A Hybrid Similarity Measure Framework for Multimodal Medical Image Registration

Thesis

How to cite:


For guidance on citations see FAQs.

© 2016 Parminder Reel

Version: Version of Record

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online’s data policy on reuse of materials please consult the policies page.
A Hybrid Similarity Measure Framework for Multimodal Medical Image Registration

Parminder Singh Reel

B.Eng, M.Sc

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

Department of Computing and Communications

Faculty of Mathematics, Computing and Technology

The Open University

Milton Keynes, United Kingdom

Submitted on 28th August 2015
I dedicate this work to …

My Parents, and I hope that this achievement will complete the dream that you had for me all those many years ago when you chose to give me the best education you could.

My Wife, whose loyalty to our best interests and willing sacrifices have been such great factors in the completion of this work. Without her patience, understanding, support and most of all love, this work would have not been possible.

My Sister, for her love, care and support at all times.
Acknowledgements

Looking back at the PhD journey, I have to agree with my supervisor, Prof. Laurence S. Dooley, who at the beginning told me that what I have started, would be at times both an intellectually and emotionally challenging journey, that would indeed require a lot of personal motivation on my part. I am deeply grateful for his endless support, guidance and wisdom. His constant academic and moral support gave me the confidence to succeed which I needed during the various highs and lows in this journey. I would also like to express my sincere gratitude to Dr. Patrick Wong for always providing me very useful comments and excellent guidance. Discussions with him have always helped me to improve the quality of my work. I am also thankful to my external supervisor Dr.-Ing. Anko Börner at German Aerospace Center (DLR), Berlin for his constant support, encouragement and insightful discussions despite his busy schedule. I especially thank him for the hospitality and personal effort in accommodating me at DLR during my three annual visits. I will forever be grateful, and could not have completed this thesis without you all.

I am thankful to my mini viva examiners Dr. David Chapman and Dr. Allan Jones for their suggestions and feedback during the early stages of this thesis work. I also owe my thanks to Dr. Asanobu KITAMOTO at National Institute of Informatics (NII), Tokyo, for giving me the opportunity to work in his laboratory during my internship at NII. I also wish to thank Dr. Adrian S. Poulton and Dr. Soraya Kouadri Mostéfaoui for their valuable feedback, as well as all the current and alumni members of neXt Generation Multimedia Technologies (XGMT) research group for their supportive and interesting technical discussions. I am also grateful to Hartwig Hetzheim, Eugen Funk and other colleagues at DLR for providing me valuable ideas as well as a friendly environment to work. I am also thankful to Thomas Fähnsing for his support and efforts during my visits to Berlin. I would also like to thank all the co-travellers of doctoral journey in Room S1005, for providing me a friendly as well as focused environment. I am thankful to Advait, Alam, Mobolaji and Faisal for taking out time from their busy schedules to go through my thesis draft.

Finally, I would like to thank The Open University, Research Degrees Office, Faculty of MCT and its Computing and Communications Department for offering me the funding and an opportunity to pursue this PhD study.
Abstract

Medical imaging is widely used today to facilitate both disease diagnosis and treatment planning practice, with a key prerequisite being the systematic process of medical image registration (MIR) to align either mono or multimodal images of different anatomical parts of the human body. MIR utilises a similarity measure (SM) to quantify the level of spatial alignment and is particularly demanding due to the presence of inherent modality characteristics like intensity non-uniformities (INU) in magnetic resonance images and large homogeneous non-vascular regions in retinal images. While various intensity and feature-based SMs exist for MIR, mutual information (MI) has become established because of its computational efficiency and ability to register multimodal images. It is however, very sensitive to interpolation artefacts in the presence of INU with noise and can be compromised when overlapping areas are small. Recently MI-based hybrid variants which combine regional features with intensity have emerged, though these incur high dimensionality and large computational overheads.

To address these challenges and secure accurate, efficient and robust registration of images containing high INU, noise and large homogeneous regions, this thesis presents a new hybrid SM framework for 2D multimodal rigid MIR. The framework consistently provides superior quantitative and qualitative performance, while offering a uniquely flexible design trade-off between registration accuracy and computational time. It makes three significant technical contributions to the field: i) An expectation maximisation-based principal component analysis with mutual information (EMPCA-MI) framework incorporating neighbourhood feature information; ii) Two innovative enhancements to reduce information redundancy and improve MI computational efficiency; and iii) an adaptive algorithm to select the most significant principal components for feature selection.

The thesis findings conclusively confirm the hybrid SM framework offers an accurate and robust 2D registration solution for challenging multimodal medical imaging datasets, while its inherent flexibility means it can also be extended to the 3D registration domain.
Declaration

The work presented in this thesis is an original contribution of the author. Parts of this thesis have appeared in the following:

**Peer-Reviewed Publications:**


**Research Prize:**

*IET Engineering Prize Award* sponsored by ABB Group at IET Midlands Power Network Engineering Prize Award Evening Event (in association with industry), Coventry University, UK, Mar 18, 2014. doi: 10.6084/m9.figshare.1004109

Parminder Singh Reel

Submitted on 28th August 2015
Contents

Contents vi

List of Figures xii

List of Tables xxviii

List of Abbreviations xxix

List of Symbols xxxi

1 Introduction 1

1.1 Overview .................................................. 1

1.2 Mutual Information-based Multimodal MIR .................. 4

1.3 Research Motivation ....................................... 6

1.4 Research Question and Objectives .......................... 9

1.5 Contributions .............................................. 12

1.6 Thesis Structure ........................................... 14

1.7 Summary .................................................. 16
## Contents

### 2 Medical Image Registration: A Review

- **2.1** Introduction .......................................................... 17
- **2.2** Brief History of Medical Imaging .............................. 18
- **2.3** Medical Image Registration ...................................... 25
  - **2.3.1** Transformation .................................................. 26
  - **2.3.2** Interpolation .................................................... 29
  - **2.3.3** Similarity Measure .............................................. 32
  - **2.3.4** Optimisation .................................................... 37
- **2.4** MI and its variants for multimodal MIR ...................... 41
- **2.5** Discussion .............................................................. 54
- **2.6** Summary ............................................................... 57

### 3 Research Methodology

- **3.1** Introduction .......................................................... 58
- **3.2** Research Methodology and Test-bed ......................... 59
- **3.3** Generic MIR Model Setup ........................................ 61
- **3.4** Image Datasets ...................................................... 63
  - **3.4.1** Mono and Multimodal BrainWeb Dataset .................. 63
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.4.2</td>
<td>Mono and Multimodal Retinal Datasets</td>
<td>64</td>
</tr>
<tr>
<td>3.5</td>
<td>Robustness Scenarios</td>
<td>67</td>
</tr>
<tr>
<td>3.5.1</td>
<td>Initial Misregistration</td>
<td>68</td>
</tr>
<tr>
<td>3.5.2</td>
<td>INU and Gaussian Noise</td>
<td>69</td>
</tr>
<tr>
<td>3.5.3</td>
<td>Low Contrast and Homogeneous Regions</td>
<td>71</td>
</tr>
<tr>
<td>3.6</td>
<td>Performance Metrics</td>
<td>73</td>
</tr>
<tr>
<td>3.6.1</td>
<td>Quantitative Assessment</td>
<td>73</td>
</tr>
<tr>
<td>3.6.2</td>
<td>Perceptual Assessment</td>
<td>74</td>
</tr>
<tr>
<td>3.6.3</td>
<td>Computational Cost</td>
<td>75</td>
</tr>
<tr>
<td>3.7</td>
<td>Simulation Platform</td>
<td>76</td>
</tr>
<tr>
<td>3.8</td>
<td>Software Code Validation</td>
<td>77</td>
</tr>
<tr>
<td>3.9</td>
<td>Summary</td>
<td>80</td>
</tr>
<tr>
<td>4</td>
<td>EMPCA-MI for Medical Image Registration</td>
<td>81</td>
</tr>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>81</td>
</tr>
<tr>
<td>4.2</td>
<td>EMPCA-MI</td>
<td>83</td>
</tr>
<tr>
<td>4.2.1</td>
<td>Step I: Image Neighbourhood Rearrangement</td>
<td>83</td>
</tr>
<tr>
<td>4.2.2</td>
<td>Step II: EMPCA-based feature extraction</td>
<td>86</td>
</tr>
</tbody>
</table>
4.2.3 Step III: Finding MI for Registration ................. 88

4.3 Choosing interpolation and neighbourhood radius for EMPCA-MI . 93

4.3.1 Choice of Interpolation Technique ...................... 94

4.3.2 Choice of neighbourhood radius ....................... 102

4.4 Summary .................................................. 111

5 EMPCA-MI-based Registration Results 113

5.1 Introduction ............................................. 113

5.2 Experimental Setup ...................................... 114

5.3 Experimental Results Discussion ....................... 116

5.3.1 Set 1: BrainWeb Dataset-based Experiments .......... 116

5.3.2 Set 2: DRIVE and Brno-Multi Dataset Experiments ..... 127

5.3.3 EMPCA-MI Framework Robustness .................... 137

5.4 Summary .................................................. 141

6 EMPCA-MI Framework Enhancements 142

6.1 Introduction ............................................. 142

6.2 EMPCA-MI Framework Enhancement 1 .................... 144

6.2.1 Neighbourhood Region Selection ..................... 146
6.2.2 Experiment Setup ........................................ 147
6.2.3 Experimental Results Discussion ..................... 148

6.3 EMPCA-MI Framework Enhancement 2 ...................... 154
6.3.1 Number of bins for MI computation .................... 155
6.3.2 Experimental Setup ...................................... 159
6.3.3 Experimental Results and Discussion .................. 159

6.4 Summary .................................................. 164

7 Adaptive EMPCA-MI Framework ............................. 166
7.1 Introduction ................................................ 166
7.2 Adaptive EMPCA-MI framework ............................ 167
7.2.1 Stopping Rules for PCA ................................ 168
7.2.2 Adaptively choosing $P$ in the EMPCA-MI framework . 169

7.3 Experimental Setup ........................................ 170

7.4 Experimental Results Discussion ........................ 172
7.4.1 Set 1: BrainWeb Dataset Registration Experiments . 172
7.4.2 Set 2: DRIVE and Brno-Multi Dataset Experiments . 178
7.4.3 Computational Cost of Adaptive EMPCA-MI Framework 185
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Layout of the framework objectives in this thesis</td>
<td>10</td>
</tr>
<tr>
<td>2.1</td>
<td>Brief historical timeline of medical imaging</td>
<td>19</td>
</tr>
<tr>
<td>2.2</td>
<td>Showing (a) the first published human Fundus photography apparatus (T. Jackman and J.D. Webster, 1886), (b) the first Fundus retinal image taken from it, with top white area as the optic disc and the lower, light area as an artefact (Saine, 1993), (c) the first ever X-ray image of A. B. Röntgen’s hand taken in 1895 (Kevles, 1997).</td>
<td>20</td>
</tr>
<tr>
<td>2.3</td>
<td>Showing (a) the first fluorescein angiogram (H. R. Novotny and D. L. Alvis, 1961) and (b) the first brain CT scan in 1971 with 80 × 80 pixels resolution (Beckmann, 2006).</td>
<td>21</td>
</tr>
<tr>
<td>2.4</td>
<td>Various medical imaging modalities for (a &amp; b) human brain, (c &amp; d) eye retina from (Montreal Neurological Institute and Hospital, nd) and (Kolar et al., 2013; Kolar and Tasevsky, 2010).</td>
<td>22</td>
</tr>
<tr>
<td>2.5</td>
<td>Generic MIR model.</td>
<td>27</td>
</tr>
<tr>
<td>2.6</td>
<td>Rigid registration followed by non-rigid registration.</td>
<td>29</td>
</tr>
</tbody>
</table>
2.7 (a) Nearest neighbour, (b) Bilinear and (c) Bicubic Interpolation Methods. ................................................................. 31

