerations

3.2.5.1 Choice of Method and Parameters

Two factors distinguish the methods described previously in Sections 3.2.1 through 3.2.3: (1) Choice of Fisher Information estimator form \{Univariate; Divergence\}; and (2) Choice of means to combine the orthogonal components derived from the 2D JMS \{Gradient; Product\}. This is achieved by choice from five methods previously named Fisher Gradient; Fisher Product; Fisher Positive Product; Fisher Negative Product; and Fisher Absolute Value Product. The Fisher Gradient methods may be implemented with either the first or second order difference, but more derivative bandwidth control is available with the first order difference. All of the product methods should be implemented with the second order difference, hence only five methods are tested.

Further adjustments are made by varying \(\sigma\) ‘\(\text{Sigma}\)’ of the JMS smoothing function applied prior to the derivative estimator. All the IT estimators are profoundly non-linear, and therefore cannot be described by a convolution. Nevertheless, the smoothing of the JMS input to the derivative operators tends to broaden and smooth Fisher Information features as depicted in a Fisher Information surface over a search space. The result is somewhat visually similar to taking a rough IT search surface and smoothing it after estimation, even though the smoothing and IT measure computations are not interchangeable. Such smoothing of the IT search surface would reduce the amplitude of spurious extrema by
blurring or distributing their values to neighboring search locations. Instead, the blurring of the JMS prior to derivative computation makes some unwanted responses vanish. The effect of the JMS blurring is a much more effective and powerful reduction of some noise responses than linear filtering of the IT search surface.

Adjustment of histogram bin size affects the results as a resolution setting, and scales the blurring function described immediately above. The range of bin values is assumed and set in this work to be beyond the range of non-zero JMS values.

The additive offset that is adjusted so that the boundary between zero JMS values outside the data range, and the bins that have small counts, does not significantly impact the Fisher Information and Entropy computations, but makes the contributions come mainly from the region of more populous JMS values - those around the inflection points of the JMS density.

A way to predict the best method and its parameters is not available at this time, and is therefore left to experimentation. Results from several methods developed in Chapter 4 of this work can be compared to determine which ones are correctly indicating a co-alignment, and which are not, based on results presented in Chapter 5.
3.2.5.2 Search Strategy

The translation search range and step size parameters must be set such that several scale issues are satisfied. To encompass the extremum, shifts must cover the mis-registration uncertainty range. Unfortunately, the number of computations increases proportionately with the number of search points. To locate the actual peak or dip, it is necessary to sample the Fisher Information surface with sufficiently fine shifts. The computation load increases with the square of the inverse of the step size, such that optimally meeting the range and resolution needs may be infeasible in one run. Of course, experimentation to optimize these search parameters incurs additional runs.

Continuity allows a successive search to be performed that follows trends toward an Fisher Information surface peak, or finds an approximate location of the peak, followed by centering and recalculation at higher resolution. The sensor’s overall point spread function imparts continuity to adjacent pixel values. Furthermore, if the resampling function used to shift an image in sub-pixel increments is continuous, then the Fisher Information surface should exhibit some continuity as well. An artificial blurring of the images prior to JMS density computation may also be used to extend continuity, allowing multi resolution searches [78]. Bandwidth limitation of the derivative may be a superior to image blurring. Setting the ranges, step sizes, and derivative bandwidths at each stage requires some experimentation.
CHAPTER 4 DESCRIPTION OF RESEARCH

Actual implementation of algorithms was done in ten major revisions of code. The details of each are not germane to the research outcome of this dissertation, and are not presented. However, a major revision chronology may be helpful in understanding the history of the results by describing three phases of algorithm development. **Phase 1**: Initial trials of Fisher Information. Here, the most basic application of Fisher Information to the alignment estimation problem was tried. Eq. (3.3) was the only Fisher Information form used. The outcomes indicated that Fisher Information does indicate location of co-alignment. All interpolation methods used to spatially transform one of the two images over a search pattern imposed profound artifacts on the results. **Phase 2**: Product forms of Fisher Information in Eqs. (3.4) - (3.7) were introduced. Interpolation artifacts were mitigated by amplitude density correction as described in Appendix D. Artifacts disappeared, but performance at high image noise levels worsened. **Phase 3**: Divergence forms of Fisher Information in Eqs. (3.11)-(3.15) were implemented. Performance substantially improved for all imagery, and interpolation artifact mitigation was discontinued because the divergence measures are less susceptible to the interpolation effects. A summary of experimental attributes in each phase is in Table 4-1

The numerical formulas were implemented in straightforward computations. The purpose of the numerical processing is validation of the proposed approach, therefore there was no effort to implement numerical trickery for computational optimizations.
<table>
<thead>
<tr>
<th>Attribute</th>
<th>Phase 1</th>
<th>Phase 2</th>
<th>Phase 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fisher Information-Gradient Form</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Fisher Information-Product Forms</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Fisher Divergence-Gradient Form</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Fisher Divergence-Product Forms</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Entropy</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Entropic Divergence</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Correlation Measures</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Full affine spatial transformation</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Artifact mitigation (density matching)</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Artifact mitigation (differential translation)</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Derivative bandwidth control</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>External/Real Imagery</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>External/Real degraded imagery (Monte Carlo)</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Synthetic imagery (Monte Carlo)</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Error magnitude density estimation</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>

*Table 4-1 Summary of experimental attributes in each phase.*
4.1 Algorithm Implementation

Matlab (Versions R12, R13, and R14) was chosen as the programming and data handling environment, due to its programming ease, realization capability of any forseeable mathematical formulae, libraries of numerous common functions, graphic visualization tools, and well structured data storage environment.

The first tests of Phase 1 were intended to verify that Fisher Information may be a substitute for Entropy in the co-registration application.

4.1.1 Image Generation/Preparation

Test images were externally generated, either subset from real images, or generated by manual calculations. To test Fisher Information as an indicator of co-alignment (Phase 1), code was written that would allow control of translation, rotation, and scaling so that Fisher Information values could be evaluated versus different transformations about a known correct alignment. The transformation was based upon Matlab’s ‘imtransform’ function. (General purpose image transformation utility.) The resampling method was chosen to be bilinear. Only one image would be transformed, as the relative transformation between the two images was to be evaluated. To test the effect of image sharpness and imager noise on the results, each image could be independently blurred with a truncated Gaussian convolutional filter, and white noise could be added independently. In
Phase 1, general affine transformations were implemented with the ‘imtransform’ function. In Phase 2 and Phase 3, the extremely slow ‘imtransform’ function was replaced by the faster ‘interp2’ function, and only translations were implemented. It is particularly fast for resampling to a uniform grid when using the ‘*linear’ interpolation method. The decision to no longer test alignment estimators with other transformations within the capability of an affine transform was made after noting that the response of both Fisher Information and Entropy to different kinds of spatial transformation was similar - an extremum would indicate the correct alignment.

To minimize the effects of resampling upon the marginal densities of each image, the translation code allowed independent shifting of each image in Phase 3. Artifacts are minimized by shifting both images away from inter-pixel grid positions, symmetrically in opposite directions.

Experiments with interpolation showed that responses to transformation were almost indistinguishable whether bilinear or bicubic interpolation was applied. For sake of efficiency, bilinear interpolation was hard coded into the translation algorithm.
Estimation of the JMS probability density was accomplished by two dimensional histogram binning. Fisher Information and Entropy are both sensitive to the image variance (or power) which is not of interest in the co-alignment problem. A normalization to zero mean and unit variance removed the effects of those parameters in each image prior to binning.

4.1.2 Fisher Information Estimation

The Fisher Information estimator initially tested in Phase 1 was a discrete approximation to the gradient Fisher Parameter realization, utilizing both the discrete estimate of the JMS probability density and the discrete approximation of the derivative.

In Phase 2, the Fisher Information was in gradient and product form. The only desired response of the IT measures should come from relations of the two images rather than from changing statistics of either one alone. Therefore marginal density correction was implemented to make marginal densities unchanging under interpolation.

In Phase 3, all the IT measures were divergences. The density correction procedures were discontinued due to the good correction for marginal density variations implemented within the divergences.
4.1.3 Comparison: IT Estimators

For comparison to Fisher Information, Entropy, Correlation Coefficient, and the Phase Correlation zero-shift value were also computed. Each of these has a straightforward implementation on discrete data, because the formulas are already in discrete form. Normally, the entire correlation function (function of x, y lag) would be computed at once using the original un-translated images, but this would not allow comparison against the IT measures because the data would not be processed through the interpolation function. To process exactly the same transformed and interpolated images through the correlation methods, only the zero shift output points are extracted from the cross-correlation functions, and these measures are repeatedly computed during the spatial search.

4.1.4 Spatial Transformation Looping

It is necessary to evaluate co-alignment indicators over a range of mis-alignments that includes co-alignment. A loop that invokes Fisher Information and other indicators on a pair of images undergoing a relative spatial transformation allows plotting the response to the transformation to see if co-alignment is correctly indicated. Initial versions (Phase 1) allowed only one spatial transform parameter to be sequenced, but later versions allowed two. This made possible not only shifting in both image coordinates, but allowed combining of various spatial transformations, for example simultaneous rotation and scaling.
4.1.5 Density Correction for Resampling (Applied Only in Phase 2)

Ideally, the goal is to spatially adjust one image with respect to another until the two match. It is not desirable to alter the image(s) in any other way. An issue with resampling is alteration of the radiometric statistics of an image due to new sample values being numerical combinations of original values.

A way to mitigate these artifacts of the spatial transformation process is to adjust the image values to match the statistics prior to transformation. An algorithm was implemented to estimate the individual densities of each image (marginal densities in the JMS) and then perform a histogram matching after spatial transformation to pre-transformation values. This is described with more detail in Appendix D.

Divergence forms of the information measures applied in Phase 3 of this dissertation innately correct for variation in the marginal densities, making the above correction less needed.

4.1.6 Detection of Alignment

In this dissertation, the location of best alignment, in terms of the transformation required to obtain it, is declared by position of an extremum an IT measure, within confines of the search space. All of the desired responses are summarized in Table 4-2. Fisher Informa-
tion in gradient form responds only with a peak (positive extremum), therefore, the maximum value is used to detect the location. Entropy-based mutual information also produces a peak, however Entropy responds with a dip. The product form of Fisher Information may respond in either positive or negative directions, therefore the extremum with the largest magnitude is used to determine location. Modifications of the product form are again unipolar. They either discard one polarity or the other, or use an absolute value operation, preventing accumulation of values with an opposite sign. In all aforementioned cases, only a peak is detected and located. A seventh estimate is obtained by computing the average location between the Fisher Information absolute value result and that of Entropy. For the Correlation methods, a positive peak is the only expected indicator of co-alignment.

4.1.7 Analysis of Results

4.1.7.1 Visual Analysis

For co-registration runs on real data, it is not possible to determine the actual alignment estimation error because true alignment is not known. Real imagery resampled to a lower resolution is also used because the unknown registration errors are reduced relative to the output sample interval by the downsampling. A surface plot display for all six IT estimators allows qualitative assessment of whether the attempts succeeded, and how easy it would be to precisely locate the extremum’s position. Such assessments are made accord-
ing to how clean and sharp the peak or dip appears, and whether other IT surface features compete with or interfere with what is believed to be the correct location. A qualitative comparison may also be made between the IT methods and the correlation methods.

All of the IT surface plots are on a scale that varies from case to case, and should be used for relative comparison within each individual graph. The negative-part-of-product graph and Entropy are inverted to help visualize a negative extremum as a peak. The Phase Correlation plot is also on its own arbitrary scale. The Correlation Coefficient is plotted on a unitless scale in the -1.0 to +1.0 range.

<table>
<thead>
<tr>
<th>Method</th>
<th>Desired Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fisher Gradient</td>
<td>Positive peak</td>
</tr>
<tr>
<td>Fisher Product</td>
<td>Extremum of largest magnitude</td>
</tr>
<tr>
<td>Fisher Positive Product</td>
<td>Positive peak</td>
</tr>
<tr>
<td>Fisher Negative Product (negated)</td>
<td>Positive peak</td>
</tr>
<tr>
<td>Fisher Absolute Value</td>
<td>Positive peak</td>
</tr>
<tr>
<td>Entropy (negated)</td>
<td>Positive peak</td>
</tr>
<tr>
<td>Correlation Coefficient</td>
<td>Positive peak</td>
</tr>
<tr>
<td>Phase Correlation</td>
<td>Positive peak</td>
</tr>
</tbody>
</table>

*Table 4-2 Responses of IT measures to co-alignment.* Response directions are the same for IT measures and divergences derived from them.
4.1.7.2 Statistical Analysis

For synthetic images created with a known alignment, precise statistics may be computed. Large numbers of synthetic images are repeatedly run, and RMS error statistics on error radius, and counts of how many times a particular method succeeded out of the total are tallied. Success is defined by results that are within an error magnitude bound. For all six methods, and for a seventh method, a mean peak position from Fisher Information absolute value and Entropy, the error vector magnitude density and cumulative distribution are estimated. (The reason why the Fisher Information absolute value algorithm is chosen for this averaging is that it appears to succeed most often of all the Fisher Information methods.) The algorithm that obtains an anti-aliased density estimate, described in part of Appendix D, the same that is used in marginal density restoration to mitigate artifacts, is tasked to allow comparison of error distance distribution over the entire range of possible errors in the search space.
4.2 Test Methods

4.2.1 Real Images

Synthetic imagery is potentially ideal for ascertaining performance of algorithms because it allows full knowledge of truth, and does not limit sample size. Unfortunately, many properties of real images, for which this dissertation is intended, are not realizable by a practical synthesis algorithm. The most confounding aspect of real images is the complexity of objects in a real terrestrial scene. Classification of them is partially attainable, but synthesis is not.

To verify alignment estimation properties of IT methods, a selection of remotely sensed image pairs is tested to qualitatively assess the algorithms. An overview flow diagram is shown in Figure 4-1. Absolute errors can not be measured on these data due to lack of co-alignment truth. The chosen image pairs include some that are visually dissimilar due to disparity in sensing physics, independent noise between them, or changes over time. Included are image pairs that are statistically uncorrelated in the region under evaluation.