2.8 Bifurcation structure comprising of a master bifurcation point and its three connected neighbours. Figure reproduced from (Chen et al., 2015). ................................................................. 36

3.1 Block diagram of the research methodology including the simulation test bed. ................................................................. 60

3.2 Example BrainWeb images of brain MRI (a) T1, (b) T2 and (c) PD from (Montreal Neurological Institute and Hospital, nd). .......... 65

3.3 Example retinal images (a) and (b) of monomodal colour Fundus retinal images from the DRIVE dataset (Image Sciences Institute, nd) and (c) colour Fundus, (d) SLO from Brno-Multi dataset (Kolar et al., 2013; Kolar and Tasevsky, 2010). ........................................ 66

3.4 Example BrainWeb image of (a) uncorrupted brain MRI T1 u, (b) multiplicative 20% INU $\alpha$, (c) additive Gaussian noise $\beta$ and (d) measured T1 image a (Collins et al., 1998; Montreal Neurological Institute and Hospital, nd; Razlighi et al., 2013). ................. 70

3.5 Example BrainWeb images of misregistered brain MRI (a) T1+$\alpha_{40}$+$\beta$, (b) T2+$\alpha_{40}$+$\beta$ and (c) PD+$\alpha_{40}$+$\beta$ from (Montreal Neurological Institute and Hospital, nd). ................................. 71
3.6 Example retinal images from *Brno-Mono* dataset (a) colour Fundus and (b) only its green channel (shown in *grayscale* colour) from (Kolar et al., 2013; Kolar and Tasevsky, 2010). .......................... 72

3.7 Visual quality assessment of (a) multimodal T1+$\alpha$+$\beta$/T2 pair registration using *Canny edge detection* overlay (with inset showing zoomed in region) and (b) monomodal retinal registration zoomed-in view using *checkerboard* overlay (reproduced from (Gharabaghi et al., 2013)). .......................... 76

3.8 *Shepp-Logan* head phantom image (Guo and Huang, 2008; Haidekker, 2011) ........................................ 80

4.1 The neighbourhood region selection for feature extraction in EMPCA-MI, showing the dimensionality $d$ with radius (a) $r=1$, (b) $r=2$ and (c) $r=3$ pixels. .......................... 85

4.2 Illustration showing EMPCA-MI computation steps on reference $I_R$ and interpolated sensed $I^*_S$ images. .......................... 90

4.3 EMPCA-MI Traces with nearest neighbour, bilinear and bicubic interpolation for $x$ and $y$ axis translation for monomodal T2/T2 image pairs (a & b) T2+$\alpha$/T2, (c & d) T2+$\beta$/T2 with inset showing zoomed-in view. .......................... 96

4.4 EMPCA-MI Traces with nearest neighbour, bilinear and bicubic interpolation for $x$ and $y$ axis translation for multimodal PD/T1 image pairs (a & b) PD+$\alpha$/T1, (c & d) PD+$\beta$/T1. .......................... 97
4.5 ART performance of EMPCA-MI using nearest neighbour, bilinear and bicubic interpolation methods for monomodal T2/T2 and multimodal PD/T1 image pairs. ................................................. 98

4.6 RE versus iterations plots for nearest neighbour, bilinear and bicubic interpolation with EMPCA-MI-based monomodal T2/T2 registration using (a) 1st initial misregistration scenario and (b) 20th initial misregistration scenario, multimodal PD/T1 registration using (c) 1st initial misregistration scenario and (d) 20th initial misregistration scenario. ......................................................... 100

4.7 RE summary boxplot of (a) monomodal T2/T2 registration and (b) multimodal PD/T1 registration using nearest neighbour, bilinear and bicubic interpolation with EMPCA-MI for all 20 initial misregistration scenarios. The mean and outliers are denoted by ♦ and + shapes. The numbers on top of the plots refer to the number of out-of-range outliers with RE higher than 9 and 14 pixels respectively. ..................................................... 102

4.8 EMPCA-MI Traces with different radius values for x and y axis translation for monomodal T2/T2 image pairs (a & b) T2+α20/T2, (c & d) T2+β/T2. ................................................................. 103

4.9 EMPCA-MI Traces with different radius values for x and y axis translation for multimodal PD/T1 image pairs (a & b) PD+α20/T1, (c & d) PD+β/T1. ............................................................... 104
4.10 Number of samples \((q)\) with respect to different neighbourhood radii for the variety of datasets. ................................. 106

4.11 ART performances for EMPCA-MI and conventional PCA-MI with respect to different neighbourhood radii. ................................. 107

4.12 RE versus iterations plots for EMPCA-MI with varying neighbourhood radius for monomodal (a) \(T2+\alpha_{20}/T2\) pair (using \(12^{th}\) initial misregistration scenario) and (b) \(T2+\beta/T2\) pair (using \(16^{th}\) initial misregistration scenario). Multimodal (c) \(PD+\alpha_{20}/T1\) pair (using \(3^{rd}\) initial misregistration scenario) and (d) \(PD+\beta/T1\) pair (using \(17^{th}\) initial misregistration scenario). ................................. 108

4.13 RE summary boxplot of (a) monomodal \(T2/T2\) registration and (b) multimodal \(PD/T1\) registration using varying neighbourhood radius with EMPCA-MI for all 20 initial misregistration scenarios. The mean and outliers are denoted by ♦ and + shapes. The numbers on top of the plots refer to the number of out-of-range outliers with RE higher than 140 and 200 pixels respectively. ................................. 110

5.1 RE versus iteration plots for MI, GMI, RMI and EMPCA-MI using \(18^{th}\) initial misregistration scenario for monomodal registration of (a) \(T1/T1\), (b) \(T1+\alpha_{20}/T1\), (c) \(T1+\alpha_{40}/T1\), (d) \(T1+\beta/T1\) and (e) \(T1+\alpha_{40}+\beta/T1\) pairs. ................................. 117
5.2 Qualitative results of monomodal $T1 + \alpha_40 + \beta / T1$ pair using $18^{th}$ initial misregistration scenario showing (a) initial misregistration, (b) ground truth and (c) MI, (d) GMI, (e) RMI, (f) EMPCA-MI output with inset showing zoomed-in central lobes. . . . . . . . . . 119

5.3 RE versus iterations plots for MI, GMI, RMI and EMPCA-MI using $15^{th}$ initial misregistration scenario for multimodal registration of (a) $T1 + \alpha_40 + \beta / T2$ pair. . . . . . . . . . . . . . . . . . . . . . . . . . . . 121

5.4 Qualitative results of multimodal $T1 + \alpha_40 + \beta / T2$ pair using $15^{th}$ initial misregistration scenario showing (a) initial misregistration, (b) ground truth and (c) MI, (d) GMI, (e) RMI, (f) EMPCA-MI output with inset showing zoomed-in central lobes. . . . . . . . . . 122

5.5 RE summary boxplot of (a) monomodal $T1 / T1$ pair and (b) multimodal $T1 / T2$ pair using MI, GMI, RMI and EMPCA-MI for all 20 initial misregistration scenarios. The mean and outliers are denoted by ♦ and + shapes. The numbers on top of the plots refer to the number of out-of-range outliers with RE higher than 65 and 85 pixels respectively. . . . . . . . . . . . . . . . 124

5.6 ART performance of various similarity measures for Set 1: Brain-Web Dataset-based mono and multimodal registration experiments. 126

5.7 RE summary bar graph of 20 monomodal Fundus retinal image pairs using GFM and EMPCA-MI for $3^{rd}$ initial misregistration scenario.

N.B. Bar graph does not show GFM result for Image Pair 9 since the registration fails for the particular image pair. . . . . . . . . . 128
5.8 Qualitative results of monomodal Fundus retinal Image Pair 5 using 3\textsuperscript{rd} initial misregistration scenario, showing (a) initial misregistration, (b) ground truth, (c) GFM and (d) EMPCA-MI output. Also (e) shows the zoomed-in regions of ground truth (1\textsuperscript{st} row), GFM (2\textsuperscript{nd} row) and EMPCA-MI output (3\textsuperscript{rd} row) respectively. . . . . . . 130

5.9 Qualitative results of monomodal Fundus retinal Image Pair 9 using 3\textsuperscript{rd} initial misregistration scenario, showing (a) initial misregistration, (b) ground truth, (c) EMPCA-MI output, with (d) zoomed-in ground truth and (e) EMPCA-MI regions. N.B. GFM failed to register this image pair. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 131

5.10 RE summary boxplot of 20 monomodal Fundus retinal image pairs using GFM and EMPCA-MI for all 4 initial misregistration scenarios. The mean and outliers are denoted by ♦ and + shapes. The numbers on top of the plot refer to the number of out-of-range outliers with RE higher than 18 pixels. N.B. GFM results for scenario no. 4 were unavailable. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 132

5.11 RE summary bar graph of 10 multimodal Fundus/SLO retinal image pairs using RIR-BS and EMPCA-MI for 2\textsuperscript{nd} initial misregistration scenario. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 133
5.12 Qualitative results of multimodal Fundus/SLO retinal Image Pair 3 using 2nd initial misregistration scenario, showing (a) initial misregistration, (b) ground truth, (c) RIR-BS and (d) EMPCA-MI output. Also (e) shows the zoomed-in regions of ground truth (1st row), RIR-BS (2nd row) and EMPCA-MI output (3rd row) respectively. 135

5.13 RE summary boxplot of 10 multimodal Fundus/SLO retinal image pairs using RIR-BS and EMPCA-MI for all 3 initial misregistration scenarios. The mean and outliers are denoted by ♦ and + shapes. The numbers on top of the plot refer to the number of out-of-range outliers with RE higher than 190 pixels. 136

5.14 Qualitative results of multimodal T1+α40+β/T1 pair using 11th initial misregistration scenario showing (a) initial misregistration with (b) ground truth and (c) EMPCA-MI output with inset showing zoomed-in central lobes. 138

5.15 Qualitative results of multimodal Fundus/SLO retinal Image Pair 4 using 3rd initial misregistration scenario, showing (a) EMPCA-MI output (with P=2) and (b) RIR-BS output. Also (c) shows the three zoomed-in regions of EMPCA-MI output (with P=2) (1st row) and RIR-BS (2nd row) respectively. 140

6.1 Schematic showing (a) the original EMPCA-MI framework and the various approaches investigated as enhancements in (b) Step I and (c) Step III respectively. 143
6.2 Illustration of the original EMPCA-MI framework with 1st order (8-pixel) and the proposed 2nd order (8-pixel and 4-pixel) region connectivity-based pre-processing in Step I, for an image pair with size of 10 × 10 pixels.

6.3 RE summary boxplot of (a) monomodal T1/T1 pair and (b) multimodal T1/T2 pair using 1st order (8-pixel), 2nd order (8-pixel) and 2nd order (4-pixel) neighbourhood region connectivity in EMPCA-MI framework for all 20 initial misregistration scenarios. The mean and outliers are denoted by ♦ and + shapes. The numbers on top of the plots refer to the number of out-of-range outliers with RE higher than 14 and 20 pixels respectively.

6.4 ART performance for using original EMPCA-MI (1st order (8-pixel)) followed by 2nd order (8-pixel) and 2nd order (4-pixel) neighbourhood region connectivity in the EMPCA-MI framework for Brain-Web dataset-based mono and multimodal registration experiments.

6.5 Qualitative results of monomodal T1+α40+β/T1 pair using 11th initial misregistration scenario showing (a) initial misregistration and (b) ground truth followed results for (c) original EMPCA-MI framework (1st order (8-pixel)), then EMPCA-MI framework with (d) 2nd order (8-pixel) and (e) 2nd order (4-pixel) neighbourhood region connectivity enhancements (with inset showing the zoomed-in central lobes).
6.6 RE summary boxplot for EMPCA-MI framework using different number of bins selection strategies for monomodal Brno-Mono retinal dataset registration. *Blue* and *red* colour represent the fixed and adaptive number of bins selection approaches, while *mean* and *outliers* are denoted by ♦ and + shapes. . . . . . . . . . . . . . . . 160

6.7 ART summary bar graph for EMPCA-MI framework for monomodal Brno-Mono retinal dataset registration using different number of bins selection strategies. *Blue* and *red* colour represent the *fixed* and *adaptive* number of bins approaches. . . . . . . . . . . . . . . . 162

6.8 Qualitative results of monomodal Fundus retinal Image Pair 22, showing (a) initial misregistration, (b) ground truth and (c) the zoomed-in regions of ground truth and different *fixed* and *adaptive* number of bins selection strategies with EMPCA-MI framework. . . 163

7.1 Flowchart for the *adaptive* EMPCA-MI framework. . . . . . . . . . 169

7.2 RE summary boxplot of (a) monomodal T1/T1 pair and (b) multimodal T1/T2 pair using EMPCA-MI \{1\} and *adaptive* EMPCA-MI \{P\} for all 20 initial misregistration scenarios. The *mean* and *outliers* are denoted by ♦ and + shapes. The numbers on top of the plots refer to the number of *out-of-range* outliers with RE higher than 10 and 20 pixels respectively. . . . . . . . . . . . . . . . . . . 174
7.3 Qualitative results of monomodal T1+$\alpha_{40}+\beta$/T1 pair using 14\textsuperscript{th} initial misregistration scenario showing (a) initial misregistration, (b) ground truth and (c) EMPCA-MI \{1\} and (d) adaptive EMPCA-MI \{3\} output (with inset showing the zoomed-in central lobes). . . . 176

7.4 Qualitative results of multimodal T1+$\alpha_{40}+\beta$/T2 pair using 13\textsuperscript{th} initial misregistration scenario showing (a) initial misregistration, (b) ground truth and (c) EMPCA-MI \{1\} and (d) adaptive EMPCA-MI \{2\} output (with inset showing the zoomed-in central lobes). . . . 177

7.5 RE summary bar graph of 20 monomodal Fundus retinal image pairs using EMPCA-MI \{1\} and adaptive EMPCA-MI \{P\} for 3\textsuperscript{rd} initial misregistration scenario. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 178

7.6 Qualitative results of monomodal Fundus retinal Image Pair 20 using the 3\textsuperscript{rd} initial misregistration scenario, showing (a) initial misregistration, (b) ground truth, (c) EMPCA-MI \{1\} and (d) adaptive EMPCA-MI \{3\} output. Also (e) shows the zoomed-in regions of ground truth (1\textsuperscript{st} row), EMPCA-MI \{1\} (2\textsuperscript{nd} row) and adaptive EMPCA-MI \{3\} (3\textsuperscript{rd} row). . . . . . . . . . . . . . . . . . . . . . . . . . . . . 180

7.7 RE summary boxplot of 20 monomodal Fundus retinal image pairs using EMPCA-MI \{1\} and adaptive EMPCA-MI \{P\} for all 4 initial misregistration scenarios. The mean and outliers are denoted by ♦ and + shapes. The numbers on top of the plot refer to the number of out-of-range outliers with RE higher than 18 pixels. . . . . . . . . . . . . . . . . 181
7.8 RE summary bar graph of 10 multimodal Fundus/SLO retinal image pairs using EMPCA-MI \( \{I\} \) and adaptive EMPCA-MI \( \{P\} \) for the 2\textsuperscript{nd} initial misregistration scenario. .................. 182

7.9 Qualitative results of multimodal Fundus/SLO retinal Image Pair 6 using the 2\textsuperscript{nd} initial misregistration scenario, showing (a) initial misregistration, (b) ground truth, (c) EMPCA-MI \( \{I\} \) and (d) adaptive EMPCA-MI \( \{2\} \) output. Also (e) shows the zoomed-in regions of ground truth (1\textsuperscript{st} row), EMPCA-MI \( \{I\} \) (2\textsuperscript{nd} row) and adaptive EMPCA-MI \( \{2\} \) (3\textsuperscript{rd} row). .................. 183

7.10 RE summary boxplot of 10 multimodal Fundus/SLO retinal image pairs using EMPCA-MI \( \{I\} \) and adaptive EMPCA-MI \( \{P\} \) for all 3 initial misregistration scenarios. The mean and outliers are denoted by ♦ and + shapes. The numbers on top of the plot refer to the number of out-of-range outliers with RE higher than 150 pixels. . . 184

7.11 ART performance for EMPCA-MI \( \{I\} \) and adaptive EMPCA-MI \( \{P\} \) frameworks for Set 1 and Set 2 registration experiments. . . . 186

B.1 EMPCA-MI Traces for monomodal T1/T1, T1+α\textsubscript{20}/T1, T1+α\textsubscript{40}/T1, T1+β/T1 and T1+α\textsubscript{40}+β/T1 image pairs for \( x \), \( y \) translation and \( \theta \) rotation. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 247

B.2 EMPCA-MI Traces for monomodal T2/T2, T2+α\textsubscript{20}/T2, T2+α\textsubscript{40}/T2, T2+β/T2 and T2+α\textsubscript{40}+β/T2 image pairs for \( x \), \( y \) translation and \( \theta \) rotation. . . . . . . . . . . . . . . . . . . . . . . . . . . . . 248
B.3  EMPCA-MI Traces for monomodal PD/PD, PD+α20/PD, PD+α40/PD, PD+β/PD and PD+α40+β/PD image pairs for x, y translation and θ rotation. .................................................. 249

B.4  EMPCA-MI Traces for multimodal T1/T2, T1+α20/T2, T1+α40/T2, T1+β/T2 and T1+α40+β/T2 image pairs for x, y translation and θ rotation. .................................................. 250

B.5  EMPCA-MI Traces for multimodal T2/PD, T2+α20/PD, T2+α40/PD, T2+β/PD and T2+α40+β/PD image pairs for x, y translation and θ rotation. .................................................. 251

B.6  EMPCA-MI Traces for multimodal PD/T1, PD+α20/T1, PD+α40/T1, PD+β/T1 and PD+α40+β/T1 image pairs for x, y translation and θ rotation. .................................................. 252

C.1  EMPCA-MI Traces for monomodal T1/T1, T1+α20/T1, T1+α40/T1, T1+β/T1 and T1+α40+β/T1 image pairs for x, y translation and θ rotation. .................................................. 254

C.2  EMPCA-MI Traces for monomodal T2/T2, T2+α20/T2, T2+α40/T2, T2+β/T2 and T2+α40+β/T2 image pairs for x, y translation and θ rotation. .................................................. 255

C.3  EMPCA-MI Traces for monomodal PD/PD, PD+α20/PD, PD+α40/PD, PD+β/PD and PD+α40+β/PD image pairs for x, y translation and θ rotation. .................................................. 256
C.4 EMPCA-MI Traces for multimodal T1/T2, T1+α_20/T2, T1+α_40/T2, 
T1+β/T2 and T1+α_40+β/T2 image pairs for x, y translation and 
θ rotation. .......................................................... 257

C.5 EMPCA-MI Traces for multimodal T2/PD, T2+α_20/PD, T2+α_40/PD, 
T2+β/PD and T2+α_40+β/PD image pairs for x, y translation and 
θ rotation. .......................................................... 258

C.6 EMPCA-MI Traces for multimodal PD/T1, PD+α_20/T1, PD+α_40/T1, 
PD+β/T1 and PD+α_40+β/T1 image pairs for x, y translation and 
θ rotation. .......................................................... 259

D.1 RE versus iterations plots for MI, GMI, RMI and EMPCA-MI using 
18th initial misregistration scenario for monomodal registration of 
(a) T2+α_40+β/T2 and (b) PD+α_40+β/PD pairs. .......... 261

D.2 Qualitative results for T2+α_40+β/T2 with inset showing zoomed-in 
central lobes for 18th initial misregistration scenario. ........ 262

D.3 Qualitative results for PD+α_40+β/PD with inset showing zoomed-
in central lobes for 18th initial misregistration scenario. ........ 263

D.4 RE versus iterations plots for MI, GMI, RMI and EMPCA-MI using 
15th initial misregistration scenario for multimodal registration of 
(a) T1/T2, (b) T1+α_20/T2, (c) T1+α_40/T2 and (d) T1+β/T2 (e) 
T2+α_40+β/PD and (f) PD+α_40+β/T1 pairs. ............... 264
D.5 Qualitative results for T2+α40+β/PD with inset showing zoomed-in central lobes for 15th initial misregistration scenario. . . . . . . . 265

D.6 Qualitative results for PD+α40+β/T1 with inset showing zoomed-in central lobes for 15th initial misregistration scenario. . . . . . . . 266

D.7 RE summary boxplot of monomodal (a) T2/T2 and (b) PD/PD pairs using MI, GMI, RMI and EMPCA-MI for all 20 initial misregistration scenarios. The mean and outliers are denoted by ♦ and + shapes. The numbers on top of the plots refer to the number of out-of-range outliers with RE higher than 65 and 85 pixels respectively. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 267

D.8 RE summary boxplot of multimodal (a) T2/PD and (b) PD/T1 pairs using MI, GMI, RMI and EMPCA-MI for all 20 initial misregistration scenarios. The mean and outliers are denoted by ♦ and + shapes. The numbers on top of the plots refer to the number of out-of-range outliers with RE higher than 60 and 95 pixels respectively. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 268