All real images have been previously registered by some other means with unknown accuracy, and are believed to be aligned with less than one pixel error. The presence of a distinct peak near the zero translation point indicates a high likelihood that this peak marks a valid co-alignment estimate made by a method under test.
Figure 4-1 Process flow for real images in Phase 1 and Phase 3. Processing of a
4.2.2 Degraded Real Images (Monte Carlo)

A compromise method of obtaining image pairs with nearly co-aligned real objects is to start with a high resolution real image pair that is aligned to better than one pixel accuracy.
at its native resolution. Such images are then downsampled to further improve the alignment relative to the lower output resolution. By blurring the higher resolution image prior to down sampling, then using aggregate average interpolation, and finally adding noise to the result, an emulation of a noisy low resolution sensor is accomplished. Details are presented in Appendix B. These images are repeatedly generated using random sampling phase and noise samples for Monte Carlo simulation. The process is summarized in Figure 4-2 flow diagram.

4.2.3 Synthetic Images (Monte Carlo)

To evaluate the usefulness of the measures and compare their performance, a statistical analysis of their absolute precision and accuracy was performed. Batches of synthetic images and degraded real images were run with all random values independently generated for each iteration. Figure 4-3 shows Phase 2 process flow, and Figure 4-4 shows Phase 3 process flow.

4.2.4 Statistical Analysis for Monte Carlo Runs

For each co-registration trial, and for each IT measure, the IT estimator extremum location is detected and logged. As the correct co-alignment is known to be at zero translation, the mean error euclidean distance (or error vector magnitude) in the x, y plane are computed.
Co-alignment indications beyond a set distance from zero shift are marked as failed. This threshold is set as a two dimensional box within the search area of half the search range for each side.

Figure 4-3  Synthetic image flow diagram for Phase 2. The search loop runs inside the Monte Carlo loop. The synthesis process is diagrammed in Appendix A, and the density restoration elements are diagrammed in Appendix D.
Using the anti-aliased density estimation methodology described in Appendix D, the density of error magnitude is plotted for all five Fisher Information measures, for Entropy, for the mean result from Entropy and Absolute Product methods, for the Correlation Coefficient, and for Phase Correlation. The cumulative distribution is similarly plotted for all measures.

**Figure 4-4 Synthetic image flow diagram for Phase 3.** The synthesis process is described in Appendix A., and the statistical analysis is explained in section 4.2.4, with further explanation of error density estimation in Appendix D.
The display exhibits the smoothing and slight random deviations imparted by anti-aliasing that is accomplished by addition of small random values to the data and by digital filtering. This randomization causes slight variations from true values that randomly vary each time the estimator is applied. The smoothing broadens the estimate range such that some density at negative magnitudes may be reported in a ‘tail’ from large real values at positive magnitudes.

4.3 Control Parameters

Listed here are the main control parameters used to define the experiments. Theoretical explanations of how these parameters affect the algorithms are given in other parts of this thesis, such as the theoretical derivations in the prior chapters or in the appendixes. Actual setting of these parameters requires a judgement call based on trial runs.

4.3.1 Fisher Information Estimation

4.3.1.1 Choice of Method

In this work, all methods are computed for each experimental run.
4.3.1.2 Derivative Bandwidth (Sigma)

In approximating the derivative, the bandwidth is restricted to less than the intrinsic response of the derivative approximator through convolution of JMS values with a 2D circularly symmetric Gaussian kernel. The width of this kernel is specified in units of JMS bin intervals referenced to the Gaussian function standard deviation, and is named Sigma. The same Sigma value is applied consistently in the Fisher Information estimator forms and variations. Typical values ranged from <0.1 (effectively zero) for sharp, clean images, to about 2 for very noisy images. It does not matter if the noise is in the images themselves, or simulated. Experimentation with Sigma should begin with low values. When a point is reached such that the apparent precision is no longer improving, Sigma should not be increased further, as accuracy may degrade.

4.3.2 Entropy Estimation

There are no controls or parameters for Entropy estimation in this implementation.
4.3.3 JMS Density Estimation

4.3.3.1 Binsize, Range

Specification of the JMS bin partitioning is in standard deviation units due to the normalization that takes place prior to the estimation. The bin range is set such that there is little or no truncation of distribution tails. For image data measured in power units, such as radiance, the left tail is inherently truncated at zero on the original scale, and is thus limited in extent after normalization. Some data, such as from Synthetic Aperture Radar (SAR) has a very heavy positive tail requiring a large bin range even after normalization. For random synthetic data, both tails need to be accommodated. The Binsize parameter controls how fine details are represented in the JMS estimate. Larger bin sizes produce less sharp responses to co-alignment, but also help reduce representation of random noise in the JMS due to an on-average larger sample size (number of image pixels) per bin. If the Sigma parameter is set to a large value, then there is already smoothing of the JMS and small bin sizes are not required. Bin range values were set to -6 to +7 standard deviations about zero in this work. Bin sizes were set at 0.05 to 1.0 standard deviations.

An offset value is added to each bin count to prevent division by zero or logarithm of zero singularities in code that utilizes the JMS estimate. This parameter is fixed at 1/number-of-pixels.
4.3.3.2 Crop Size

This determines the square window size used to sample the images for JMS estimation. A single number specifies the size of each side of the square in pixel interval units. The input images must be larger to allow room for the search ranges, and any other processing to take place, such as convolutional filtering, so that border effects will be entirely outside this crop window. Best performance has been observed with window sizes from 64x64 to 192x192 pixels. For smaller windows, the estimates of the joint density have a higher variance such that co-registration precision and success rate decrease. In a too-small window, the diversity of objects may inadequately represent the scene. Joint density estimation in larger windows may be too time consuming.

4.3.3.3 Initial Offset \((x, y)\)

The central point of the search range may be specified in order to center it about approximately known mis-alignments. If the initial mis-alignment is unknown, this parameter should be set to \([0, 0]\).
4.3.3.4 Search Range (x, y, shift, scale, rotation)

In versions of the code that allow affine spatial transformations (Phase 1), two parameter lists are specified that provide pairs of values that are any two variables in an expression realizable by the transformation. The program then performs the specified spatial transformations, spanning all combinations of parameters on the lists and finds IT measure values for those points. Surface plotting of the IT surfaces is then done against the x and y plot axes at values given in the two lists. Later versions of the code (Phases 2 and 3) allowed only translation in x and y to be specified. Two automatically generated lists of x and y shifts are generated from +/- a single range value, and subdivided by step size. Computation time is roughly proportional to the total number of points - the square of the range.

The recommended range must be greater than the initial mis-alignment plus the breadth of the response to co-alignment. Experimentation is needed to determine the correct range. Initial estimates may be applied as an Initial Offset and used to refine the search into a smaller area in which computation can take place in smaller steps. For real imagery already coarsely co-registered to within one pixel spacing, the search range was +/- 0.5 to 3.0 pixels.
4.3.3.5 Step Size

Step Size specifies the x and y shift increment between adjacent search points. The computation time for one image pair is roughly proportional to the square of the inverse of this parameter. Experimentation is required to evaluate how fine a grid is necessary to correctly locate the extremum and/or to display a visually continuous Fisher Information surface graph. Typical values range from 1 pixel for blurry, noisy data to 0.05 pixels for sharp, clean images. Surface graphs with around 30x30 points were found to be optimal for graphical presentation using Matlab’s ‘meshgrid’ function.

4.3.4 Degraded Real Image Generation

Details of the methodology and details of application of these control parameters are presented in Appendix B.

4.3.4.1 Output Size

Size in pixels (row, column) of the resulting images to be processed. See Section 4.3.3.2 for the alignment estimator requirements.
4.3.4.2 Downsampling Factor

The scaling factor by which the images will be downsampled. This is an odd integer, usually 5 or greater.

4.3.4.3 Window Center

Window Center coordinates of the crop window that is applied to the source images. These coordinates must be chosen such that the crop window will be entirely within the source image frames. The actual size of the crop window applied to the input image is the downsampling factor times the output size. The actual window center is randomly dithered +/- 0.5 * Downsampling Factor pixels for one input image and +/- 0.5 or 1.0 * Downsampling Factor pixels for the other.

4.3.4.4 Fixed/Random Relative Sampling Phase

The sampling phase of one image is uniform random on the range of up to +/- half of an output pixel. This parameter determines if the second image is sampled in the same phase, or if its phase with respect to the first image is uniform random on the range of up to +/- half of an output pixel.
4.3.4.5 Gaussian Blurring Standard Deviation

Gaussian Blurring Standard Deviation sets the isotropic Gaussian blurring kernel one standard deviation width in units of output pixel spacing. Blurring is implemented by convolution with a Gaussian function truncated to a square kernel with each side at Gaussian Blurring Standard Deviation * 4.5, rounded up to the nearest integer. The value used in this work was 0.5 chosen to mimic a hypothetical sensor that has a Gaussian PSF with a standard deviation width of one half of a pixel.

4.3.4.6 SNR

SNR is the ratio of standard deviations of downsampled image to that of i.i.d. noise to be added to the downsampled result. This assumes that all variance of the downsampled image is signal, regardless of how noisy the original input images are. Any non-negative value may be set. Values used were in the infinity to 0.1 range.

4.3.5 Synthetic Image Generation

Details of the generation process and how these parameters are used are presented in Appendix A
4.3.5.1 Object Scale (Blurring Factor for Object Generation)

Object scale of synthetic objects is determined by width of a Gaussian blurring function set by this parameter. The use of a larger value results in larger but fewer objects. Units are standard deviations of a Gaussian process in output image pixel size units. This value was typically set to 1.0.

4.3.5.2 Number of Type I Classes

Number of Type I Objects is the number of objects that will be filled in both images with random constant values chosen from a standard zero-mean Gaussian distribution. Any non-negative integer may be used. Values in the range 0 to 11 were tried.

4.3.5.3 Number of Type II Classes

Number of Type II objects sets the number of synthetic objects that will be filled with Gaussian noise, and offset from zero mean by random values chosen from a standard Gaussian distribution in one output image, and will be filled with random constant values chosen from a standard Gaussian distribution in the other image. The generator will deterministically alternate the objects chosen with each type of fill such that if this parameter is
set to an even number, exactly half of the objects of this type will have a constant value in image A and the other half of the objects will be filled with noise. Any non-negative integer may be used. Values in the range 0 to 11 were tried.

4.3.5.4 Number of Type III Classes

This sets the number of objects that will be filled in both images with Gaussian noise, and offset from zero mean by random values chosen from a standard Gaussian distribution. Any non-negative integer may be used. Values in the range 0 to 11 were tried.

4.3.5.5 Gaussian Blur (PSF Standard Deviation)

Gaussian Blur sets a width parameter for a simulated sensors’ optical point spread function prior to the spatial effects of the detectors themselves. The width is specified in pixel units corresponding to the one standard deviation point on the Gaussian function. It is implemented as convolution with a sampled Gaussian function truncated by a square of size Gaussian Blur * 4.5, rounded up to the nearest integer. Value used in this work was 0.5 chosen to mimic a hypothetical sensor that has a Gaussian PSF with a standard deviation one half of a pixel in width.
4.3.5.6 SNR

The simulated signal to noise ratio (SNR) is equal to the ratio of image variance to added white gaussian noise variance. The SNR is established at the simulated sensor’s detector stage to conform to the imaging model. Any non-negative value may be set. Values used were in the infinity to 0.1 range.

4.3.6 Monte Carlo

4.3.6.1 Number of Runs (N)

N applies to both synthetic image generation and to degraded real image generation. This was the number of image pairs to be tested for co-alignment response from the IT measures. Good error magnitude density estimates were obtained for N around 400. The objective was density and distribution estimates for IT measure performance. For each point at which the estimate is made (error magnitude), there should be several Monte Carlo iterations with that resulting error magnitude. If error magnitudes are clustered in a narrow range, the number of runs needed is less than if they are distributed throughout the estimation range. Ultimately, the sample size for each point must be large enough so the graphs distinguish sought performance features. Experimentation with different N within computing time restrictions reveals sufficient numbers.
4.4 Datasets

4.4.1 Permanent Data

The Permanent Data (Table 4-3) was available for testing of IT co-alignment estimators and is available for further testing in the future.

4.4.2 Volatile Synthetic Data

Synthetic data for repetitive testing was generated with random parameters, and was not stored beyond its application in each Monte Carlo iteration. Parameters are described in Section 4.3.5. and explained in Appendix A.
<table>
<thead>
<tr>
<th>Image Pair</th>
<th>Sources</th>
<th>date A/ date B</th>
<th>Attributes/ purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Independent random pixel</td>
<td>White Gaussian noise/ Same image as A</td>
<td>n/a</td>
<td>Complete dependence between pair; Maximal spatial bandwidth/ Determination of ultimate performance on ‘easiest dataset’ ( r = 1.0 )</td>
</tr>
<tr>
<td>Urban SAR-TM 3</td>
<td>ERS SAR/ TM5 band3</td>
<td>960111/ 960112</td>
<td>Multisensor; Phoenix commercial area/ Testing of SAR-optical co-registration</td>
</tr>
<tr>
<td>Urban SAR-TM 4</td>
<td>ERS SAR/ TM5 band4</td>
<td>960111/ 960112</td>
<td>Multisensor; Phoenix commercial area/ Testing of SAR-optical co-registration</td>
</tr>
<tr>
<td>Agricultural TM 3/4 uncorrelated</td>
<td>TM5 band3/ TM5 band4</td>
<td>960111</td>
<td>Multispectral; Agricultural area near Phoenix/ Registration testing on uncorrelated data; ( r \approx 0 )</td>
</tr>
<tr>
<td>Agricultural TM 3/4 multitemporal</td>
<td>TM5 band3/ TM5 band4</td>
<td>960111/ 960602</td>
<td>Multispectral; Multitemporal Agricultural area near Phoenix with 5 month time separation/ Registration testing on data with partial alteration of scene between images.</td>
</tr>
<tr>
<td>Agricultural TM 3/6</td>
<td>TM5 band3/ TM5 thermal band (6)</td>
<td>011196</td>
<td>Multispectral; Agricultural area near Phoenix; band 3 downsampled to 120m/ Registration testing on optical-thermal data</td>
</tr>
<tr>
<td>Ikonos 3/4 Maricopa</td>
<td>Ikonos band 3/Ikonos band 4</td>
<td>010726</td>
<td>Multispectral; Maricopa, AZ agricultural area; Includes areas with zero correlation/ Testing on uncorrelated data; ( r \in (-0.85, 0.9) )</td>
</tr>
</tbody>
</table>

**Table 4-3 Permanent datasets.** Description of image sources, attributes, and intended purposes
CHAPTER 5 DESCRIPTION OF RESULTS

This chapter describes the preliminary tests of all IT measures related to Fisher Information. Initial tests are described briefly indicating plausibility of utilizing simple forms of the IT measures (Section 5.1), then the chapter progresses to testing with direct correction of interpolation artifacts in Section 5.2 that were discovered in the initial tests, and then concentrates of the latter forms of the IT measures (in Section 5.3), with tests on real imagery, Monte Carlo tests on low resolution imagery derived from real high resolution imagery, and Monte Carlo tests on synthetic imagery. The goal of all testing is to ascertain the sub-pixel precision of the alignment estimators, as well as their reliability.