E.1 RE summary boxplot of monomodal (a) T2/T2 and (b) PD/PD pairs using 1st order (8-pixel), 2nd order (8-pixel) and 2nd order (4-pixel) neighbourhood region connectivity in EMPCA-MI framework for all 20 initial misregistration scenarios. The mean and outliers are denoted by ♦ and + shapes. The numbers on top of the plots refer to the number of out-of-range outliers with RE higher than 35 and 15 pixels respectively. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 270
E.2 RE summary boxplot of multimodal (a) T2/PD and (b) PD/T1 pairs using 1st order (8-pixel), 2nd order (8-pixel) and 2nd order (4-pixel) neighbourhood region connectivity in EMPCA-MI framework for all 20 initial misregistration scenarios. The mean and outliers are denoted by ♦ and + shapes. The numbers on top of the plots refer to the number of out-of-range outliers with RE higher than 21 and 16 pixels respectively. . . . . . . . . . . . . . . . . . . . . . . . . . . . 271

F.1 RE summary boxplot of monomodal (a) T2/T2 and (b) PD/PD pairs using the EMPCA-MI {1} and adaptive EMPCA-MI {P} for all 20 initial misregistration scenarios. The mean and outliers are denoted by ♦ and + shapes. The numbers on top of the plots refer to the number of out-of-range outliers with RE higher than 11 and 15 pixels respectively. . . . . . . . . . . . . . . . . . . . . . . . . . . . 273

F.2 RE summary boxplot of multimodal (a) T2/PD and (b) PD/T1 pairs using the EMPCA-MI {1} and adaptive EMPCA-MI {P} for all 20 initial misregistration scenarios. The mean and outliers are denoted by ♦ and + shapes. The numbers on top of the plots refer to the number of out-of-range outliers with RE higher than 17 and 16 pixels respectively. . . . . . . . . . . . . . . . . . . . . . . . . . . . 274
## List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Comparative summary of major existing similarity measures.</td>
<td>56</td>
</tr>
<tr>
<td>3.1</td>
<td>Parameter settings for the generic MIR model.</td>
<td>63</td>
</tr>
<tr>
<td>3.2</td>
<td>Various clinical image dataset details (Kolar et al., 2013; Kolar and Tasevsky, 2010; Montreal Neurological Institute and Hospital, 2000, nd).</td>
<td>64</td>
</tr>
<tr>
<td>3.3</td>
<td>Simulation platform specifications and their details.</td>
<td>78</td>
</tr>
<tr>
<td>5.1</td>
<td>Parameter settings for EMPCA-MI framework.</td>
<td>114</td>
</tr>
<tr>
<td>5.2</td>
<td>Experiments and their relevant parameter details.</td>
<td>115</td>
</tr>
<tr>
<td>7.1</td>
<td>The different parameter selections for the <em>adaptive</em> EMPCA-MI framework.</td>
<td>171</td>
</tr>
<tr>
<td>A.1</td>
<td>Initial misregistration scenarios for the <em>BrainWeb</em> (Collins et al., 1998) dataset as in (Kotsas and Dodd, 2011).</td>
<td>244</td>
</tr>
<tr>
<td>A.2</td>
<td>Initial misregistration scenarios for the retinal datasets as in (Ghara-baghi et al., 2013).</td>
<td>245</td>
</tr>
</tbody>
</table>
List of Abbreviations

ART Average Runtimes
CC Correlation Coefficient
cMI conditional Mutual Information
CR Correlation Ratio
CT Computed Tomography
DB-ICP Dual-bootstrap Iterative Closest Point
DE Difference Image
DRIVE Digital Retinal Images for Vessel Extraction
EMPCA Expectation Maximisation for Principal Component Analysis
EMPCA-MI EMPCA with MI
FNMI Feature Neighbourhood MI
GC Gradient Correlation
GD Gradient Difference
GFM Geometrical Features Method
GMI Gradient MI
GPVE Generalised Partial Volume joint histogram Estimation
HO-MI Higher-order MI
INU Intensity Non-Uniformities
LBP Local Binary Pattern
LC Local correlation
MC-MI Multi-channel MI
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>MF-MI</td>
<td>Multi-feature MI</td>
</tr>
<tr>
<td>MI</td>
<td>Mutual Information</td>
</tr>
<tr>
<td>MIR</td>
<td>Medical Image Registration</td>
</tr>
<tr>
<td>MRI</td>
<td>Magnetic Resonance Imaging</td>
</tr>
<tr>
<td>MRI-PD</td>
<td>Proton Density weighted MRI</td>
</tr>
<tr>
<td>MRI-T2</td>
<td>T2 relaxation time weighted MRI</td>
</tr>
<tr>
<td>NCC</td>
<td>Normalised Correlation Coefficient</td>
</tr>
<tr>
<td>NHS</td>
<td>National Health Service</td>
</tr>
<tr>
<td>NMI</td>
<td>Normalised MI</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear Magnetic Resonance</td>
</tr>
<tr>
<td>OF-MI</td>
<td>Organ Focused MI</td>
</tr>
<tr>
<td>PC</td>
<td>Personal Computer</td>
</tr>
<tr>
<td>PCA</td>
<td>Principal Component Analysis</td>
</tr>
<tr>
<td>PCA-MI</td>
<td>Conventional PCA with MI</td>
</tr>
<tr>
<td>PI</td>
<td>Pattern Intensity</td>
</tr>
<tr>
<td>PVI</td>
<td>Partial Volume Interpolation</td>
</tr>
<tr>
<td>RE</td>
<td>Registration Error</td>
</tr>
<tr>
<td>RIR-BS</td>
<td>Vascular Bifurcation Structures Method</td>
</tr>
<tr>
<td>RIU</td>
<td>Ratio Image Uniformity</td>
</tr>
<tr>
<td>RMI</td>
<td>Regional MI</td>
</tr>
<tr>
<td>SAD</td>
<td>Sum of Absolute Difference</td>
</tr>
<tr>
<td>SE-MI</td>
<td>Spatially Encoded MI</td>
</tr>
<tr>
<td>SLO</td>
<td>Scanning Laser Ophthalmoscopy</td>
</tr>
<tr>
<td>SSD</td>
<td>Sum of Squared Difference</td>
</tr>
<tr>
<td>WC</td>
<td>Woods Criterion</td>
</tr>
</tbody>
</table>
## List of Symbols

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\triangleq$</td>
<td>Is equal by definition to</td>
</tr>
<tr>
<td>$\approx$</td>
<td>Approximately equal to</td>
</tr>
<tr>
<td>$\in$</td>
<td>Is an element of</td>
</tr>
<tr>
<td>$\not\in$</td>
<td>Is not an element of</td>
</tr>
<tr>
<td>$</td>
<td>$</td>
</tr>
<tr>
<td>$\sim N(\mu, \sigma^2)$</td>
<td>Normal distribution</td>
</tr>
<tr>
<td>$\nabla_x(\psi)$</td>
<td>Gradient vector at point $x$ of scale $\psi$ in GMI</td>
</tr>
<tr>
<td>$\alpha_{20}$</td>
<td>20% INU</td>
</tr>
<tr>
<td>$\alpha_{40}$</td>
<td>40% INU</td>
</tr>
<tr>
<td>$\beta$</td>
<td>Gaussian noise</td>
</tr>
<tr>
<td>$\gamma_k$</td>
<td>Unit vector in the $x_k$-coordinate direction</td>
</tr>
<tr>
<td>$\epsilon$</td>
<td>Tolerance threshold for Powell optimisation</td>
</tr>
<tr>
<td>$\theta$</td>
<td>Rotational angle in $xy$ plane</td>
</tr>
<tr>
<td>$\theta_{initial}$</td>
<td>Initial rotational angle in $xy$ plane</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>Weighting function of GMI</td>
</tr>
<tr>
<td>$\lambda_P, \lambda_{P'}$</td>
<td>Eigenvalue of $P$ and $P'$</td>
</tr>
<tr>
<td>$\mu$</td>
<td>Mean</td>
</tr>
<tr>
<td>$\mu_{MN}$</td>
<td>Central moment of order $(M + N)$ of binary 2D region</td>
</tr>
<tr>
<td>$O$</td>
<td>Computational complexity</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>Variance</td>
</tr>
</tbody>
</table>
List of Symbols

\( \tau \) Transformation parameters

\( \tau_{\text{final}} \) Final registration transformation parameters

\( \tau_{\text{initial}} \) Initial misregistration transformation parameters

\( \phi_{x,x'}(\psi) \) Angle between the gradient vectors for given \( \psi \) in GMI

\( \psi \) Scale of Gaussian kernel in GMI

\( \Lambda \) Bifurcation structure vector for RIR-BS

\( \Sigma_d \) Covariance matrix of \( d \)-dimensional normal distribution

\( \Upsilon_i \) Branch angle for RIR-BS

\( \Phi \) Transforming variable

\( \Phi^{\text{new}} \) Updated value of \( \Phi \)

\( \Phi^T \) Transpose of \( \Phi \)

\( \Omega_S \) Unknown states

\( \Omega_S^T \) Transpose of \( \Omega_S \)

\( a \) Measured MRI image

\( b \) Number of bins

\( B \) Sliding window with radius \( r \)

\( B^* \) Processed column vector using 2\(^{nd}\) order connectivity

\( c \) Neighbourhood region connectivity

\( d \) Dimensional Subspace

\( \det \) Determinant

\( D_{\infty} \) Chebyshev distance

\( e \) Euler’s number constant (\( \approx 2.71828 \))

\( f(x) \) Sample function for Powell optimisation

\( g \) Center pixel of the neighbourhood sliding window

\( G \) Gradient term in GMI
List of Symbols

$h$  
Pixel location within the sliding window

$H_g$  
Entropy term in RMI

$H(I_R), H(I_S)$  
Marginal entropy of $I_R$ and $I_S$ respectively

$H(I_R, I_S)$  
Joint entropy between $I_R$ and $I_S$

$H(X_R), H(X_S)$  
Marginal entropy of $X_R$ and $X_S$ respectively

$H(X_R, X_S)$  
Joint entropy between $X_R$ and $X_S$

$i$  
Number of rows to be scanned by the sliding window

$I$  
Sample image

$I_R$  
Reference image

$I_S$  
Sensed image

$I^*_S$  
Interpolated sensed image

$I^*$  
Interpolated sample image

$I_{11-44}$  
coordinates of sample image $I$

$I_1-4$  
First four affine moment invariant descriptors for GFM

$j$  
Number of columns to be scanned by the sliding window

$k$  
Natural number having value $1 < k < q$

$K$  
kurtosis

$l_i$  
Normalised branch length for RIR-BS

$l_k$  
Starting vector of Powell optimisation with index $k$

$m$  
Number of columns in the image

$n$  
Number of rows in the image

$\mathbb{N}$  
Set of natural numbers

$P, P'$  
Number of principal components

$p$  
Probability function
List of Symbols

$q$  Number of observed data vectors
$Q_R, Q_S$  Processed $I_R$ and $I_S^*$ respectively using $r$
$r$  Neighbourhood radius
$R$  Range of $Z$
$S_k$  Generalised column vector $s$
$s$  Column vector
$S$  Uniform scaling in $xy$ plane
$S_{initial}$  Initial uniform scaling in $xy$ plane
$SM$  Similarity measure
$t_x$  Translation in $x$ axis direction
$t_y$  Translation in $y$ axis direction
$u$  Uncorrupted MRI image
$U$  Count of possibilities for computing probability
$v$  Normally distributed Gaussian noise
$V$  Covariance Matrix
$w$  Number of elements in $Z$
$W$  Design variables for Powell optimisation
$x$  $x$ axis
$x_0$  Initial starting point of Powell optimisation
$x_{initial}$  Initial translation in $x$ axis
$(x, y)$  Original pixel location in an image
$(x', y')$  Transformed pixel location in an image
$(\bar{x}, \bar{y})$  Pixel location of center of gravity in binary 2D image
$X_R, X_S$  EMPCA output for $Q_R$ and $Q_S$ respectively
$y$  $y$ axis
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( y_{\text{initial}} )</td>
<td>Initial translation in ( y ) axis direction</td>
</tr>
<tr>
<td>( z )</td>
<td>Latent variable</td>
</tr>
<tr>
<td>( Z )</td>
<td>Sample data for selecting number of bins</td>
</tr>
<tr>
<td>( \bar{Z} )</td>
<td>Mean of ( Z )</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