5.1 Initial Tests on Fisher Information

Initial testing (Phase 1) was performed to determine if Fisher Information applied to the JMS is at all an indicator of co-alignment. The single Fisher Information measure tested in this section is described in Eq. (3.3). In this section, stressing IT alignment estimators by addition of noise was not done.
5.1.1 Correlated Image - i.i.d. Pixels

Figure 5-1 shows response of Fisher Information to a white Gaussian noise image shifted with respect to itself. The JMS Entropy is also plotted, showing its opposite response. Both measures are sensitive indicators of co-alignment.

This image pair shows interpolation artifacts despite the strong responses of the IT measures. In this case, the artifacts are periodic with a 1 pixel period. They bias the measures differently at different search (shift) positions, and therefore distort the shape of the IT responses in a way that may move the location of the extremum away from where it should be. The artifacts in the Fisher Information response are substantially smaller than in the Entropy response relative to the extrema indicative of co-alignment (34:1 versus 6:1).
5.1.2 Uncorrelated Image - Real Imagery

To verify that Fisher Information does not require correlation between the two images, an image pair that is nearly uncorrelated was extracted from a Landsat 5 TM scene. The images, shown in Figure 5-2, are from the red and near infrared bands at a location where the two have little correlation in the 128x128 window. In Figure 5-3, the result shows correct indications of co-location from both IT measures. An artifact pattern is superimposed.

Figure 5-2 Uncorrelated image set from Landsat5 TM agricultural scene. These subsets of (left) red band and (right) near infrared band are 191x191 pixels to accommodate processing in a 128x128 window.
on the Fisher Information response. For comparison, the correlation coefficient (Eq. (1.1)), and the response of phase correlation (Eq. (1.3)) are presented. They both failed to respond to the co-alignment.

Figure 5-3 Alignment estimation of uncorrelated images using Fisher Information, Entropy, correlation coefficient, and phase correlation. Green is JMS Entropy; blue is Fisher Information with respect to JMS gradient; red is the correlation coefficient; and cyan is phase correlation. Circles identify the locations of ‘on-pixel-center’ points. Window Size: 128x128; Search: +/- 4 pixels in 0.05 pixel steps; Binsize = 0.5 standard deviations.
5.1.3 Correlated Image - Rotation, Scale and Translation Estimation.

To estimate correct scale and rotation of images, in addition to rigid translation, the search space included these spatial transformation parameters for three examples. Each search was performed over two parameters at a time to combine all combinations of translation in one direction, rotation about the image center, and scale variation from the image origin at the upper left hand image corner.

*Figure 5-4* ERS SAR test image. Phoenix, AZ urban area. The image is square root stretched for visual enhancement.

The test image was an ERS SAR image of a Phoenix, Arizona commercial and urban area (Figure 5-4). Results from the combination of translation and rotation are presented in Figure 5-5. Combination of translation and scale search are depicted in Figure 5-6. Finally, rotation and scaling are searched together with the result in Figure 5-7. It is thus demonstrated that the response of IT alignment indicators is an extremum at the co-align-
Figure 5-5 Responses to translation and rotation. (a) shows response of Fisher Information with respect to JMS gradient; (b) is JMS Entropy; (c) is correlation coefficient. Image is Phoenix ERS SAR image. Window Size: 64x64; Search: +/- 4 pixels 1D translation in 0.075 pixel steps, and +/- 0.05 radian rotation in 0.0005 radian steps (2.3 pixel total displacement at window corners); Binsize = 0.25 standard deviations.
Figure 5-6  Responses to translation and scale. (a) shows response of Fisher Information with respect to JMS gradient; (b) is JMS Entropy; (c) is correlation coefficient; and (d) is phase correlation response. Image is Phoenix ERS SAR image. Window Size: 64x64; Search: +/- 4 pixels 1D translation in 0.05 pixel steps and 0.925 to 1.075 scale in 0.002 steps (+/- 4.8 pixel change in alignment at one window border; zero at the other); Binsize = 0.25 standard deviations.
Figure 5-7 Responses to scaling and rotation. (a) shows response of Fisher Information with respect to JMS gradient; (b) is JMS Entropy; (c) is correlation coefficient; and (d) is phase correlation response. Image is Phoenix ERS SAR image. Size: 64x64; Search: 0.925 to 1.075 scale in 0.002 steps and +/- 0.05 radian rotation in 0.0005 radian steps; Binsize = 0.25 standard deviations.
ment point for each of these spatial transformations. The IT measures only indicate similarity of the two images, without any directional sensitivity to the transformation. This helps explain the similarity in these results and also supports a prediction that the type of continuous spatial transformation does not determine the nature of the response, i.e. it is an extremum. A search with any chosen spatial transformation that encompasses co-alignment and inverts the distortion that caused mis-alignment, identifies the transformation needed to re-align the objects. Because of this, testing with different spatial transformations was discontinued, and only translation searches were made from this point on.

### 5.2 Marginal Density Correction Tests

Phase 2 algorithms are the same as for Phase 1, but the images subjected to interpolation are corrected to pre-interpolation density using the method described in Appendix D. Fisher Information derivatives are also bandwidth limited, hence the ‘Sigma’ parameter is introduced. The testing was done with independent noise added to both images in the image synthesis to determine if the methods are suitable for low SNR image pairs.
5.2.1 Synthetic Test Images

Testing of density correction was only performed on synthetic images generated as described in Appendix A. In Figure 5-8 is shown a typical result. The interpolation artifacts were mitigated, successfully for the Fisher Information measure, however, noise handling was poor. Figure 5-8 shows successful performance at the highest noise level that allowed observation of correct alignment estimations with several successes observed out of about 50 trials. At a pixel SNR of 1.0, the IT alignment response plots are rough, leading to high variance of estimated position. In this particular run, the Fisher Information measure produced a peak near the location of correct alignment, but Entropy failed. This

Figure 5-8 Responses to Density corrected data.
Synthetic data with 12 object classes: 6 Type I; 4 Type II; and 2 Type III classes. The spatial blur used to generate the objects was 1.0 standard deviations; image blurring was 1.0 standard deviations; method chosen was positive product; Window Size: 64x64; Search: +/- 2.0 pixels in 0.08 pixel steps; Sigma = 2.0; Binsize = 0.25 standard deviations; SNR = 1.0
figure may be compared to the divergence results presented in section 5.3.3. that show comparable performance at worse noise scenarios. Little characterization was done with direct IT measures because higher performance was sought, and was realized with divergence measures.

5.3 Fisher Information Divergence Forms

In this section (representing Phase 3), the latest alignment measures are tested. They are all derived from Fisher Divergence. Mitigation of interpolation artifacts through marginal density restoration is discontinued. Testing of the measures is extended to lower SNRs. The preliminary tests of the divergences indicated that they are successful methods, prompting more extensive testing with real and synthetic imagery.

5.3.1 Real Images

Information about the datasets, including dates of acquisition are tabulated in Table 4-3.
5.3.1.1 TM Band3/Band4 Uncorrelated Agricultural

To test and demonstrate functionality of the divergence-based IT measures on uncorrelated data - an agricultural area with near-zero correlation between TM bands 3 and 4 is processed. The results indicate that correlation is not necessary for alignment of the images using IT-based divergences. The responses, with exception of the positive product divergence method clearly show an extremum near the (0,0) shift in x, y, where the correct alignment is believed to be (Figure 5-9).

In Figure 5-10 are shown the responses to band 3 co-registered with itself. Figure 5-11 is likewise for band 4. The responses are substantially sharper than for the band-to-band case in Figure 5-9. Total correlation and lack of differential noise explain the reduced broadening of the main response. The responses in Figure 5-10 and Figure 5-11 are not delta functions because of intrinsic scene spatial correlation, spatial correlation due to sensor PSF, and some low level of spatially correlated sensor noise. The interpolation scheme also broadens the response. Extremely sharp responses of auto-co-alignment to spatially uncorrelated data is demonstrated in Section 5.3.3.1.
Figure 5-9  Divergence responses to uncorrelated TM agricultural scene. (a) is TM5 red band; (b) is JMS scattergram; (c) is TM5 near infrared band image; and (d) are responses of IT and correlation measures. Window Size: 128x128 pixels; Search: +/- 3 pixels in 0.2 pixel steps; Sigma = 0.25; Binsize = 0.25 standard deviations.
Figure 5-10 Responses to Figure 5-9 red band data co-aligned with itself. (a) is the TM5 red band image; (b) is JMS scattergram; (c) is responses of the estimators. The settings of the estimators are the same as for the red/near infrared run.
Figure 5-11 Responses to Figure 5-9 TM5 near infrared band data co-aligned with itself. (a) is the near infrared band image; (b) is JMS scattergram; (c) is responses of the estimators. The settings of the estimators are the same as for the red/near infrared run.
5.3.1.2 TM Band3/Band4 Multi-Temporal Agricultural Scene

To meet the needs of change detection, the assumption that the object boundaries are identical in both images may be necessarily violated. In multi-temporal data, the objects may change within their boundaries in terms of texture parameters, much like multi-modal imaging, but there may also be alteration of some object shapes, and some objects may be replaced with different ones. Such a case is tested in this section. The multitemporal Landsat 5 TM scene is of farmland North-West of Phoenix, AZ. In Figure 5-12, the scene shown is in the near infrared band acquired in June 02, and 01 November, 1996 showing substantial changes in vegetation density, but not layout (shapes and positions of objects are similar).

![Image B]

Figure 5-12 TM Change detection scene acquired over 5 month interval. TM5 band 4 acquired on (a) 960602, and (b) 961101.
The difference between this situation and the runs in Section 5.3.1.1 is that the images in
the two bands were acquired approximately 5 months apart. The scene location is also dif-
ferent, and some correlation exists between the two images (bands). Results are shown in
Figure 5-13. Only the gradient and the absolute product Fisher Information forms have
successfully found evidence of alignment at reasonable locations. Despite a correlation
coefficient of about +0.14, both correlation-based methods have failed. This is the first
presented test of partially correlated images.

5.3.1.3  TM Band 3/Band 6 Agricultural

Although the red band and thermal infrared band are from the same sensor in this instance,
the physical processes are differentiated not just by wavelength, but also by sensing mech-
anism (Solar reflective versus thermal emissive). The ground sample interval (GSI) is
120m for the thermal band, but was up-sampled to 30m by the provider of the dataset.
The resolution of both was reduced by the author to 120m (native GSI of TM thermal
band) using pixel averaging interpolation prior to processing.

Results are shown in Figure 5-14. All eight methods show reasonable performance. The
correlation coefficient between these bands is approximately 0.3. Each image co-align-
ment with itself is shown in Figure 5-15 and Figure 5-16. The response of the IT-based
Figure 5-13 Divergence responses to multi-temporal and multi-spectral TM agricultural scene. (a) is TM5 red band acquired on 961101; (b) is JMS scattergram; (c) is TM5 near infrared band acquired on 960602; and (d) are responses of IT and correlation measures. Window Size: 128x128 pixels; Search: +/- 3 pixels in 0.2 pixel steps; Sigma = 0.25; Binsize = 0.25 standard deviations.
estimators has an apparently sharper shape than for the two band case as seen earlier (Figure 5-9 through Figure 5-11). Interestingly, the sharp identical noise induced peaks are not appearing on the self co-alignment estimates. This may be explained for the band 3 image because it is downsampled using pixel average interpolation making its noise more spatially correlated and attenuated. This makes the sharp peaks attenuated and broadened such that they blend in with the broad peaks generated by the spatially correlated data. The lack of the sharp peaks on the band 6 data remains unexplained.

5.3.1.4 ERS SAR/TM Band 4 Commercial Phoenix

This data combination was the most challenging image pair examined in this dissertation. It was difficult to ascertain alignment visually, to manually co-register the images, and the co-registration indications are weak.

Data for the experiment consisted of a part of the Phoenix area that has some streets, large buildings, and homes. The SAR image and the TM image were acquired on subsequent days. A manual co-registration was performed to align the lower resolution TM image to the SAR, thus up-sampling the TM image in the process. The original resolution of each image is different, violating the requirement of identical spatial imaging properties. Results of the IT-based alignment response are in Figure 5-17.
Figure 5-14  Divergence responses to agricultural scene in TM red and thermal bands.  (a) is TM5 red band; (b) is JMS scattergram; (c) is TM5 thermal band image; and (d) are responses of IT and correlation measures.  Window Size: 112x112 pixels; Search: +/- 2 pixels in 0.2 pixel steps; Sigma = 0.5; Binsize = 0.25 standard deviations.
Figure 5-15  Responses to Figure 5-14 red band data co-registered with itself.  (a) is the TM5 red band image; (b) is JMS scattergram;  (c) is responses of the estimators. The settings of the estimators are the same as for the red/thermal infrared run.
Figure 5-16 Responses to Figure 5-14 thermal infrared band data co-registered with itself. (a) is the TM5 thermal infrared band image; (b) is JMS scattergram; (c) is responses of the estimators. The settings of the estimators are the same as for the red/near infrared run.
Figure 5-17 Responses to ERS SAR image shifted against Landsat 5 TM band 4 (NIR) of urban/commercial Phoenix area. Shifts are +/- 2.25 pixels in x and y in 0.15 pixel steps. Image Window Size is 96x96 pixels; Sigma=1.5; Binsize=0.5. (a) is image A (SAR) without bandwidth reduction; (b) is the JMS scattergram at estimated best alignment; (c) shows image B (TM5); and (d) indicates IT measure alignment responses. Numerical results are in (row, column) coordinates.
The Fisher Information Divergence, Negative Product, and Absolute Product measures indicate a peak at (-0.8, +0.75) (x,y) pixels that is believed to be near the correct location of the co-alignment, possibly revealing a manual co-registration error. Manual verification of the alignment errors was accomplished by rotation, scale, and translation (RST) distortion modeling using 11 matching control points. The average translation position was determined at (-0.64, +0.30) that is well within a global RMS error of 0.98 pixels. The Entropy measure shows a prominent peak further away that has not been reconciled with object offsets in the images.

To compare how sharp the indications may potentially be, the images were co-registered with themselves. In Figure 5-18, the SAR image is co-registered with itself. This revealed the sharpness limitation imposed by the SAR image properties alone. The same limitation due to the TM image is shown in Figure 5-19. It is evident that the TM image limits the resolution more that the SAR image as is expected due to its lower innate resolution. Figure 5-19 is repeated with the spatial transformation resampling set to perform a bicubic interpolation rather than bilinear (Figure 5-20). The results appear identical with exception of slight changes in the Phase Correlation sidelobes, and the small variation atop the broad peak of the Positive Product response. This justifies use of the faster bilinear interpolation
Figure 5-18  *ERS SAR image co-registered with itself.* This indicates the limit to sharpness imposed by the properties of this image in a co-registration of it with another image. (a) shows the SAR image; (b) shows the JMS scattergram; and (c) are the IT measure values versus shifts. All parameters are same as in Figure 5-17.
Figure 5-19  Responses to TM band 4 image co-aligned with itself.  This is a counterpart to Figure 5-18, run with the same parameters.
In both auto co-registrations, the Entropy divergence shows a very sharp central peak upon the broad peak. This is due to uncorrelated noise, probably from the sensor and quantization noise, which have a much greater bandwidth than the imaged objects. This peak does not show up on the Fisher Information surfaces, due to the severe bandwidth reductions imposed on the Fisher Information derivative. A reduction in the bandwidth sigma parameter and use of smaller bin size reveals the peak on all plots as shown in Figure 5-21. Figure 5-19 was generated with settings of Figure 5-17 for comparing the two.

**Figure 5-20**  Figure 5-19 results repeated with spatial transformation realized through bicubic interpolation. It is otherwise a repeat run of Figure 5-19.
The settings of Figure 5-21 applied to the run of Figure 5-17 would result in poor noise immunity and failure of the alignment estimation. Without identical noise in both images, the peak disappears, and therefore is not seen in Figure 5-18 and Figure 5-19.

**Figure 5-21 Sharpening effect of finer JMS estimation.** This is same as Figure 5-19, but with binsize = 0.1 and Sigma = 0.25.
To approximately meet the assumption of same imaging PSFs, the SAR image was blurred with a Gaussian convolutional filter of 1.5 pixel standard deviation as a PSF matching measure (See Figure 5-22). The sharpness matching was approximated visually by varying the Gaussian standard deviation to match the blur. Even though the peaks that indicate the correct alignment are still visible, they are broadened, and other anomalous features have become relatively larger, obscuring the exact extrema.

Only the Negative Product and Absolute Product measures indicates a reasonable alignment location of ( -0.45, +0.3), as compared to the manually determined (-0.64,+0.30). It is difficult to say whether this result or that from un-matched image PSFs is actually more precise, as both are well within the RMS error of the manual estimate. The Gradient measure fails to indicate a reasonable result with the blurred data. Qualitative assessment of the IT measure surfaces indicates that blurring to match PSFs might be detrimental to alignment estimator precision and may not be prudent.

It may be important to note that on the SAR-optical results is an elongation of many features along the x shift axis. This is approximately the cross-track direction for the SAR, in which error is brought about by vertical dimension of ground targets. The author speculates that the resulting spatial disparity may be responsible for the elongation. Another potential reason is different resolution in the cross-track and along-track directions.
Figure 5-22  Responses to bandwidth-reduced SAR co-registered with TM5 band 4.
Parameters and individual captions are the same as for Figure 5-17.
Visual inspection of the larger SAR scene in which very bright point reflectors are seen, including ones that display a symmetric cross shaped response, indicate a fairly equal PSF component in the x and y directions that extends many pixels from the center, however, the main lobe of this impulse response appears wider in the cross-track direction possibly supporting a SAR spatial anisotropy hypothesis.

5.3.2 Degraded Real Imagery - Monte Carlo

Ikonos is a high resolution commercial remote sensing system with 4m resolution, multi-spectral and 1 m panchromatic imagery. It is operated by Space Imaging LLC. It serves here as a high resolution approximation to ‘truth’ for simulating lower resolution imagery. The IT measures were tested on Ikonos imagery that was blurred, downsampled, and made more noisy, to emulate a lower resolution sensor operating at poor signal to noise ratio conditions. A portion if an Ikonos scene was located and subset where the correlation coefficient between red and near infrared bands is close to zero, and remains so with perturbations of the location by a few pixels. The test image generation process, i.e. degradation and location choice are described with more detail in Appendix B.

The same image with randomized sampling phase in the downsampling, and different image noise samples, was used to study performance over a variety of phase and noise situations. This emulated the lack of control over these two aspects of imaging, as is often the case in real image acquisition. The process was repeated for seven noise levels: 0.0;
0.2; 0.5; 1.0; 2.0; 3.0; 4.0 corresponding to pixel SNRs of: infinity; 5; 2; 1; 0.5; 0.33333; and 0.25. Noise was added independently to both images using zero-mean Gaussian i.i.d. values for each pixel. Forty-four repetitions were performed for each noise level; the only change between repetitions was the noise sample applied to the images and the random sampling phase.

Statistics indicate how many trials succeeded out of each of the runs. These are plotted as fractions of 100 percent success. Success is declared if the estimate obtained is within a box of +/- 0.5 pixel in x, y. All seven IT methods are tested, and some results for the correlation-based measures are included, even though they are not successful with this dataset. Accuracy statistics are also plotted. The 90\textsuperscript{th} percentile of the error is shown for the seven IT measures. This indicates the level of estimation error magnitude that is not exceeded 90 percent of the time.

Errors are statistically described by estimating their density versus the magnitude of the error. For translation errors, this is the Euclidean distance in the x, y search space from zero error coordinates. The plot abscissa ranges from zero error to the Euclidean distance from the center of the square search space to its corners. Density plot ordinates are in units of fractional occurrence rate, and the plot points sum up to one. Cumulative distribution plots have the same abscissa, and indicate cumulative occurrence rate from zero to one.
The dataset is shown in Figure 5-23 at a SNR of infinity (no noise added) and a SNR of 0.5. At one extreme, no noise was added (Figure 5-24). The correlation-based measures failed to indicate an alignment. The positive part of Fisher Information product measure performed poorly, and the remaining IT measures successfully located the co-alignment with some error.
Figure 5-23 Test images subset from IKONOS scene. Maricopa, Arizona. Top row: left is red band; right is near infrared band with no noise added. Middle row is the same with a SNR = 5. Bottom row is same with a SNR = 0.5.
Figure 5-24  Responses to degraded IKONOS image at SNR = infinity. X direction scale is to the left, and y is to the right. This will be the case in subsequent surface plots.
A plot of the error distance cumulative distributions is in Figure 5-25, and the corresponding error distance density plot is shown in Figure 5-26. These figures allow comparison of accuracy among the different estimators.

For the SNR = 0.5 case, the estimator responses are shown in Figure 5-27, and corresponding error distance plot is in Figure 5-28 and Figure 5-29. At this noise level, and in this particular run, all of the IT-based estimators responded with indications of co-alignment location. The peaks appear to be more noisy, but the extrema are still at correct loca-
tions within reasonable error. The error distance distribution plot shows a consistency of performance for all of the measures, however, the Fisher Information gradient and positive-part-of-product measures are performing distinctly worse at this noise level than other IT methods. The density plot appears scattered because the sample size of 44 runs per SNR was not insufficient provide for a smooth density estimate for this high noise level. At higher noise levels, the IT measures become intermittently successful.
To compare the performance of the estimators over the operational noise level range, the failure rate and the 90th percentile of error distance are plotted versus the noise level in Figure 5-30. Fisher Information positive-part-of-product, absolute-value-of-product, and Entropy maintained the best success rate over all noise levels. As for lowest error magnitudes in the 90th percentile, the Fisher Information positive-part-of-product, absolute-value-of-product, Entropy, and the mean result between the last two are the most consistent. This leaves only Fisher Information absolute product and Entropy as the two methods that can be distinguished as having the highest number outcomes in the 0 - 0.1 pixel error magnitude range.

*Figure 5-27 Typical responses to degraded IKONOS image at SNR = 0.5.*
With no added noise, the imprecision of the alignment estimates is usually better than 0.3 pixels. Interestingly, the most common errors are at 0.1 rather than 0.0 pixels. This may be due to resampling effects, imprecise estimation, or mis-registration of the input images. Space Imaging LLC does not specify the co-registration accuracy of their product. Visual inspection of this IKONOS image indicates that bands 3 and 4 are aligned to no worse than one pixel. Lack of truth makes understanding the exceptionally frequent 0.1 pixel error difficult to explain. At a $\text{SNR} = 0.5$, the imprecision of the estimators is commonly reported to be around 0.4 pixels, and IKONOS data mis-registration can no longer be suspected as the main cause.

\textit{Figure 5-28  Cumulative error magnitude distribution at pixel SNR = 0.5.}
5.3.3 Synthetic Images - Correlated i.i.d.

5.3.3.1 Verification with Single Synthetic White Noise Image

As was done in tests of the simplest Fisher Information form (Section 5.1), the Divergence forms were tested to determine their response to a white noise image. This image paired with itself, represents an unlimited bandwidth within the constraints of uniform sampling,
Figure 5-30  Performance of IT alignment estimators vs. seven noise levels. Upper graphs shows success rate; lower shows 90th percentile on error magnitude (vertical scale) in pixels. The horizontal scales are noise variance - inverse of SNR.
Figure 5-31 Responses of the six divergence-based algorithms to a white noise image shifted with respect to itself. Shifting is +/- 1.5 pixels in 0.1 pixel steps; Sigma = 0.25; Crop size is 192x192 pixels; Bin size = 0.25.
and no differences between the images in the pair. In Figure 5-31, results are shown for alignment estimation with the estimators and search performed using settings that were appropriate for real imagery.

Experiments with these data indicated that co-registration precision was practically unlimited. By not applying bandwidth limits on the derivatives, and making the JMS bin size very small (in order of 0.05 standard deviations.), then precision of 0.00001 pixels is readily achieved as shown in Figure 5-32. The responses in this run could have been further sharpened by use of a much smaller bin size. Such extreme performance is not available with real imagery due to noise, limited spatial bandwidth, and in some cases, subtle spatial dissimilarity between the same objects in each image.

The product form measures failed in this run because their innately low bandwidth derivatives (derived from a second order difference) do not respond to the extremely fine structure in the JMS that occurs with perfect co-alignment. They actually respond better to the broader, more diffuse structure corresponding to slight mis-alignment. The Fisher Gradient method responded with a peak because it uses the first order difference derivative approximator that responds well to fine structure.
Figure 5-32 Demonstration of 0.00001 pixel co-registration precision. Data is same as for Figure 5-31 showing the extrema at much finer sampling. Fisher Information gradient and Entropy-based divergences properly indicate the correct alignment. Other variations invert due to bandwidth limitations of derivative approximations.
5.3.4 Synthetic Images - Monte Carlo on Random Scenes.

One purpose of experiments in this section is determination of IT measure performance on estimating alignment of random scene imagery. Thousands of real images are not available to the author for determining IT measure universality on an ensemble of various imagery. In this section, random objects were generated in large numbers and filled with particular textures.

Another purpose of testing alignment estimators on synthetic imagery is absolute estimation error assessment. In the previous section, registration error was reduced from less than +/- 1 pixel in the real image to less than +/- 0.2 pixels in a degraded version. Here the objects are perfectly aligned allowing determination of smaller errors.

The basis of generating the synthetic images is synthesis of object boundaries that are common to both images. The interiors of the objects are then filled with a texture, independently for each image. Textures may be either a random constant or i.i.d. random values offset by a random constant. Objects in the image pair belong to classes. Each class has a pair of fixed probability distribution parameters that determine texture independently for objects in each image for that class. The classes are of three types according to how objects are filled: (I) constant in each image; (II) constant in either one image and i.i.d. in the other image; and (III) i.i.d. in both images. These three types are intended to represent statistical ‘endmembers’ spanning object textures. Textures of real objects are
presumed to be somewhere between the characteristics of these types. Testing of the IT measures on each type alone and on combinations of types verifies the suitability of the measures for any imagery ‘spanned’ by these statistical types.

In Appendix A, the synthetic image generation is described in more detail, including explanations of the types and numbers of objects generated. A set of visual examples are presented in this appendix to visually show the variety of images with different types of textures. Noise levels in this Section were set much higher such that the differences in algorithm performance were most distinguishable and were set by experimentation. At higher noise levels, the failure rate would have prevented comparison, and at lower noise levels, excellent performance of all measures also would have prevented comparison. To determine which algorithms are at all suitable for a particular type of data, runs were made with no added noise.