1.1 Overview

Medical imaging is a technical process to create visual representations of the interior of a body for medical inspection. In 1895, Röntgen firstly discovered that X-rays could identify bone structures (Kevles, 1997) and since then they have been widely used for the detection of pathology of the skeletal system. In the past decades, they have also been used to identify diseases such as pneumonia, lung cancer and kidney stones, though somewhat less useful in the imaging of soft tissues like the brain or muscle (Sherrow, 2006). Recent advancements in medical physics have led to the development of generic medical imaging modalities/technologies such as Computed Tomography (CT), Magnetic Resonance Imaging (MRI) (Analoui et al., 2012) as well as application-specific technologies such as mammography, angiography and fluoroscopy (Paragios et al., 2015) which can be used to image almost the whole human anatomy.
According to the National Health Service (NHS), in England, a total of 37.9 million imaging tests were conducted during the year ending March 2014, with an increase of more than 5% compared to the previous year (NHS England Analytical Services, 2014). Medical imaging has thus become one of the most indispensable components of healthcare technology by providing complementary anatomic and functional information which helps the physician in detecting and diagnosing diseases and health conditions (Bourland, 2012).

The medical images acquired using the imaging technologies need to be aligned amongst themselves as this allows the physician to examine the advancement of a disease over a period of time or to compare the pre/post-surgery condition of a patient (Gunderman, 2013). Similarly retinal scans for a patient over a course of time will inevitably have been taken using different orientations, resolutions and/or magnification settings, and these will need to be physically aligned to help facilitate the physician perform clinical analysis (Holz and Spaide, 2010). Most of the advanced medical applications such as image guided therapy/surgery require the acquired images to be in alignment (Khalifa et al., 2011).

The process of geometrically transforming a source medical image in order to attain the best physical alignment with a reference medical image is called *medical image registration* (MIR) (Modersitzki, 2004). It uses an optimisation method to maximise a similarity measure with known transformations between the two images. It is usually considered to be a mathematical optimisation problem with the similarity measure playing a critical role in its successful convergence (Hajnal and Hill, 2001). This process is termed monomodal or multimodal depending on whether the images to be aligned are from the same or different modalities.
Medical image registration can be broadly classified into feature and intensity techniques (Zitová and Flusser, 2003). For example, feature-based approaches for registering the retinal images primarily use optical disc, fovea (Xu et al., 2007) and vascular structural details (Lowe, 2004; Sofka and Stewart, 2006), while brain images use edge detection and surface extraction followed by the least mean square of extracted features (Hsu and Loew, 2001). Although feature-based approaches are fast, their performance is very dependent on the segmentation quality and degree of overlap of the extracted features. This is especially evident in the case of challenging retinal Fundus and Scanning Laser Ophthalmoscopy (SLO) images which normally have inherently low contrast and large homogeneous non-vascular region characteristics, which can undermine their registration performance.

In contrast, intensity-based registration employs pixel intensity information by means of similarity measures such as normalised correlation coefficient (NCC) (Rogelj and Kovačič, 2003), phase correlation (Kolar et al., 2013), Fourier-based methods (Maintz and Viergever, 1998) and mutual information (MI) (Pluim et al., 2003). Although all these similarity measures are commonly used for monomodal registration, MI has been proven to be popular because of its ability to perform multimodal image registration of various parts of human anatomy (Tagare and Rao, 2015).

Multimodal registration is considered more demanding in comparison to the monomodal registration since it aligns different modalities having structural and functional imaging information effectively which is not always complementary in nature (Analoui et al., 2012). Multimodal registration also acts as a prerequisite for applications such as multimodal image fusion and its visualisation.
For example, MI has been used to perform multimodal registration between MRI, CT as well as *positron emission tomography* and *single photon emission computed tomography* images (Studholme et al., 1996a; Yin et al., 2010; Zhu and Cochoff, 2002) which enable the physician to provide an accurate diagnosis. Furthermore, in comparison to feature and other intensity-based similarity measures, MI does not require any prior segmentation or pre-processing steps. It is for these reasons that MI has been readily adopted for a variety of registration problems in the fields of microscopy, histology, computer vision and remote sensing (Paragios et al., 2015). In the last decade, few variants of MI-based similarity measures have emerged and offer solutions to the multimodal MIR problem (Fernandez-de Manuel et al., 2014).

### 1.2 Mutual Information-based Multimodal MIR

MI establishes a statistical relationship between the intensity values of the images to quantify the level of alignment between a source and reference image by utilising the underlying fundamental concepts of information theory (Hartley, 1928). However during MIR, it is sensitive to the interpolation artefacts, and its performance can be severely compromised when the overlap region between the images is small (Pluim et al., 2000b). While *normalised* MI (NMI) has been shown to successfully register partially overlapping images (Loi et al., 2008), it along with MI is unable to consistently and accurately register images containing *intensity non-uniformities* (INU) which are an omnipresent feature in MRI images. INU is a smooth intensity variation caused by factors such as radio frequency excitation field inhomogeneity, non-uniform reception coil sensitivity and eddy currents (Zhuang et al., 2011).
Furthermore, a degraded registration performance is also observed for retinal images having innately challenging characteristics such as low contrast and large homogeneous regions (Kubecka and Jan, 2004). This has led to the recent development of hybrid approaches which combine different aspects of the aforesaid feature and intensity methods (Holden et al., 2004; Legg et al., 2008; Studholme et al., 1996b). Higher-order MI (HO-MI) (Rueckert et al., 2000) and gradient MI (GMI) (Pluim et al., 2000a) for instance, help align images even at lower sampling resolutions, achieving better global maxima with low interpolation induced local minima. However, their performance degrades as noise is introduced.

On the other hand, methods such as regional MI (RMI) and its variants incorporate neighbourhood features within MI by segmenting the image into several regions for feature extraction to reduce the influence of INU and noise in the case of MRI images. Similarly for retinal images, the extracted vascular structures with spatial information are used for RMI and feature neighbourhood MI to improve the registration performance (Russakoff et al., 2004). In computing the associated entropies, these MI-based approaches employ a covariance matrix instead of high-dimensional histograms to reduce data complexity, though as the size of a neighbourhood grows, the computation overheads also increase correspondingly.

As will be discussed in Chapter 2, the current literature highlights the presence of inherent artefacts and noise in medical images as one of the major challenges in both rigid and non-rigid registration (Argyriou et al., 2015; Wilson and Laxminarayan, 2007). The two key examples are INU and noise in MRI images (Zhuang et al., 2011) and large homogeneous region with low contrast in retinal images (Legg et al., 2015), which make their registration especially challenging.
Since patterns of INU fields vary in different MRI images, the inconsistency of intensity values of same tissue region in two images lead to large errors when existing techniques are used (Razlighi et al., 2013). Similarly for retinal images, the quality of images is further compromised by the presence of different pathologies like haemorrhages and retinal scars caused by laser treatment (Holz and Spaide, 2010; Saine and Tyler, 2002).

Previous contributions (Butz and Thiran, 2001; Chappelow and Madabhushi, 2010; Kang et al., 2011; Luan et al., 2008; Yang et al., 2006; Yi and Soatto, 2011; Zhang et al., 2011; Zheng, 2010; Zouqi et al., 2010) have also combined spatial information in different ways for these challenging cases but the computation overhead remains high. This provided the key motivation for this research as there exists the clear prospect to exploit these opportunities and develop a new hybrid similarity measure which is able to provide better accuracy and robustness in the registration performance alongside lower computational costs.

1.3 Research Motivation

The inclusion of spatial features along with MI for multimodal MIR similarity measure has been considered as a promising solution (Fernandez-de Manuel et al., 2014; Woo et al., 2015) because it provides better robustness and performance. Although it can significantly improve the registration process, there are still many underlying challenges to resolve in realising a robust and flexible hybrid similarity measure for MIR.
Some of the main challenges to be addressed include:

- In the presence of inherent artefacts in particular medical modalities, rigid and non-rigid registration processes can be difficult, and convergence is not always guaranteed. For example, in case of MRI and retinal Fundus images, having high INU and large homogeneous non-vascular regions, respectively.

- Local features used along with information-theoretic similarity measures help in registering the medical images with the aforementioned characteristics by utilising higher order dimensionality, but incur high data and computational complexity. Conversely using the fast but rough approximating dimensionality reduction techniques underperform, especially in the presence of artefacts and noise.

- While dimensionality reduction for feature selection is possible, the stopping rule is a challenging problem in machine learning (Cangelosi and Goriely, 2007; Jackson, 1993; Villacampa, 2015) and can lead to performance degradation.

- Although the similarity measure plays a critical role in the image registration process, inappropriate choice of interpolation technique for a given transformation and optimisation can degrade the convergence rate of the registration process.

- For information theoretic similarity measures such as MI, there are various choice of parameters which impact significantly on the probability distribution estimation, especially when the medical images have varying image intensity characteristics.
1.3 Research Motivation

These challenges and the general aim of improving the robustness and performance of multimodal MIR without increasing the computational costs were the main motivations behind this research. These are important requirements both in terms of being able to register challenging multimodal images and practical computational costs.

As mentioned in Section 1.2, the performance of both rigid and non-rigid registration is known to be compromised in the presence of inherent artefacts in medical images. Commonly, rigid registration is performed prior to non-rigid registration being undertaken for two reasons (Holden, 2008; Rueckert et al., 2000). Firstly, it helps in ascertaining the global correspondences between the images, and secondly, it reduces the computation time needed for the non-rigid registration stage. Hence, it is essential in addressing the aforementioned challenges for the rigid registration scenario to be considered in order to reduce the likelihood of the propagation of errors within the subsequent non-rigid registration step (Argyriou et al., 2015; Modersitzki, 2004; Zhao and Jia, 2015).