The test images contained objects with random pixel values, and much noise was added, resulting in images that have nearly indistinguishable objects. An example pair of images used in 5.3.4.2 is presented in Figure 5-33. Due to the lack of visually discernible information (and they all look characteristically alike), examples are not shown for the other experiments.
For these synthetic data, the correlation-based methods rarely succeeded. To obtain success and precision statistics, Monte Carlo runs were made with randomized image generation and noise samples. The number of runs (sample size) was set to obtain distinguishable performance characteristics between the IT methods. Success is defined for section 5.3.4 as half of the search range, or any result in a +/- 1.5 pixel box around zero error. Precision estimates were sampled at the same interval as the search step size (0.2 pixel spacing).

Figure 5-33  Examples of synthetic image pair, as generated for Section 5.3.4.2. They contain two Type I Classes. SNR = 0.143. Type I objects are the easiest to visually distinguish. This illustration shows that the IT alignment estimators can function at noise levels where human vision barely recognizes objects at all.
Errors are statistically described by estimating their density versus the magnitude of the error. For translation errors, this is the Euclidean distance in x, y space from zero error coordinates. The plot abscissa ranges from zero error to the Euclidean distance from the center of the square search space to its corners. Density plot ordinates are in units of fractional occurrence rate, and the plot points sum up to one. Cumulative distribution plots have the same abscissa, and indicate cumulative occurrence rate from zero to one.

5.3.4.1 Zero SNR behavior

In order to statistically identify failed behavior that is caused by excessive noise, a run was made at an SNR of zero. The two input images had no common information within their frames, forcing all methods to fail, as no correct solution exists. The experiment performed here was intended to help identify this mode at very low SNRs, but there may be failure modes with other causes than a high differential noise level and not represented by this example.

In Figure 5-34 is the set of IT measure surfaces. Table 5-1 indicates the random chance outcomes classified as successes, and Figure 5-35 and Figure 5-36 respectively show the result error magnitude density and cumulative distribution. Even though the two images had no common information and co-alignment did not exist, computation proceeded as though the input images are perfectly aligned. The alignment estimation error magnitude occurrence rates for IT measures were close to what would be expected if the measures
have their extrema distributed uniformly over the +/- 3 pixel square search space. The
Entropic Divergence appears to follow a straight line for a 3 pixel error magnitude range,
whereas the Fisher Information methods tend to have a higher error density in the 1.5 to
3.0 pixel error magnitudes than for 0.0 to 1.5 pixels.

The method constituted by averaging the Fisher Absolute Product and Entropy estimates
shows error magnitudes substantially reduced from the error magnitudes of its compo-
nents. This is indicative of strong independence of the alignment estimates from the two
IT measures given the same unrelated input images.

\begin{table}[h]
\centering
\begin{tabular}{|l|c|}
\hline
Method   & Success rate \\
         & 361 trials \\
\hline
Gradient & 0.211 \\
Product  & 0.263 \\
Negative & 0.255 \\
Positive & 0.230 \\
Abs. Val. & 0.186 \\
Entropy  & 0.313 \\
\hline
\end{tabular}
\caption{Success rate for SNR = 0 case. Images are evaluated in 128x128 pixel
window. The search is +/- 3.0 pixels in 0.2 pixel steps; \(\text{Sigma} = 2.0;\) \(\text{Binsize} = 0.5.\)
The content of the image is irrelevant.}
\end{table}
5.3.4.2 Two Type I Classes

This configuration contains only objects that are constant throughout their extent. It was used for testing at three noise levels. Table 5-2 (a) indicates the success rates for these conditions. The no-added-noise error magnitude density is displayed in Figure 5-37. Most of the methods show zero error performance except for the Fisher Information product measure that occasionally has a small error.

Testing at a high noise level was done at a SNR = 0.143. The success rate table is shown in Table 5-2 (b). The cumulative error magnitude curve (Figure 5-38) indicates that there are some instances where the IT measures performed well, but most of the time the errors were uniformly distributed in the manner seen in the zero SNR case. In between relatively precise operation and failure, there is little or no continuity, i.e., the methods succeed and produce a high precision result, or they fail. Figure 5-39 demonstrates this by error magnitude densities are clearly bi-modal for a SNR = 0.143. At this time, the bimodal behavior can not be explained. The relative dearth of results with poor precision, that provide accurate information about the actual alignment (and therefore should not be considered failure) is not known. No tools exist at this time to peer into the formation of the IT mea-
Figure 5-34  IT measure example responses at SNR = 0. Parameters are listed in the Table 5-1 caption.
Figure 5-35  IT alignment estimate error magnitude density for SNR = 0. The approximately linear rise between 0 and 3 of the Entropic Divergence curve indicates a near uniform error distribution within a 3 pixel error magnitude relative to the search space center (zero error). Fisher Information measures appear to have a higher density in the 1.5 - 3 pixel error magnitude range compared to the 0 - 1.5 pixel range. The non-linear drop-off above 3 pixel radius appears consistent with uniformly-distributed errors in the x-y plane, appearing in vanishing corners of a 6x6 pixel square search box, as a function of increasing error radius. Parameters are listed in the Table 5-1 caption.
Figure 5-36  IT alignment estimate error magnitude cumulative distribution for SNR = 0. The approximately parabolic rise of the curves (integration of a straight line) up to a three pixel error magnitude indicates a near uniform error distribution up to a 3 pixel search radius around zero shift. The deviation from uniformity that was apparent in the density plot of Figure 5-35 is more difficult to see. Parameters are listed in the Table 5-1 caption.
sure extrema that would allow connecting the image noise samples to the results and would indicate the nature of the sharp failure threshold. The separability seen here suggests that failures may be robustly detectable according to error magnitude.

The black curve in Figure 5-38 and 5-39 shows the cumulative error magnitude and error magnitude density, respectively, for the Entropy divergence extremum position averaged in 2D with that of the Fisher Information divergence absolute product method. At the right part of the graph, where the individual results are considered failures, the averaging shows substantially reduced error compared to both if its components, demonstrating that Fisher Information and Entropy are fairly independent in their responses. The lower errors of the average are due to the central tendency of the averaging. Unfortunately, these random outcomes do not contain alignment information. At the left side of the graph, where all measures are successful, the black curve runs between the error curves of its component measures indicating high dependence.

At a SNR = 0.5, the estimators were usually successful. The success rate table is shown in Table 5-2 (c). A typical IT estimator surface set for this SNR is shown in Figure 5-40. Comparison of performance may be made by looking at the error magnitude density (Figure 5-41), and cumulative distribution (Figure 5-42). The Fisher Information gradient and
absolute value of product measures had a slightly higher occurrence rate than all of the other measures. The Entropy divergence performance was worst in terms of precision despite highest success rates.

5.3.4.3 Two Type II Classes

This alternation of textures between images within objects is a case where statistical dependence is at a minimum if the images are co-aligned and the number of classes is even. With no noise added, all estimators were successful in a run of 361 trials. An exam-
Figure 5-37  Error magnitude density for Type I objects with no noise added.
Parameters are listed in Table 5-2 (a) caption.

Sample of the results of this case at a pixel SNR = 0.35 are shown in Figure 5-43. Estimated probabilities of error are plotted in Figure 5-44. The success rates for each algorithm are summarized in Table 5-3.
The divergence measures tested with Type II objects responded with generally lowest error magnitudes and lowest failure rates as compared to results on objects synthesized with the same type of texture fill in both images.

Figure 5-38  Cumulative error magnitude. Type I objects, SNR = 0.143.
Parameters are listed in Table 5-2 (b) caption.
Figure 5-39  Error magnitude density for run in Figure 5-38 - Type I objects.

Parameters are listed in Table 5-2 (b) caption.
Figure 5-40  Type I object IT alignment responses at SNR = 0.5. This example shows 1 of 361 trials. Parameters are listed in Table 5-2 (c) caption.
Figure 5-41  Error magnitude density for Type I objects at SNR = 0.5 Parameters are listed in Table 5-2 (c) caption.
Figure 5-42  Cumulative error magnitude for Type I objects at SNR = 0.5

Parameters are listed in Table 5-2 (c) caption. The horizontal (error magnitude) scale is magnified to 0 - 1.6 pixels out of the 0 - 4.24 pixel search range.
5.3.4.4 Two Type III Classes.

In this case all objects in both images are filled with random valued pixels. The objects are only distinguished by different mean values of their pixel value distributions. Again, the baseline performance was established by not adding any noise to the texture filled images. The no noise success rate is listed in Table 5-4 (a), and the error magnitude density is shown in Figure 5-45. The noisy nature of the objects is similar to constant objects measured with large random errors. It was indicated by a larger alignment estimate imprecision and some failures at infinite SNR in all algorithms, not seen for the constant pixel

<table>
<thead>
<tr>
<th></th>
<th>Success rate 361 trials</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gradient</td>
<td>0.864</td>
</tr>
<tr>
<td>Product</td>
<td>0.695</td>
</tr>
<tr>
<td>Negative</td>
<td>0.859</td>
</tr>
<tr>
<td>Positive</td>
<td>0.837</td>
</tr>
<tr>
<td>Abs. Val.</td>
<td>0.856</td>
</tr>
<tr>
<td>Entropy</td>
<td>0.801</td>
</tr>
</tbody>
</table>

*Table 5-3 Success rates for Type II objects. Images are evaluated in 128x128 pixel window. The search is +/- 3.0 pixels in 0.2 pixel steps; Sigma = 2.0; Binsize = 0.5; 2 Type II classes.*
181

Figure 5-43  **IT alignment responses for Type II objects.**  This example shows 1 of 361 trials. Parameters are listed in Table 5-3 caption.

textures (Type I) in section 5.3.4.2. At a pixel SNR of 0.5, the success rates are listed in Table 5-4 (b), and error magnitude density is shown in Figure 5-46. A typical IT alignment response set is shown in Figure 5-47.
Figure 5-44  Cumulative error magnitude for Monte Carlo run using Type II objects. Parameters are listed in Table 5-3 caption.
This configuration of the image synthesis is to represent more complex scenes with numerous and different objects. All three object types were generated. To establish performance on the object responses themselves, Figure 5-48 shows infinite SNR error magnitude performance. As in previous cases, the Fisher Information product measure occasionally failed in this zero added noise case.

### Table 5-4 Success rates for Class III objects.

(a) with no added noise; (b) at SNR = 0.5. Images are evaluated in 128x128 pixel window. The search is +/- 3.0 pixels in 0.2 pixel steps; Sigma = 2.0; Binsize = 0.5; 2 Type III classes.
Figure 5-45  Error magnitude densities for Type III objects and no added noise.
The run parameters are listed in Table 5-4 caption.
Figure 5-46  Error magnitude densities for Type III objects. Pixel SNR = 0.5. The run parameters are listed in Table 5-4 caption.
Figure 5-47 Representative IT measure responses for Type III objects. Pixel SNR = 0.5. This example shows 1 of 361 trials. The run parameters are listed Table 5-4 caption.
Testing at a pixel SNR = 0.6 demonstrated that the remaining 5 IT measures can still work on this complex object imagery at substantial noise levels. The success rates are presented in Table 5-5. An example of the IT alignment responses is presented in Figure 5-49. For comparison of performance between the different measures, Figure 5-50 indicates the error magnitude density.
Figure 5-48  Error magnitude density for combined object type mixture and no added noise.  11 Type I classes; 3 Type II classes; and 5 Type III classes. Images are evaluated in 128x128 pixel window. The search is +/- 3.0 pixels in 0.2 pixel steps; Sigma = 2.0; Binsize = 0.5.
**Table 5-5 Success rate table for complex object type mixture.** Parameters are listed in Figure 5-48 caption.

<table>
<thead>
<tr>
<th>Method</th>
<th>Successes rate in 324 trials</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gradient</td>
<td>0.901</td>
</tr>
<tr>
<td>Product</td>
<td>0.759</td>
</tr>
<tr>
<td>Negative</td>
<td>0.843</td>
</tr>
<tr>
<td>Positive</td>
<td>0.840</td>
</tr>
<tr>
<td>Abs. Val.</td>
<td>0.898</td>
</tr>
<tr>
<td>Entropy</td>
<td>0.855</td>
</tr>
</tbody>
</table>
Figure 5-49  IT measure surfaces for combined object type mixture at pixel SNR = 0.5. This example shows 1 of 324 trials. Parameters are listed in Figure 5-48 caption.
Figure 5-50  Error magnitude density for combined object type mixture at pixel SNR = 0.6. Parameters are listed in Figure 5-48 caption.
Development of Fisher Information-based measures has yielded a new approach that can augment or replace Entropy-based co-alignment methods. The responses of the Fisher Information-based measures are different from Entropic measures, making Fisher Information complementary to Entropy in this application.

All of the IT methods of estimating co-alignment evaluated in this work share a strength over correlation-based methods in handling arbitrarily sensed images of a common scene. Unlike correlation-based methods, that rely on a presumed relationship of pixel measurements between two sensors, the IT methods are sensitive to shapes of features in the JMS rather than where the features are located in the JMS. This makes the methods less dependent on sensor responsivities and focuses them on properties of scene objects.

Testing on a broad range of imagery strongly suggests that the methods are sound for any imagery that depicts a set of objects in two images with equivalent spatial boundaries, but which have spectrally, statistically, or radiometrically arbitrary differences.
Experimentation with a number of Fisher Information-based image co-alignment estimators resulted in development of a new set of divergence formulas suitable for the task of sub-pixel alignment estimation. Preliminary results of Monte Carlo co-alignment estimations using synthetic images indicate that for some data, several of the Fisher Information-based divergences are simultaneously more robust and yield lower variance estimates of correct alignment compared to Entropy-based divergence - Mutual Information.

To verify performance of co-alignment estimators on real images, image pairs were selected from a variety of remote sensors. Most of the pairs were chosen to be difficult or impossible datasets for correlation-based techniques. Tests were performed on real data that included multispectral (optical/thermal), multispectral uncorrelated (optical), multispectral/multi-temporal (optical), and multimodal (SAR/optical). All IT methods were successful. The results were qualitatively interpreted, as the true co-alignment of these images is not known and therefore neither is the error. Fisher Information often appeared better as a co-alignment indicator than did Entropy.

Emulation of a low resolution image derived from a real high resolution image showed that Fisher Information-based divergences perform equally well with this data as does Entropic Mutual Information. Precision better than 0.25 pixels was achieved at a SNR as low as 4.0 with a success rate of 100%.
Testing of the IT measures on synthetic imagery, using independently generated image pairs with objects in common, generated to mimic textures of various objects sensed by various means, indicated that all five Fisher Information-based divergences are sometime suitables for random imagery. The method named Positive Product was least often successful, and it usually under performed the other four in terms of precision. All imagery tested had correlation coefficients between zero and one. It is expected that for anti-correlated imagery, the Negative Product method will perform poorly, and the Positive Product method will prevail. It is not understood why the Negative Product method performs better on uncorrelated imagery, even though the JMS patterns are aligned so as to not favor granularity sensitivity of either product method.

Some simple indications of conditions that differentiate the methods are: The Product forms are appropriate for imagery that have substantial noise and/or posses relatively much spatial correlation. The product forms are suitable for the most difficult co-registration problems. One of them usually substantially outperforms the other, but the one with poor performance also has a lower response magnitude making the response of the Absolute Product measure primarily from the better one. They may also fail to respond to very sharp JMS features such that responses to co-alignment of sharp and clean imagery are inverted, preventing identification of the co-alignment location. Sharper responses are obtained from the Gradient measure, and it may be operated with a very low Sigma on sharp and clean imagery. Entropy is suitable for low to moderate noise and/or spatial correlation, as it provides moderately sharp responses and usually produces usable responses
without adjustment (as it has no adjustments). For difficult cases, manual adjustment of Sigma for best performance of one or more of the Fisher Information measures results in outperformance of Entropy. Otherwise, no pattern was found that would predict which method is most suitable for a given type of imagery. Neither the scene type (urban, agricultural, or natural), sensing combination (optical, thermal, or SAR), nor texture type (I, II, or III) predict which method works best. Perhaps it is only the nature of the image noise sample or other degradation that determines the quality of a particular outcome, rather than the desired sensing characteristics.

On synthetic imagery, precision varied from unlimited for completely correlated images with uncorrelated pixels in each image to about 0.3 pixels at SNRs set to the breaking point of the method. Among the results of the six cross-compared measures, tested under conditions that cause high failure rates of any single measure, in most cases, at least one yielded plausible results (no failure). A comparison of non-failed results may reveal which ones are more precise. This implies that, as a set, the consensus of the methods usually recovers the needed alignment information and outperforms, on average, any single measure.

Computation of IT measures over a search space constitutes sampling of the IT values over the spatial domain. The true IT surface is not smooth, and has fractal properties over the translation ranges typically of interest, therefore aliasing must be taking place. The usual measures of anti-alias filtering are not directly applicable because the surface is not
available in a continuous domain prior to sampling, but is sampled through computation of each point. Reduction of the Fisher Information derivative bandwidth appears to smooth the response, but not with the usual frequency response control obtained by linear filtering. Fisher Information, as implemented here, nevertheless allows adding continuity to spatial searching, directly in the estimator in a way not available with Entropy.

All of the IT-based measures are slow compared to correlation-based measures because the lack the advantages of FFT efficiency, and unlike FFT methods, do not indicate correct alignment in one trial, but require a spatial search. A Monte Carlo run with 400 iterations and a 30x30 point search space typically takes about 30 hours to complete on Matlab R14 running on a 2.4 GHz PC. This is without numerical or algorithmic optimizations. For real imagery, the processing requires much less time, as one spatial search per image pair can provide the alignment estimate. Computation time is mostly spent on JMS estimation (sorting), and image transformation (interpolation and new pixel array assignment).

Main contributions summarized are:

(1) Presentation of a new divergence formula that is derived from Fisher Information. This work demonstrates that Entropies are not the only available information measures for direct information estimation. It also demonstrates that other information measure forms may also be converted into divergence measures that are non-redundant with Entropic Mutual Information;
(2) Derivation of five co-alignment estimators from the Fisher Divergence. A first use of Fisher Information as a solution of the alignment estimation problem is demonstrated. Multiple estimators may allow better understanding of co-alignment as manifested in the JMS. Availability of multiple, non-redundant measures, allows for combining of the results to build a more accurate or more reliable estimator than is possible using one measure alone. The results indicate that Entropy, Gradient, and at least one Product Form provide a set of at least three independent alignment estimates;

(3) Addition of bandwidth control to the likelihood term of an IT measure. Capability of controlling the scale of features in the JMS that are processed by an information measure allows optimization of response in the presence of differential noise. As JMS features are blurred by noise, the alignment estimator may be set to respond to only larger scale features that are still available despite the degradation. Continuity between search points over a search grid is also augmented by blurring the JMS input to the likelihood term; and

(4) design of a density estimator suitable for irregularly spaced data values. The estimator is suitable for density estimation of data that is either deterministically spaced data taken at irregular intervals or randomly spaced data. The results are sampled with controlled aliasing onto a uniformly sampled domain with sampling interval independent of the data acquisition.
6.1 Future Improvements

Below are suggestions of future research in Fisher Information-based co-registration. Despite good initial results, much is to be learned about the properties of the new measures. Their form is also considered preliminary, rather than ultimate.

6.1.1 JMS Density Estimation

Even though density estimation methods better than histogram bin counting have been touted in this dissertation, the JMS density estimator implemented is just a 2D histogram. The methods in Appendix D can be extended to two dimensions. The effect of aliasing in the JMS probability density estimate upon the IT measures is unknown, and may be detrimental, therefore the resulting anti-aliased density estimates should more accurately depict the real JMS structure for a pair of related and co-aligned images.

An better method yet of estimating densities in 1D and 2D would be based on Extreme Physical Information (EPI) championed by Frieden [64]. This methodology involves minimizing Fisher Information within constraints of measured data such that the JMS estimates would possess minimal roughness not supported by the data.
6.1.2 Fusion of Results Presented

Each of the six divergence-based co-alignment algorithms excels with some data. Furthermore, failure of all six measures occurs much less frequently than that of any one. By estimating the quality of each individual result, a consensus on the correct result may be formed. A framework that allows fusion of multiple results that have attributed to them an intermediate level of belief in viability is evidence theory. A good overview of Dezert-Smarandache theory, the most complete evidence theory, is given by Dezert in [79]. This framework allows for belief levels on individual results as well as subsets of the pool of results that mutually corroborate or conflict with each other.

6.1.3 Further Experimentation with New Fisher Information Variations

Choice of the Fisher Information parameter (differentiation is with respect to it) should be considered as a way to generate more powerful measures.

Just as there are algebraic variations on Entropy that put its constituent density functions in polynomial terms, density functions in Fisher Information may be modified similarly. This may improve reliability and precision for some values of the algebraic parameters.
6.1.4 Hierarchical Resolution Processing

Iterative co-alignment estimation by first computing a coarse-search estimate on a down-sampled image, then refining it on progressively higher resolution versions, should work with the divergence measure methods described herein. Computational time could be saved primarily by not having to compute the divergence over a large range of translations with a fine grid interval. To save time using coarse searches, the bin spacing would have to be increased. This is allowable, as the goal of the coarse co-registration is to find an approximate result, therefore unsharp responses are acceptable. High Sigma values will help maintain continuity between the sampled points. Hierarchical processing has been applied successfully to Entropic registration methods [19], [69], [80].
Appendix A

GENERATION OF SYNTHETIC TEST IMAGES

Appendix A describes the rationale and algorithm for generating synthetic images that were used in the dissertation.

A.1 Rationale

The goal of generating synthetic image pairs is control over both their spatial and statistical properties. Quantitative performance assessment of co-registration algorithms requires knowledge of correct alignment and probabilistic makeup of the process that generates the images, neither of which is available in real data. An overview of the processing sequence is presented in flow chart of Figure A-1, and its nomenclature is summarized in Table A-1.
Figure A-1  Image synthesis flow diagram.
A.2 Procedure

**Step 1 - Object creation and localization.** A random base image of three times the size in x and y as the desired output images is made up of independent, identically distributed (i.i.d.) random pixels chosen from a uniform distribution. A Gaussian convolutional filter is used to correlate adjacent pixel values. The result is segmented according to levels, producing a class map with regions classified according to level. The number of classes to be generated is a function of user specified control parameters as will be explained below. Levels are set to produce approximately equal numbers of pixels per class. Unequal level spacings are used, as the distribution of blurred image pixels, due to convolution with the Gaussian, is no longer uniform. An example of a class map is given in Figure A-2. Numerous discontiguous regions are formed for each class. The geometry of these regions specifies the geometry of simulated objects. Regions in a particular class will now be referred to as objects in that class.

**Figure A-2  Objects in class map - four level segmentation.** Spatial blur set to 2.5 standard deviations to demonstrate relatively large objects. Target image size is 64x64.
Step 2 - Assign textures to objects. Using the above class map as a template, two images are formed that depict imaging of the same objects with perfect co-alignment. For each region class in each image, a unique random texture is assigned. In either one of the images, all objects that are in the same class are assigned the same texture, but an independent texture is assigned to those same objects in the other image. An image pair then has a number of assigned textures equal to the number of classes times two. These textures are limited to two kinds: either (a) i.i.d. standardized Gaussian (zero mean; unit variance) distributed pixels randomly chosen from a standardized Gaussian distribution and offset by a random constant from a standard Gaussian distribution, or (b) a random constant chosen from a standard Gaussian distribution. The number of possible combinations is then four for each class in an image pair. The actual number of each kind of combination assigned is determined and specified as a control parameter prior to generating the image pair therefore it is not random. The total number of classes to be generated is then the sum of the numbers of all texture combinations specified.

Three object class types are defined according to how one of four texture combinations is assigned to an object: Type I - objects of this type in each image are filled with a constant, randomly chosen from a standard Gaussian random variable. This constant is independently chosen for each object class in each image; Type II - in one image, the texture of an object belonging to a Type II class is a constant, randomly chosen from a standard Gaussian random variable, and in the other image, the texture consist of i.i.d pixel values chosen from a standard Gaussian random variable offset by a constant chosen from a standard...
Gaussian random variable. The texture combination with an exchanged texture assignment from that above also makes an object Type II; Type III - i.i.d pixel values chosen from a standard Gaussian random variable offset by a constant chosen from a standard Gaussian random variable are assigned independently to a Type III object of a given class in each image.

**Step 3 - Simulation of imaging sensor.** To simulate the blurring that occurs in the sensor, the synthetic images are both convolved with the same assumed Gaussian PSF. Simulation of sampling is done by a three fold decimation in x and y with pixel average interpolation. The sampling phase is the same for both images. Sensor noise is then simulated by adding i.i.d. Gaussian noise. The pixel signal to noise ratio (SNR) is defined as the ratio of synthetic image standard deviation after blurring, to standard deviation of noise added. This emulates the SNR of an imaging detector, i.e., the blurred image prior to downsampling is considered entirely signal. It is precisely set for each image by computation of the blurred image standard deviation and adding noise of known standard deviation in controlled proportion.
<table>
<thead>
<tr>
<th><strong>Output Images</strong></th>
<th>Pair of images to be used for alignment estimator testing.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Base Image</strong></td>
<td>i.i.d. pixel image to be processed into object shapes through blurring and amplitude segmentation.</td>
</tr>
<tr>
<td><strong>Segmentation Levels</strong></td>
<td>Set of amplitude levels classifying blurred Base Image according to amplitude ranges.</td>
</tr>
<tr>
<td><strong>Classes</strong></td>
<td>Names given to amplitude intervals used in the segmentation and Regions or Objects distinguished by these intervals.</td>
</tr>
<tr>
<td><strong>Regions</strong></td>
<td>Spatial shapes resulting from amplitude segmentation of blurred Base Image into classes. To each class belong numerous discontiguous Regions.</td>
</tr>
<tr>
<td><strong>Segment Map</strong></td>
<td>Spatial map indicating Regions and their Class membership.</td>
</tr>
<tr>
<td><strong>Objects</strong></td>
<td>These are geometric shapes, that of Regions, emulating real objects. They retain class membership of regions, and they are filled with textures.</td>
</tr>
<tr>
<td><strong>Texture</strong></td>
<td>Statistical properties of random pixel values assigned to fill an object. There are two kinds of texture.</td>
</tr>
<tr>
<td><strong>Texture Combinations</strong></td>
<td>The pairing of Textures assigned to an object in each of the two synthetic images. There are four possible Texture Combinations.</td>
</tr>
<tr>
<td><strong>Object/Class Type</strong></td>
<td>Object/Class distinguished by the Texture Combinations assigned to Objects in each of the two images. There are three Object Types. A Class can have only objects of one type.</td>
</tr>
</tbody>
</table>

*Table A-1  Image Synthesis Nomenclature.*
The control parameters are summarized in Table A-2:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image size</td>
<td>Output image size - 2D vector.</td>
</tr>
<tr>
<td>Object scale</td>
<td>Controls correlation distance in base image - sets size of objects by setting Gaussian blur of base image.</td>
</tr>
<tr>
<td>Number of classes with Type I objects</td>
<td>Number of classes with objects containing constant value in both output images.</td>
</tr>
<tr>
<td>Number of classes with Type II objects</td>
<td>Number of classes with objects containing constant in one output image and i.i.d. random values in other output image.</td>
</tr>
<tr>
<td>Number of classes with Type III objects</td>
<td>Number of classes with objects containing i.i.d. random pixels in both output images.</td>
</tr>
<tr>
<td>PSF standard deviation</td>
<td>Gaussian PSF sigma in output image pixel Std. Dev. units.</td>
</tr>
<tr>
<td>SNR</td>
<td>Output image pixel SNR.</td>
</tr>
</tbody>
</table>

*Table A-2 Summary of Control Parameters.*

The algorithm is based on shapes emulating objects in a real scene. Because of the way they are generated, they are delineated by smooth random curves, and are not realistic with respect to shapes of most real objects, especially anthropogenic objects, as those are often bounded by straight lines or other simple geometric figures. Many of the simulated objects generated by this algorithm are entirely surrounded by one or more other objects. While this is not entirely unrealistic (example: lake surrounded by meadow, surrounded by forrest, surrounded by farmland), it occurs more frequently than in real imagery. Another
realism shortfall is lack of simulation of objects that have a non-stationary texture within their bounds, or have no distinct boundary between objects. The image model does not support such object complexities.