For these reasons, this thesis will only focus on rigid image registration. Furthermore, while multimodal registration is the primary objective of this research, without loss of generality both mono and multimodal rigid registration techniques will be considered and critically investigated. This all serves to provide the context for the overarching thesis research question and the related objectives which are discussed in the next section.
1.4 Research Question and Objectives

From the above discussion, the following research question was framed:

*How can the multimodal medical images be efficiently and accurately registered?*

After a detailed review of the existing methods and a critical evaluation of prospective techniques, a strategy combining local regional features with MI for a robust and flexible similarity measure was identified as a particularly promising area for investigation as it has considerable potential to achieve superior rigid registration performance on challenging mono and multimodal medical images. Such a strategy would not only enable the integration of image features within an entropy-based approach, but also offer greater flexibility in terms of both computational cost and registration performance by employing effective dimensionality reduction.

A set of three research objectives was framed to underpin the above overarching research question, and these are shown in Figure 1.1. The objectives are specifically:

1. *To develop and critically evaluate a hybrid similarity measure framework which utilises regional spatial features for efficient registration performance of multimodal datasets.*

*Justification:* The objective is to develop a new similarity measure, which utilises dimensionality reduction to not only extract spatial features but to lower the computational overheads with minimal impact on the registration
1.4 Research Question and Objectives

Objective 1

Objective 2

Objective 3

Adaptive feature selection process

Development of local regional features based hybrid similarity measure

Enhancement 1: Neighbourhood region relationships

Enhancement 2: Mutual information parameter selection

Flexibility & Performance

Figure 1.1: Layout of the framework objectives in this thesis.

performance. It will address the potential problem of degraded registration performance of medical images having inherent characteristics such as INU and non-vascular regions. The use of spatial regions for the dominant feature selection from the image dataset is followed by MI computation. The assumption to choose only the most significant features aims to lower the computational cost, while observing the corresponding impact upon the quality of the registration. Furthermore, by investigating appropriate spatial radius selection and interpolation techniques for the new similarity measure, faster convergence will be feasible.

2. To critically analyse key neighbourhood region relationships and framework parameter choices.
1.4 Research Question and Objectives

Justification: This objective seeks to enhance the capability of the new hybrid similarity measure framework by relaxing both the spatial region relationships and fixed parameter selection assumptions made in Objective 1, while still using only the most significant features. Since the neighbourhood region represents these features, it is essential that an accurate representation is used. A neighbourhood region scheme needs to be developed and incorporated into the similarity measure to achieve this. Similarly, choosing the best parameters for the MI computation will allow more accurate registration. Figure 1.1 shows these two enhancements which offer a refined and more flexible similarity measure while for computational efficiency, using only the most significant features.

3. To develop and critically evaluate a mechanism to improve robustness by adaptively selecting the best number of features, based on image characteristics.

Justification: Many existing hybrid similarity measures use high-dimensional representation of the spatial features and consequently experience high computational cost (Russakoff et al., 2004; Zhuang et al., 2011). Reducing the computation by estimating the similarity measure generally leads to poorer registration performance especially for images with artefacts (Legg et al., 2009). Hence, there is a need to develop new mechanisms which provide the flexibility to incorporate a varying number of features, based upon the dataset characteristics. As shown in Figure 1.1, this objective will enable the framework to have the flexibility to trade different levels of granularity for different mono and multimodal registration scenarios, as well as a corresponding trade-off with the computational cost incurred.
1.5 Contributions

In summary, the overall aim is to develop a flexible similarity measure framework which provides consistently robust performance with lower computational overheads by utilising the hybrid combination of spatial features with MI for multimodal MIR of challenging image datasets. The new hybrid similarity measure achieves this aim by successfully fulfilling Objectives 2 and 3, which build upon the basis of new similarity measure framework developed as an outcome of Objective 1.

1.5 Contributions

To accomplish the aforementioned three objectives, this thesis presents a new robust hybrid *Expectation Maximisation for Principal Component Analysis with MI* (EMPCA-MI) *similarity measure framework*, which combines the neighbourhood feature information with MI. Initially the framework is developed with a series of underlying assumptions, which are subsequently relaxed to provide some important framework enhancements. It explores the impact of neighbourhood region relationships and parameter selection for MI computation and integrates an adaptive mechanism to select the most appropriate number of features for the EMPCA-MI similarity measure framework during the registration process. Each of these enhancements increases the flexibility of the framework as well as increasing its robustness to registering multimodal medical images.

The three original scientific contributions made in this thesis to the multimodal MIR domain are as follows:
1. A new EMPCA-MI similarity measure framework which utilises local regional features is introduced and analysed to address the challenge of multimodal image registration in the presence of modality artefacts. It involves a three step process: firstly, local regional features are identified and processed by employing neighbourhood region connectivity. Secondly, an iterative *Expectation Maximisation for Principal Component Analysis* (EMPCA) technique is applied to extract the most dominant features. Finally, the spatial alignment is quantified using MI. The appropriate selection of the interpolation technique and neighbourhood radius is determined, and the registration performance and computational efficiency of the new framework are benchmarked against existing comparative similarity measures.

2. Two novel enhancements are subsequently integrated into the EMPCA-MI framework designed in contribution 1. Firstly, higher-order neighbourhood region connectivity is exploited and, secondly, parameter selection strategies for MI are analysed. Higher-order neighbourhood region connectivity preserves the intensity relationships and so better represents the features, especially, for example, in MRI images which contain artefacts. Similarly, the new, variable parameter strategies are able to accommodate low contrast and large homogeneous regions present in many retinal image datasets.

3. Finally, an *adaptive* EMPCA-MI similarity measure framework is developed which can automatically determine the most significant features, based on the image characteristics to achieve further improvements in the registration performance. Furthermore, it affords a flexible trade-off mechanism between overall registration quality and the computational time involved.
1.6 Thesis Structure

The rest of the thesis is organised as following:

- *Chapter 2* presents a historical background for medical imaging, before a thorough literature review of the different components of MIR algorithms is presented. It also focuses on the strengths and weaknesses of existing MI-based similarity measures and their associated variants which are used for MIR.

- *Chapter 3* discusses the research methodology deployed, including the description of MIR test-bed and its parameters settings. It also presents the various testing scenarios employed, together with details on the clinical datasets, performance metrics and software validation methods used.

- *Chapter 4* details the formulation of the new EMPCA-MI similarity measure framework for MIR. It also provides an empirical evaluation for the choice of interpolation technique and neighbourhood radius. Work from this chapter has been published in (Reel et al., 2012b) and (Reel et al., 2012a).

- *Chapter 5* rigorously assesses the performance of the EMPCA-MI similarity measure framework presented in *Chapter 4*. Using various test clinical images under different scenarios, a critical evaluation is undertaken in terms of both the registration quality and computational efficiency against existing similarity measures. Work from this chapter has been published in (Reel et al., 2013c) and (Reel et al., 2013b).

- *Chapter 6* leverages from the key findings of *Chapter 5* to boost the EMPCA-
MI similarity measure framework with two enhancements. *Enhancement 1* extends the idea of neighbourhood region connectivity to a higher order in the pre-processing stage, while *Enhancement 2* investigates the impact of introducing different fixed and parameter selection strategies for the MI computation. Both these enhancements enable the EMPCA-MI similarity measure framework to provide improved registration performance in comparison with the original framework proposed in *Chapter 4*. Work from this chapter has been published in (Reel et al., 2013a) and (Reel et al., 2014a).

- *Chapter 7* formulates the *adaptive* EMPCA-MI similarity measure framework which has the ability to choose the best number of features instead of empirically choosing the most significant features as in *Chapter 4* and integrating the two enhancements in *Chapter 6*. It provides the flexibility in terms of robust registration performance and corresponding trade-off in computational cost. Work from this chapter has been published in (Reel et al., 2014b).

- *Chapter 8* presents potential future directions for the new EMPCA-MI similarity measure framework in the MIR domain.

- Finally, *Chapter 9* summarises the main findings and contributions emanating from this research.
1.7 Summary

In this chapter, the overarching research question and three principal research objectives have been framed to address some of the key challenges in multimodal MIR, particularly in the presence of artefacts in clinical images and the correspondingly high computational overheads incurred by existing techniques. The next chapter will provide a deeper context for this by presenting an overview of medical imaging history and a critique of MIR literature.
Chapter 2

Medical Image Registration: A Review

2.1 Introduction

Imaging is one of the fundamental tools used in the modern medical science. Today, imaging technologies are used by physicians to detect cancerous lesions, diagnose broken bones as well as perform image guided surgeries and track the progression of a disease (Paragios et al., 2015). For instance, diabetic retinopathy is the first cause of blindness in the working age population in the United Kingdom and had led to the recently announced NHS diabetic eye screening programme which will extensively employ retinal imaging to diagnose diabetic retinopathy and will be offered annually to all people with any type of diabetes aged 12 or over (Keenan et al., 2013; Martín-Merino et al., 2014).
2.2 Brief History of Medical Imaging

Different medical imaging technologies are used to create different pictures of the human anatomy. While these modalities generally provide different information, there are few similarities. For example, the external contour of two different modalities images can be similar, while the information they contain can be complementary. The integration of this imaging information from different modalities creates a compilation, which can be used to improve diagnosis and treatment of the subject. Hence, image registration is used to spatially align the medical images from the same or different modalities and is one of the key tasks in the medical image processing domain (Goshtasby, 2012).

While MIR is a vital and commonly used process, no one technique or method is known to work efficiently for all applications. Also as mentioned in Chapter 1, the clinical images with INU and large homogeneous non-vascular regions present a challenge for registration. This chapter discusses the historical background of medical imaging in brief and the need of image registration in medical imaging and its current challenges. The goal of this chapter is to provide a broad overview of the MIR literature and a detailed discussion on existing MI-based similarity measures highlighting their limitations.

2.2 Brief History of Medical Imaging

Medical imaging is a technical process which is used to create images of the human body for clinical purposes. This ability of looking into the human body is an essential diagnostic tool in the field of medical science and prime topic of research in healthcare technology (Damas et al., 2011b; Liu et al., 2008).
Medical imaging technologies have improved significantly since the late nineteenth century, since many different imaging modalities have been developed and are in practical use clinically. All these modalities are based on different physics principles and are usually more or less suitable for different part of the human anatomy (Hendee and Ritenour, 2002; Hobbie, 2001). Figure 2.1 shows the major technologies discovered and invented for imaging of human brain and eye retina. This timeline summarises the various modalities and is not exhaustive since new techniques are added every few years (Brooks, 2001; Cherry, 2009). Although the research and development of imaging the various anatomy parts (such as heart, lungs and abdomen) has progressed considerably, this section addresses the focus of this thesis which is related to modalities containing INU and large homogeneous non-vascular regions.
2.2 Brief History of Medical Imaging

The first published account of in-vivo human retinal Fundus photography is from 1886 (T. Jackman and J.D. Webster, 1886), which used the novel apparatus of albo-carbon burner shown in Figure 2.2(a) and with an exposure of 2½ minutes, it was used to photograph the first very fair image of large blood vessels of the living human retina. As observed in Figure 2.2(b) the image obtained was blurry and needed more advancement in the instrumentation and acquisition techniques before the high quality Fundus photography could become a routine procedure (Saine, 1993). In 1895, the first X-ray picture was taken by the German physicist W. C. Röntgen, showing the skeletal composition of his wife’s hand, as shown in Figure 2.2(c) (Kevles, 1997). Initially, X-rays were a novelty but by the end of the year they were used by physicians around the world (Linton, 1995).

Around half century later in 1946, F. Bloch and E. Purcell developed the concept of using magnets for taking pictures of a living being, which led to the nuclear magnetic resonance (NMR) spectroscopy (Gunderman, 2013) and they were awarded the Nobel Prize in Physics in 1952.
2.2 Brief History of Medical Imaging

Furthermore in 1961, Novotny and Alvis published the first successful retinal fluorescein angiogram (H. R. Novotny and D. L. Alvis, 1961) which showed the vascular pattern and blood flow in a human retina as shown in Figure 2.3(a). Ten years later, the CT scanning machine was invented by G. Hounsfield and A. Cormack (Analoui et al., 2012). Figures 2.3(b) shows the first brain CT scan taken by them (Beckmann, 2006).

In the next four years, MRI was invented by P. Lauterbur and E. Mandfield for which they were awarded the Nobel Prize for Physiology or Medicine in 2003 (Bushong and Clarke, 2013). The MRI in particular, is considered to be a special modality since using different pulse sequences, the modality produces different images. For example, consider two contrasts e.g. proton density weighted MRI (MRI-PD) and T2 relaxation time weighted MRI (MRI-T2). A proton density sequence used in MRI detects the proton density of the subject while the T2 relaxation time sequence in MRI detects the transverse relaxation time of a proton in its environment.
2.2 Brief History of Medical Imaging

![MRI-T2 and MRI-PD images](image1)

![Fundus and SLO images](image2)

Figure 2.4: Various medical imaging modalities for (a & b) human brain, (c & d) eye retina from (Montreal Neurological Institute and Hospital, nd) and (Kolar et al., 2013; Kolar and Tasevsky, 2010).

Figures 2.4(a) and (b) show the trans-axial brain images of MRI-T2 and MRI-PD respectively. Here, MRI-PD image displays the dense tissues but shows very little detail of the brain tissue, while MRI-T2 image on the other hand, displays brain tissue details more clearly (Ackerman, 1998). However, MRI is known to be prone to non-anatomic INU variations due to radio frequency non-uniformities and static field inhomogeneity. INU along with the Gaussian noise causes the corruption of tissue images leading to their poorer visual quality (Simmons et al., 1994; Sled and Pike, 1998).
In 1979, G. Hounsfield and A. Cormack shared the Nobel Prize for Physiology or Medicine for their part in developing the computer assisted tomography (Beckmann, 2006). In the meantime, the retinal diagnosis also progressed considerably with the development of better quality Fundus imaging and the invention of SLO and Optical Coherence Tomography (Saine and Tyler, 2002).

Figures 2.4(c) and (d) show the Fundus and SLO image, respectively, of the same human retina. These retinal imaging methods have specific clinical applications for identifying eye conditions and diseases like myopia, glaucoma and diabetic retinopathy (Saine and Tyler, 2002; Sanchez-Galeana et al., 2001). They provide complementary information with the colour Fundus image showing the boundary of the optic nerve head, while the SLO images reveal deep layer reflectivity and surface topology of both the optic nerve head and retina. However, retinal images exhibit the inherent characteristics of low contrasted intensity distributions, as well as having large homogeneous non-vascular regions (Holz and Spaide, 2010).

Some of imaging modalities in the literature can be classified as structural or functional imaging systems. For example, the MRI and CT provide high-resolution information about the structure and anatomy, while functional MRI imaging gives the functional information (Cherry, 2009). Furthermore, they have been prominently used for both clinical and industrial research. Many derivative medical imaging modalities also exist such as magnetic resonance angiography from magnetic resonance imaging, computed tomography angiography from computed tomography and digital subtraction angiography from X-rays (Schneider et al., 2005).
The imaging activity conducted by NHS, England from February 2014 to February 2015 comprises of X-rays having the biggest share of all tests performed followed by Ultrasound, CT, MRI and others (NHS England Analytical Services, 2015). In the past decade, research and development trends in medical imaging have also been significant with the utilisation of methods/techniques such as biomarkers and gene expression along with the use of big data mining and informatics for the new cellular and molecular level imaging methods (Analoui et al., 2012; Gunderman, 2013). Hence, they give the unprecedented ability to detect, diagnose and monitor the various pathological as well as physiological conditions.

Most of these medical imaging modalities broadly deploy at least one of the key image processing tasks namely segmentation, registration and visualisation (Stytz et al., 1991). While medical image segmentation is usually performed as a pre-processing step to define a region of interest, on the other hand, medical image visualisation acts as a visual aid to see the medical images from different perspective and angles in order to facilitate in the diagnosis process (Brock, 2013). But, in order to fuse and visualise more than one image it is important that a correspondence is drawn among them using MIR (George, 2011; Wang et al., 2008). It can also be further used to study the time series information of mono or multimodal images, which can help to evaluate the growth rates in the region of interest (Alberto, 2011).

As mentioned in Chapter 1, the focus of this thesis is MIR in particular. The next section gives a detailed literature review of the MIR process.
2.3 Medical Image Registration

Broadly, the process of MIR utilises the feature or intensity-based information for various parts of human anatomy (Hill et al., 2001). The feature-based registration typically requires some features to be extracted or identified before the registration process. For example for retinal images, these features can be edges, contours, surfaces, statistical features, control points or high level structural descriptors (Khalifa et al., 2011). On the other hand, intensity-based methods do not require any feature extraction or segmentation. Hence, raw pixel values are used straightforwardly for registration process, as in case of brain images (Maintz and Viergever, 1998). Overall, intensity-based registration has particularly gained popularity, due to its ability to register multimodal images more effectively in comparison to the feature-based registration (Pluim et al., 2003).

Usually, two images are used in the generic image registration process (or more in case of group-wise registration (Paragios et al., 2015)), where one of the medical images is considered as reference image $I_R$ while the other image as sensed image $I_S$. Image registration is usually considered to be an optimisation problem with the prime goal of finding the spatial mapping which enables the alignment of $I_S$ with respect to image $I_R$. The generic MIR model is a multi-step, iterative process (Zitová and Flusser, 2003) which involves:

1. Transforming the coordinates of the sensed image $I_S$ with known transformation parameters $\tau$ in a given reference space (transformation step).

2. Generating a new interpolated sensed image $I^*_S$ in the reference space (interpolation step).
3. Comparing $I_S$ with the reference medical image $I_R$ using a similarity measure (compute similarity measure step).

4. Optimising the transformation parameters iteratively to achieve the best possible alignment at $\tau_{\text{final}}$, where its similarity measure value is maximum between the images (optimisation step).

5. This value of final transformation is saved.

This process of registration can be generalised as a maximisation problem:

$$
\tau_{\text{final}} = \arg \max_{\tau} SM(I_R, \tau(I_S)) \quad (2.1)
$$

where $SM$ denotes the similarity measure used and $\tau$ is the class of allowed affine and similarity transformations.

The aforementioned steps for a generic MIR model are shown in Figure 2.5 and discussed in detail as follows:

### 2.3.1 Transformation

In order to determine the position of corresponding points in reference and sensed images, a combination of translation, rotation and scaling parameters is considered. Image registration methods employ transformations such as rigid, affine and elastic (non-rigid) transformations (Zitová and Flusser, 2003). The rigid transformation considers $t_x$ and $t_y$ translations along the $x$ and $y$ axes, and a rotational angle $\theta$ for the registration process (Brown, 1992). It assumes that the subject in the image
maintain its shape and size (Yu, 2005). Affine transformations offer a high degree of flexibility in accommodating linear distortions by allowing scaling and shearing in addition to translation and rotation (Jenkinson and Smith, 2001).

The choice of using a particular transformation model for registration depends on the anatomy parts and their characteristics. For example, in the case of retinal images, the similarity transformation as in (Gharabaghi et al., 2013; Legg et al., 2008, 2009) is popularly used because it allows the magnification changes either due to the use of different optical instrument for image acquisition or the changes due to the motion in the direction of the optical axis (Ryan et al., 2004), represented as a uniform scaling $S$. It is a special form of global affine transform which includes $S$ along with $t_x$, $t_y$ and $\theta$ parameters. Hence, the similarity transformation can be defined as:

![Generic MIR model](image_url)
2.3 Medical Image Registration

\[
\begin{bmatrix}
x' \\
y'
\end{bmatrix}
= \begin{bmatrix}
t_x & S \cos \theta & S \sin \theta \\
t_y & -S \sin \theta & S \cos \theta
\end{bmatrix}
\begin{bmatrix}
1 \\
x \\
y
\end{bmatrix}
\tag{2.2}
\]

Here, \((x, y)\) and \((x', y')\) are the original and transformed pixel positions, respectively, and (2.2) becomes a rigid transformation when \(S=1\).

Elastic transformations provide more degrees of freedom as compared to rigid and affine transformations. It is also called non-rigid transformation due to its non-linear nature. Non-rigid transformations can be broadly classified on the basis of physical models or basis function expansion. While linear elasticity (Moshfeghi, 1991), viscous fluid flow (Christensen et al., 1997) and optical flow (Horn and Schunck, 1981) are examples of physical model-based transformations, radial basis functions (Fornefett et al., 2001), multiquadrics (Zhang et al., 2002), thin-plate splines (Bookstein, 1989), B-spline (Denton et al., 1999), wavelets (Wu et al., 2000) and piecewise affine transforms (Crum, 2004) are some of basis function expansion transformations.

Various surveys on transformation are conducted in the past and it is considered very difficult to robustly conclude which transformation is the most accurate or best suited to a particular application (Crum, 2004; Holden, 2008; Klein et al., 2009). Although physical models are considered to have an advantage in non-rigid registration since they offer realistic solutions, some of their properties hinder the registration. For example, linear elasticity can only accurately model small deformations, which is a limitation because often soft tissues exhibit large deformation. Fluid flow offers a solution for this and can ensure that the topology is preserved.
However, it cannot model the elastic component of the tissue deformation.

In reality, tissue exhibits a complex behaviour, which in certain conditions considered as elastic or visco-elastic. Demons (Crum, 2004; Thirion, 1998) is one of the most popular optical flow models, but it also has no constraints on the displacement and does not necessarily preserve the tissue structure.

In practice, most of the time, non-rigid registrations are not performed directly (Zhao and Jia, 2015). Rigid registration is used as a pre-processing step to non-rigid registration, in order to compute the global transformation parameters and then later non-rigid registration is performed to analyse the local deformation of the subject (Holden, 2008; Rueckert et al., 2000, 1999) as shown in Figure 2.6. The primary focus of this thesis is on the rigid registration stage.