A.3 Examples

The following examples demonstrate the capability of the synthetic image generation scheme described above. All examples show generation of a 64x64 pixel image pair. In all of them, the Gaussian blurring used to generate the class template was set to one standard deviation at 2.5 pixels in output image pixel scale. Other parameters are described for each case. In Figure A-6 through Figure A-9, the same class template is generated to show effects of different noise and blurring levels on exactly the same base images. This is accomplished by forcing the same random number generator seed for each variation on the example. Gaussian blurring kernels are Gaussian functions truncated within square kernels of size 4.5 times the standard deviation, rounded up to nearest integer.

The type I and type III objects are more likely to represent real objects than a Type II. Real scenes are likely to consist of various mixtures of objects with properties that are intermediate between the three types generated in this work. Simulating many of each type and adding noise should bring some realism to the results.
These examples show single random realizations of a simulation. In application, unlimited numbers of image pairs may be generated for Monte Carlo simulation.
A.3.1 Two Type I Classes

Two constant levels in each output image, with no sensor blurring or noise (Figure A-3).

![Images](image.png)

**Figure A-3 Synthetic image pair: Two Type I classes.** (a) and (b) are output images with corresponding histograms in (d) and (e). (c) shows object template image. The JMS scattergram is shown in (f). The remaining figures of this type in this chapter show the same arrangement of elements.

Note that the scattergram shows the two pairs of constants, but also shows mixing due to the interpolation that takes place due to the pixel average interpolation in the factor of three downsampling.
A.3.2 Two Type II Classes

Two classes: one constant level texture and one i.i.d. random texture per output image, with exchanged texture types between images. No sensor blurring or noise modifies the generated images. In this case, the two images are independent at co-alignment, and become dependent upon mis-alignment. The independence condition is depicted in the scattergrams of Figure A-4(g) as circular, horizontal, or vertical features related to a common center. With a shift of one image by one pixel in x and y, the dependence is indicated in Figure A-4(h) as a bright dot at the center of the scattergram where the two constants coincide in the JMS.

This class emulates multi-modal sensing of objects that always produce a constant texture in one image and a random texture in the other.

A.3.3 Two Type III Classes.

With this type, statistical dependences between the images or even the regions are weak. It represents a high variance texture in all objects but with relatively little change in the means. It is useful as a most difficult test case, because it possesses the worst separability in the JMS of all the synthesized texture combinations considered herein. This exam-
Figure A-4  Two Type II classes. One object constant in one image, random, i.i.d. in other, and exchanged assignment for second object. (g) shows magnified scattergram with images in alignment. (h) shows effect of misalignment - bright point in center is indicative of dependence. The legend is in the caption for Figure A-3.
ple is a most realistic emulation for imagery containing objects that dominate the pixel value variation through texture rather than through changes in mean responses, for example forests with mixed species of vegetation in patches. (Figure A-5.)

Figure A-5 Two Type III classes. This represents imagery of objects all having a ‘noisy’ texture. The legend is in the caption for Figure A-3.
A.3.4 One Type I, Two Type II, and One Type III Class

Four classes are populated: one has a random constant in both images; one has a constant in first image, and i.i.d. values in the second image, one has i.i.d. values in first image and a constant in the other image, and one has i.i.d. values in both images. No blurring PSF is applied, and no ‘detector’ noise is added. (Figure A-6.)

**Figure A-6** Mixed class synthetic image pair. Four object classes are represented. One class is constant in both images, two have the mixed and exchanged texture as in A.3.2, and one i.i.d in both images. The legend is in the caption for Figure A-3.
As A.3.4 with noise - SNR = 7.0.

This example readily demonstrates the blurring effect of added random noise upon the features in the JMS. (Figure A-7.)

![Figure A-7](image)

**Figure A-7** Mixed classes at an SNR of 7.0. Same as Figure A-6 with i.i.d. noise added. The legend is in the caption for Figure A-3.
A.3.5 As A.3.4 with PSF Blurring.

Spatial blurring results in interpolation between un-blurred JMS features. (Figure A-8.)

Figure A-8  Mixed class image pair with blurring and noise. Figure A-6 data as presented in Figure A-7 with blurring. The legend is in the caption for Figure A-3.
A.3.6 All Elements of Synthesis Simultaneously.

This is a more comprehensive case in terms of the algorithm’s capabilities and is closest to real imagery. The blurred images of A.3.5 have noise added for a pixel SNR = 2.0. The JMS scattergram is visually void of features even though the noisy images still have visible objects in them. This demonstrates the discarding of spatial information that occurs in computation of the JMS density.

![Image](image_url)

Figure A-9  Mixed class image pair with blurring and noise. Figure A-6 data with blurring as presented in Figure A-7, and i.i.d. noise added to set the SNR at 2.0. The legend is in the caption for Figure A-3.
Appendix B

GENERATION OF DEGRADED REAL IMAGE PAIRS

B.1 Rationale

A methodology for obtaining images that have constrained alignment error, and controlled spatial and statistical properties, using real objects from real imagery, is described in this appendix.

Imagery from a multi-band sensor is usually available in a co-registered form. The numerical ‘pipeline’ used in standard processing of sensor data is usually tasked to correct systemic image mis-alignments such that the residual error is less than one pixel. If data from a high resolution sensor is downsampled such that the resulting image is on a sparser grid, the original mis-alignments remain the same on object scale, but are reduced relative to the output pixel interval by the downsampling factor. Therefore, such a relatively lower resolution image may be considered as aligned to subpixel precision. The downsampling factor may be chosen high enough that the result is practically in perfect co-alignment.
The change in scale allows control of other image characteristics. If the discrete representation in the high resolution image is treated as a spatially continuous data source, then the downampling can be treated as an original spatial sampling. To emulate the spatial response of square pixel detectors that perfectly adjoin each other, pixel averaging interpolation is applied in the downampling.

To emulate the spatial response of the image forming optic’s PSF, the high resolution image may be convolved with a blurring kernel, prior to downampling. Detector noise may also be emulated by adding appropriate noise to the downsampled image.

A disadvantage of using real imagery compared to purely synthetic imagery is the limitation on the number of different scenes, due to a finite supply of the high resolution images. The advantage is that real image spatial and spectral content is used, rather than an idealized approximation via image synthesis.

**B.2 Implementation**

The source of the data is high resolution (4 m) multi-spectral IKONOS imagery, from which two bands are used. To obtain the proper output image size, an appropriately large portion of the high resolution scene is subset. The location of the subsetting is controlled to allow choice of objects that are appropriate for good tests of the alignment estimators.
Exact positioning of the subsetting is randomized on a sub-pixel scale such that different sampling phases in the downsampling may be achieved. For one of the images in a particular scene, the number of different phases possible is the square of the downsampling factor, and they occur within +/- 0.5 pixel range of x, y. There are that many unique pixel centers available in the high resolution image over the area of one low resolution pixel.

The sampling phase of the first image (A) is always random. The positioning of the second image (B) may be set to be the same (in lock-step phase) or its phase may be set to be random with respect to the first image in a manner similar to the phasing of the first image.

*Figure B-1 Degraded image generation flow diagram.*
with respect to the scene. The random phase distribution for image $A$ is uniform over the $ +/- 0.5$ high-resolution-pixel box, and so is the relative phase difference between the images when not in lock-step. This makes the phase distribution of the second image sampling pyramidal over $ +/- 1.0$ pixels (with downsampling factor to the fourth power possible phase pairs) when it is made random with respect to the first image phase.

Blurring prior to the downsampling is accomplished by convolution with an isotropic Gaussian kernel. The standard deviation is set in output pixel units, therefore, the actual standard deviation is that value multiplied by the downsampling factor. The Gaussian function is truncated to a kernel size that is an integer value rounded up from 4.5 times the standard deviation.

Noise is added to the low resolution result to set a signal to noise ratio (SNR). The variance of each image is computed, and a pair of noise images is formed with a variance to give the desired SNR. The SNR is defined as the ratio of image standard deviation to added noise standard deviation. The noise source is i.i.d. with a Gaussian distribution.

Figure B-1 summarizes the degradation process in a flow diagram.
B.3 Example Utilizing IKONOS Data

This example uses an Ikonos agricultural scene of Maricopa, Arizona (Figure B-2). A partly vegetated scene was chosen with the goal of testing co-registration methods on an image pair that lacks linear correlation. Between the red and near infrared (NIR) bands, the responses to vegetative objects and bare soil objects are roughly orthogonal in the JMS. An appropriate object mix should exist in some image subset where the responses cancel in the second moment. To prevent the correlation coefficient from changing appreciably from zero during a spatial search for co-location, the best center to pick has a near-zero correlation over a neighborhood of several pixels in x and y. To locate such an appropriate subset of the scene, first, a window size, downsampling factor, and blurring level are selected for IT measure evaluation. In this case they were 127x127 pixels downsampled from Ikonos by a factor of 5. Blurring was with a Gaussian that has standard deviation = 0.5 output pixels.

Figure B-2 Original IKONOS image of Maricopa agricultural fields. Red band image of entire scene. Area used for simulating lower resolution is shown.
The entire Ikonos scene is first blurred and downsampled in the manner described above. A moving window of 127x127 pixels was passed over the lower resolution result, and the correlation coefficient was computed for all positions of the window. This provided the ‘correlation map’ shown in Figure B-3. The arrow in the figure indicates a location with near zero correlation that persists if the window is moved a few pixels in any direction.

*Figure B-3  Correlation map for IKONOS bands 3/4 image pair.* Arrow shows location of a 127x127 pixel window center to obtain near zero correlation at that location, and in its vicinity. Axes coordinates are of the upper right hand moving window corner. Colorbar indicates correlation color mapping.
A pair of 143x143 output pixel windows was cropped, with the center chosen per above, from the original image pair (Figure B-4). These source images were blurred, downsampled, and made noisier as described above. The result has a correlation coefficient of +0.023 in the central 127x127 portion. At a SNR of 4.0, and blurring of 0.5 pixel standard deviations, is the result shown in Figure B-5.

Figure B-4 Ikonos data prepared for downsampling and degrading. Left image shows red band, and right image shows near infrared band.
Figure B-5  Example: Downsampling and degrading IKONOS image set. Left image shows red band; right image shows near infrared band. Images in Figure B-4 are blurred with 0.5 output pixel standard deviation Gaussian, downsampling by factor of
Appendix C

PROOF: EQUALITY OF FISHER INFORMATION FORMS

Fisher Information is commonly presented in two equivalent forms. This proof is derived in this dissertation to show that under condition in Eq. (C.1), the forms presented in Eq. (C.2) are interchangeable.

Theorem.

If probability density function \( p(x|\theta) \) is doubly differentiable and integrable such that

\[
\frac{\partial^2}{\partial \theta^2} \int p(x|\theta)dx = \int \frac{\partial^2}{\partial \theta^2} p(x|\theta)dx \tag{C.1}
\]

then

\[
\int p(x|\theta) \left[\frac{\partial}{\partial \theta} \log p(x|\theta)\right]^2 dx = -\int p(x|\theta) \frac{\partial^2}{\partial \theta^2} (\log p(x|\theta)) dx . \tag{C.2}
\]
Proof.

Starting with the right hand side:

\[
-\int p(x|\theta) \cdot \frac{\partial^2}{\partial \theta} \log p(x|\theta) dx = -\int p(x|\theta) \cdot \frac{\partial}{\partial \theta} \left( \frac{\partial}{\partial \theta} \log p(x|\theta) \right) dx \quad (C.3)
\]

\[
= -\int p(x|\theta) \frac{\partial}{\partial \theta} \left( \frac{1}{p(x|\theta)} \cdot \frac{\partial}{\partial \theta} p(x|\theta) \right) dx \quad (C.4)
\]

\[
= -\int p(x|\theta) \left( \frac{\partial}{\partial \theta} \frac{1}{p(x|\theta)} \cdot \frac{\partial}{\partial \theta} p(x|\theta) + \frac{\partial^2}{\partial \theta^2} p(x|\theta) \cdot \frac{1}{p(x|\theta)} \right) dx \quad (C.5)
\]

\[
= -\int p(x|\theta) \frac{\partial}{\partial \theta} \frac{1}{p(x|\theta)} \cdot \frac{\partial}{\partial \theta} p(x|\theta) dx - \int p(x|\theta) \left( \frac{\partial^2}{\partial \theta^2} p(x|\theta) \cdot \frac{1}{p(x|\theta)} \right) dx . \quad (C.6)
\]

First prove that the second integral is zero:

\[
\int p(x|\theta) \left( \frac{\partial^2}{\partial \theta^2} p(x|\theta) \cdot \frac{1}{p(x|\theta)} \right) dx = \int \frac{\partial^2}{\partial \theta^2} p(x|\theta) dx \quad (C.7)
\]
The first integral is proven to be the left hand side of assertion:

\[ \frac{\partial^2}{\partial \theta^2} \int p(x|\theta) \, dx = 0 . \]  

\[ (C.8) \]

\[-\int p(x|\theta) \left( \frac{\partial}{\partial \theta} \frac{1}{p(x|\theta)} \cdot \frac{\partial}{\partial \theta} p(x|\theta) \right) \, dx \]

\[ (C.9) \]

\[ = -\int p(x|\theta) \left( \frac{-1}{(p(x|\theta))^2} \frac{\partial}{\partial \theta} p(x|\theta) \cdot \frac{\partial}{\partial \theta} p(x|\theta) \right) \, dx \]

\[ (C.10) \]

\[ = \int p(x|\theta) \left( \frac{1}{p(x|\theta) \partial \theta} p(x|\theta) \right)^2 \, dx \]

\[ (C.11) \]

\[ = \int p(x|\theta) \left( \frac{\partial}{\partial \theta} \log p(x|\theta) \right)^2 \, dx \]

\[ (C.12) \]
Appendix D

DENSITY ESTIMATION AND HISTOGRAM MATCHING FOR CORRECTION OF INTERPOLATION ARTIFACTS

D.1 Overview

Ultimately, the procedure of shifting information in an image so that its objects line up with the same objects in a second image of same scale but different pointing would result in a new image with identical spatial and radiometric properties. Barring detector noise, the image information would be identical to that in the first image, as if the first image was acquired with the sensor already pointed exactly the same way as the one that took the second image. Achieving this goal through co-registration is not possible because interpolation functions directed by spatial transformations are capable of repositioning spatial information, but they alter radiometric statistics. The alteration takes place because output pixel values are computed by functions that combine values from several source pixels. Nearest Neighbor interpolation nearly preserves the global radiometric information, but introduces spatial jitter of +/- 1/2 pixel for each pixel. This jitter is generally not random,
as it is determined by interaction of a uniform pixel grid against one that is slightly dis-
torted by a spatial transformation. If a sub-pixel spatial re-positioning precision is desired, a resampling scheme with continuous spatial response to target location is required.

Simultaneous preservation of spatial and radiometric fidelity under spatial transformation appears to be an unsolved problem. The procedure described in this appendix is an attempt at optimization of the spatial/radiometric trade-off.

**D.2 Purpose**

In this dissertation, the control of spatial parameters of an image (such as object position with respect to a reference frame) is only accomplished by resampling of an interpolated original image rather than real-time control of cameras. Information Theoretic (IT) measures are intentionally sensitive to statistics of an image, and are therefore subject to radiometric artifacts caused by the interpolation. The artifacts are not a spatially stationary transformation of the image statistics, but rather vary substantially according to where the interpolation takes place with respect to the actual input data values. Details of the effect are described by Pluim [78].
The approach described here to mitigate the artifacts is to estimate the probability density of an image prior to any radiometric processing, then restore the image density to the estimate after spatial and statistical alteration.

**D.3 Theory**

**D.3.1 Density Estimation**

The objective is to estimate the global probability density of data from a single sample of a random process - in this dissertation, a single image.

A common and perhaps simplest method is histogram binning and counting. It can be easily demonstrated that substantial aliasing occurs due to the discrete bin ranges, for example, by estimating the density of a sample with a complex density and then repeating after slightly shifting or scaling the bin ranges and comparing the results. The optimal setting of the histogram bins is data dependent. In this dissertation, prior assumptions about the data probability densities are not made, therefore histogram methods are avoided.

The next most popular family of density estimation methods is collectively called the ‘Kernel Method’ or ‘Parzen Method’ [81]. In its simplest form, an ordered list of data values is made, and at each data point location, a fixed kernel function is centered. By super-
imposing all such kernel responses for all data point values, and normalizing the sum to unit area, a data density consisting of a discrete set of values (one per data point) may be transformed into a reasonable, continuous estimate of the density. Most commonly, the kernel function is bell shaped and symmetric about its center.

This method is also called non-parametric because there is no strict connection between the parameters of the actual probability distribution of the data and the operation of this estimator. One does, however, choose the shape and width of the kernel function, so some information about the probability density granularity to be estimated has to be known. A too narrow kernel will result in numerous peaks with little overlap that bear little resemblance to the shape of the probability distribution. This is a problem if the data density has more modes than the true probability distribution underlying the imaging process and the actual simple distribution shape will be misrepresented. A too wide kernel will result in substantial overlapping and thus blurring of a complex density shape. Even an optimized kernel often gives poor results because the data points are much more dense near a density mode, and can be extremely sparse in the tails.

Two modifications yield more agreeable results [82]: (1) A transformation of the amplitude ordinate to compress the density tails results in a more uniform spacing of the data points. In fact, if the actual density is substantially known, an equalization of the data point spacing may be accomplished to aid estimation by these kernel methods; (2) The kernel function may be adjusted for varying data density by varying its width and
inversely its height to maintain the same area. This preserves the weight of each point, but
uses a smoothing in proportion to data interval spacing. This may be done adaptively such
that different parts of the amplitude scale possessing different data point spacings may
have the density estimate optimally smoothed. The method does not adequately deal with
cases where the data interval to adjacent points is substantially different on the two sides
of a data point because one symmetric kernel can not be optimal for both intervals.

The method used in this project is a further evolution of approach (2) presented above. It
adaptively adjusts the kernel by conforming its width to each amplitude interval between
points rather than attempting to optimize it at each data point. A rectangular kernel func-
tion of width equal to the interval and height equal to its reciprocal is used to model the
density between measured data points. No knowledge of the true data density is assumed
at all, making the algorithm entirely data driven.

All published methods attributed above to parzen and devroye have been shown to be con-
sistent. It is yet to be proven that the new method is also consistent.

D.3.2 Anti-Aliasing

So far in the discussion, the result of a probability density estimate is a continuous func-
tion. If a probability density estimate is to be represented in an array of sample values,
there is a likelihood of aliasing of fine detail in the data density. The underlying probabil-
ity density of pixel values may be made up of shapes sharp enough to have frequency content in the amplitude domain that would be aliased. An actual single image from which the probability density is to be recovered is a random sample with a data density shape that contains stochastic variation upon the probability density shape and therefore will have frequency content that will be aliased. Bandwidth in such a sample is generally unlimited because the finite sample is represented on an amplitude continuum as a series of impulses (counts).

By estimating the density at finer amplitude sampling steps, one increases the range of frequencies in the fundamental range and therefore the un-aliased information. This approach is limited by computational resources preventing it from being a thorough solution, and there is no limit to how fine the sample steps should be to prevent aliasing.

Another approach is to smooth the data distribution from one image by adding amplitude noise to the image. On average, the distribution of the sum is the image distribution convolved with the distribution of the added noise. In one trial, adding noise does not smooth the distribution. However, by repeating noise addition several times with independent noise samples to copies of the same image, and processing all of the results as one larger concatenated data sample, a smoothing takes place. This increase in data size also increases processing load. The smoothing can make aliasing negligible, but there is no limit on the number of noise samples needed to make it disappear.
D.3.3 Density Matching

To then realize the estimated density with the interpolated data, two approaches have been considered: 1) Histogram matching such as is commonly implemented in image processing software. This method alters or corrects density values by keeping the same bin counts, but it adjusts the bin values such that, averaged over some relatively small range of bin values, the density is adjusted to a target density. This alteration of values does not effect spatial positions of data values, but the values themselves. The shape of the resulting distribution is altered on a small amplitude scale. The data under amplitude transformation is either discrete sampled, or it is a discrete set due to the finite number of pixels in the image. The resulting distribution after processing is also discrete valued, but with a non-uniform sampling interval and potentially large steps in some parts of the amplitude range. These are manifested as zero bin values on a uniformly spaced binning; 2) Another way to accomplish density matching is by adjusting amplitude values across all the data points independently, such that a best approximation to the target density is achieved. Approximation is necessary because of two factors. The desired probability density usually has values on a continuous domain, yet the resulting data density is accomplished with a finite integer number of data points, i.e. N total pixels in the image, making it impossible to exactly realize any of the fractional target density values. Once a set of N values to be assigned is established, locations are to be assigned to them in such order as to best preserve the spatial information of the original image. The result is in effect noise added to the original image information because the amplitude transformation is not
exactly the same for each pixel. For example, if an image has some number of pixels with
the same value, but the number of new values to assign for them is different, then the rela-
tion of values within this pixel set is not preserved because they will receive different val-
ues.
D.4 Implementation for this Dissertation

In this work, the actual density function estimation is computed in a cumulative distribution function (CDF) equivalent to the method that operates on density functions directly described above in D.3.1. First, the CDF of the image data is computed. Then the adap-
tive aspect of kernel selection is achieved by linear interpolation between the midpoints of each CDF step. This forms a continuous piecewise linear CDF estimate.

Figure D-2 Diagram: Density function estimation based on cumulative data distribution. This example, with three data points, shows conversion from discrete data, to a continuous cumulative distribution accruing at data point masses. The estimated cumulative distribution is a linear interpolation of data distribution step midpoints. This is equivalent to uniform densities between the data points with values inversely proportional to the corresponding data interval lengths.
To obtain the density estimate, this CDF estimate is differentiated by means of a first order difference at discrete and uniformly spaced sample points chosen for the discrete density output array with four-times oversampling. A diagrammatic example of conversion from discrete data points to a continuous CDF and discontinuous density is shown in Figure D-2.

D.4.2 Anti-aliasing

To reduce the aliasing problem, the sampling interval of the 1D estimator described so far is four times finer than that of the final output. The density estimate abscissa (output) points are chosen to be the same as the 2D histogram bin centers of the JMS estimator, thus their interval is exactly four times the JMS bin spacing. The decimated output is obtained by low pass filtering the finer resolution estimate through convolution with a non-negative kernel. In particular, a 30 dB Chebychev window is used for the filtering prior to downsampling. The window width is adjusted to obtain an approximate 30 dB rolloff down the main lobe at the Nyquist frequency.

Anti-aliasing for the initial sampling is achieved by taking four copies of the input image data, and to each, an independent sample of white gaussian noise is added, then the estimator processes the four images as one random process realization. The level of noise is
adjusted such that amplitude spectrum components at or above the Nyquist frequency are attenuated to 0.2 or less of their initial level. The filtering and downsampling of a small data set is shown in example of Figure D-3

Figure D-3 Density estimation of random data. The graph shows the estimated data cumulative distributions at four times target amplitude sample interval for bottoms and tops of the intervals (magenta and blue lines), and the resulting density estimate (green line). The black plot shows the anti-alias filtered density that is four times downsamped to target sample interval. Scale for the cumulative distribution is [0,1], and a rel-
D.4.3 Density Matching

Given that the desired density, specified at K points is to be replicated and the input and output images each have N pixels, N new values need to be computed such that they best meet the desired density. A simple approximation is to take the target density (which sums to one) and multiply each of its K values by N, to obtain the number of pixels to be assigned for each of the desired density’s ordinate values. This results in fractional values that are rounded to integer counts that may not exactly sum up to N. The actual algorithm slightly adjusts the scaling of the target cumulative distribution up or down from one, obtaining exactly N total new pixel values.

Then these values are assigned to pixels in a way that approximately reconstructs the spatial information in the original image as follows. The pixel values of the original image are sorted to determine their rank order. Then the replacement values are assigned to specific output image pixel locations, such that original pixel rank order is preserved.

A problem arises if the input image has some pixels with equal values as is often the case with quantized images, because the ranking process is ambiguous with respect to the spatial order of tied values. A secondary decision process determines ordering of new pixel values in a tie. 3x3 neighborhood averages are computed as secondary sorting variables around each original image pixel to determine ordering within the tied values. 5x5 neigh...
borhood averages are used to break ties not resolved in the 3x3 neighborhoods. The new pixel values do not require tie breaking because their order carries no spatial information - they are just points in a distribution.

D.5 Examples

In both examples, a test images undergoes a bilinear interpolation for a shift of 0.5 pixels in both x and y directions. The probability density of the original image is estimated, and the density of the interpolation result is matched to it.

D.5.1 White Noise Image

A 192x192 image is generated with i.i.d. standard gaussian pixels. In Figure D-4, the original density is shown in black. After resampling, the density becomes smoother and has lower variance around the original mean due to the central tendency of the averaging process in bilinear interpolation (magenta). After density matching, the result has a density closely approximated to the original (blue). The resulting density appears smoother than the original because the anti-aliasing smooths the estimate. There are two estimates applied sequentially: the original density estimate used for reconstruction, and the one used to evaluate the reconstruction.
Similarly, a 64x64 Ikonos red band image subset of agricultural crops is processed.

Shown in Figure D-5 (a), the original density is widely distributed (black), but becomes concentrated into two sharp peaks (magenta) upon resampling. The density is closely corrected to the original shape as shown by the blue line. Grayscale examples are min-max stretched. The original image is shown in (b). The resampled version in (c) appears sub-

**Figure D-4 Density matching example: white noise image.** Black line is density estimate of original 192x192 white Gaussian noise image. After resampling at a 0.5 pixel shift in x, y, the density is more central as shown by magenta line. Corrected density is shown by blue line that nearly overlays the black line. The density is sampled in 0.05 standard deviation steps.

### D.5.2 Ikonos Image

Similarly, a 64x64 Ikonos red band image subset of agricultural crops is processed. Shown in Figure D-5 (a), the original density is widely distributed (black), but becomes concentrated into two sharp peaks (magenta) upon resampling. The density is closely corrected to the original shape as shown by the blue line. Grayscale examples are min-max stretched. The original image is shown in (b). The resampled version in (c) appears sub-
stantially blurred. This effect is analyzed in depth by Park and Schowengerdt [83]. After correction (d), the spatial structure remains even though the global pixel density has been re-distributed to closely approximate the original density.

**D.6 Conclusion**

The result of these methods is an approximation to the goal of restoring the radiometric properties of an image that underwent interpolation and resampling. The imprecision in attaining the desired density, and the resulting noise in the spatial and radiometric information may be detrimental to IT measure-based estimator precision, even if the primary goal of erasing statistical artifacts is achieved. Whether the replacement of radiometric artifacts with radiometric imprecision improves the quality of IT measure results partially depends on the image size. For increasing image size, the method described in this appendix will asymptotically become accurate and precise.

The visual appearance of images processed to correct for resampling is excellent due to the re-distribution of pixel values away from the neighborhood mean values towards which the interpolation modified them. The noise that is effectively created by assignment of pixel values not in exact proportion to that in the original image is not visually identifiable as noise.
Figure D-5  Density matching example: IKONOS image. The original image is shown in (b), and its density in (a) black line. After translation in x, y by 0.5 pixels using bilinear resampling, it is shown in (c), and its density in (a) magenta line. Application of density correction renders (d), with density in (a) blue line.
REFERENCES