### 2.3.2 Interpolation

Interpolation is one of the important techniques used in medical image processing in a variety of different applications (Lehmann et al., 1999; Thevenaz et al., 2000). During registration the mapping of corresponding points as per transformation model requires them to be mapped to a non-grid position. Interpolation interpol-
ates the position of the points with respect to the grid. Intensity values can be estimated in the mapped position using a particular interpolation method (Ibanez et al., 2003). There are various types of interpolation methods used for medical image processing. They can be broadly classified as either filter or function-based. Mean, Gaussian, B-spline and median (Maeland, 1988) based interpolation methods are filter-based interpolation, while nearest neighbour, bilinear, bicubic, Lagrange and splines (quadratic, cubic, B-spline) (Grevera et al., 1999; Lehmann et al., 1999; Shi and Reichenbach, 2006) interpolations are based on functions.

Figure 2.7 shows the most commonly used interpolation methods namely, nearest neighbour, bilinear and bicubic. It shows the new pixel value assigned at position $I^*(x, y)$ using pixel values from image $I$. Both nearest neighbour and bilinear interpolations are less computationally intensive but offer poorer sub-pixel accuracy. Bicubic interpolation is more popular due to its higher accuracy in comparison to nearest neighbour and bilinear as well as being less computationally complex than spline interpolation (Wang et al., 2008). Spline interpolation (Pan et al., 2012), which represents the image intensity using a spline function, but incur higher computational cost. While higher-order spline functions are smoother, providing considerable accuracy in some cases, if however they are generalised, they can lead to overfitting. Meanwhile Fourier analysis-based windowed sinc interpolation (Meijering et al., 1999), is formed by multiplying a sampled sinc function (Olver et al., 2010) by a windowing function. This produces very high accuracy, but is not chosen for an iterative registration process since it involves a high computational overhead. It is used later, once the optimal transformation values are achieved, to re-sample the final registered image (Ibanez et al., 2003).
2.3 Medical Image Registration

\[ I^*(x, y) = I[x + 0.5][y + 0.5] \]

\[ I^*(x, y) = (1-t_x)(1-t_y)I_{11} + (1-t_x)t_yI_{12} + t_x(1-t_y)I_{21} + t_xt_yI_{22} \]

\[ I^*(x, y) = \begin{bmatrix} f(1+t_x) & f(t_y) & f(1-t_y) & f(2-t_y) \end{bmatrix} \times \begin{bmatrix} I_{11} & I_{12} & I_{13} & I_{14} \\ I_{21} & I_{22} & I_{23} & I_{24} \\ I_{31} & I_{32} & I_{33} & I_{34} \\ I_{41} & I_{42} & I_{43} & I_{44} \end{bmatrix} \times \begin{bmatrix} f(t_x) \\ f(t_y) \\ f(1-t_x) \\ f(2-t_x) \end{bmatrix} \]

(a) Nearest Neighbour

(b) Bilinear

(c) Bicubic

Figure 2.7: (a) Nearest neighbour, (b) Bilinear and (c) Bicubic Interpolation Methods.

While all the above mentioned interpolation methods, especially higher order methods, introduce new intensity values for the interpolated image, partial volume interpolation (PVI) (Maes et al., 1997) ensures that no new intensity values are
generated. In fact in traditional sense, it does not qualify as an interpolation method, since it directly updates the joint histogram, without creating any new pixel values at non grid positions. While PVI also results in abrupt changes in the joint histogram dispersion and, hence, has a serious influence on registration process. Later generalised partial volume joint histogram estimation (GPVE) was proposed (mei Chen and Varshney, 2003), which reduced the abrupt changes using B-spline instead of linear interpolation in processing GPVE, reducing the local optima effectively (Jacquet et al., 2010). Recently, Gaussian, cubic spline and notch general partial volume interpolation are proposed and their perform is better than others in its class (Pan et al., 2012). Although most of the medical registration applications are software-based, many have been implemented in hardware due to their high computational cost (Lehmann et al., 1999; Moses et al., 2011).

2.3.3 Similarity Measure

A similarity measure gives the ability to determine the level of global correspondence between two images. During the registration process the parameters of a given transformation model are changed, based on the optimisation technique until the similarity measure reaches a maximum value (Wilson and Laxminarayan, 2007). Hence the choice of similarity measure along with optimisation method plays a crucial role for a successful outcome of a registration process. In all, similarity measure quantifies the spatial alignment between two images. Various intensity-based similarity measures such as sum of squared difference (SSD) (Friston et al., 1995), sum of absolute difference (SAD) (Lemieux et al., 1994), correlation coefficient (CC) (Cideciyan, 1995), NCC (Rogelj and Kočačič, 2003) and ratio image
uniformity (RIU) (Holden et al., 2000) have been proposed for monomodal registration processes. These measures do not perform well in all cases. While SSD is highly sensitive to Gaussian noise, SAD is less sensitive to outliers on the subject boundaries. CC, NCC and RIU perform well in these conditions, but are highly sensitive to non uniform illumination in the images. Other similarity measures are entropy of difference image (DE), pattern intensity (PI), gradient correlation (GC) (Wu et al., 2009) and gradient difference (GD) (Penney et al., 1998). While DE and PI are robust against large difference of intensity values between monomodal images, GC and GD are highly sensitive to thin line structures.

Most of the these similarity measures have been prominently used for monomodal images and are known for their inability to perform multimodal registration (Gao et al., 2008). This is due to the fact that all of them assume a linear relationship between the intensity values of the two images. This is mostly the case when both the images belong to the same modality and a monomodal registration is undertaken. On the other hand, multimodal images do not share any global linear relationship between each other in terms of intensity values. However, predominant relationships can be established between the intensity values of smaller neighbourhoods of the images. This makes local correlation (LC) (Netsch et al., 2001) and correlation ratio (CR) (Roche et al., 1998) suitable for multimodal registration, but they suffer from expensive computations. Meanwhile, feature-based approaches instead of using the pixel intensity information, extract key structures like optic disc (Li and Chutatape, 2004) and vascular structural features (Lowe, 2004; Sofka and Stewart, 2006) from the retinal images. Bifurcation point matching approaches such as dual-bootstrap iterative closest point (DB-ICP) (Stewart
et al., 2003) and its variant (Pereira et al., 2012) accomplish registration by using vascular features to grow a bootstrap region. While these methods are computationally fast, their performance is very dependent on the segmentation quality and degree of overlap of the extracted features.

The geometrical features method (GFM) (Gharabaghi et al., 2013) is a recent feature-based method for monomodal retinal image registration. It extracts closed-boundary regions from the green channel of retinal fundus image and computes the affine moment invariant descriptors for each of the binary regions in both reference and sensed images. The affine moment invariant descriptors are moment based descriptors of planar shape, which are derived by means of the algebraic invariants theory and are invariant under general affine transformations (Flusser and Suk, 1993). GFM utilises the first four simplest affine moment invariant descriptors denoted as $I_1, I_2, I_3$ and $I_4$ (Flusser and Suk, 1993), which are defined as:

\begin{align*}
I_1 &= (\mu_{20}\mu_{02} - \mu_{11}^2)/\mu_{00}^4 \\
I_2 &= (\mu_{30}\mu_{03}^2 - 6\mu_{30}\mu_{21}\mu_{12}\mu_{03} + 4\mu_{03}\mu_{12}^3 + 4\mu_{21}\mu_{03}^3 - 3\mu_{21}^2\mu_{12}^2)/\mu_{00}^{10} \\
I_3 &= (\mu_{20}(\mu_{21}\mu_{03} - \mu_{12}^2) - \mu_{11}(\mu_{30}\mu_{03} - \mu_{21}\mu_{12}) + \mu_{02}(\mu_{30}\mu_{12} - \mu_{21}^2))/\mu_{00}^7 \\
I_4 &= (\mu_{20}^3\mu_{03}^2 - 6\mu_{20}\mu_{11}\mu_{12}\mu_{03} - 6\mu_{20}\mu_{02}\mu_{21}\mu_{03} + 9\mu_{20}^2\mu_{02}\mu_{12}^2 + 12\mu_{20}\mu_{02}\mu_{30}\mu_{12} - 8\mu_{11}\mu_{30}\mu_{03} - 6\mu_{20}\mu_{02}^2\mu_{30}\mu_{12} + 9\mu_{20}\mu_{02}^2\mu_{21}^2 + 12\mu_{11}\mu_{02}\mu_{30}\mu_{12} - 6\mu_{11}\mu_{02}\mu_{30}\mu_{12} + \mu_{02}^3\mu_{30}^2)/\mu_{00}^{11} \tag{2.3}
\end{align*}

Here, the central moment $\mu_{MN}$ of order $(M+N)$ of binary 2D region is defined as:

\[
\mu_{MN} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x-\overline{x})(y-\overline{y})dxdy \quad M, N = 0, 1, 2, \ldots \tag{2.4}
\]
Here, \((\bar{x}, \bar{y})\) are the coordinates of the center of gravity of the binary region. For example, the central moment using \(M = 1\) and \(N = 2\), is denoted as \(\mu_{21}\). The complete proof and detailed discussion on the the properties of invariants is discussed in (Maitra, 1979) and (Amanatiadis et al., 2009). Once the affine moment invariant descriptors are computed, region matching is performed using the feature distance (Gharabaghi et al., 2013) between them. The three most likely corresponding regions are then selected and used to estimate the registration parameters between the reference and sensed images. While GFM has inherent computational simplicity, it exhibits poor resilience to noise in images, especially when higher-order invariant descriptors are selected (Rodrigues, 2000).

Similarly, for multimodal retinal image registration - vascular bifurcation structures (RIR-BS) has performed satisfactorily for linear and affine transformation models (Chen et al., 2015, 2011). Unlike other conventional feature point-matching methods, whom largely depend on the branching angles of single bifurcation point, RIR-BS employs a feature structure-matching approach. RIR-BS utilises the bifurcation structure which comprises a master bifurcation point and its three connected neighbours, as shown in Figure 2.8. The characteristic vector of each bifurcation structure consists of the normalized branching angle and length, which is invariant against translation, rotation, scaling, and even modest distortion. The characteristic vector for each bifurcation structure in the reference and sensed image is formulated as:

\[
\Lambda = \{ \text{lengths, angles} \}
\]

\[
= \{ l_1, l_2, l_3, \gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_5, \gamma_6, \gamma_7, \gamma_8, \gamma_9, \gamma_{10}, \gamma_{11}, \gamma_{12} \}
\]
Figure 2.8: Bifurcation structure comprising of a master bifurcation point and its three connected neighbours. Figure reproduced from (Chen et al., 2015).

Here, $l_i$ and $\Upsilon_i$ represent the normalised length and angle respectively:

$$l_i = \frac{\text{ith branch length}}{\text{sum \{length1, length2, length3\}}}$$

$$\Upsilon_i = \frac{\text{ith branch angle in degree}}{360^\circ}$$

The corresponding bifurcation structure pairs from reference and sensed image are matched in order to align them. Although, these normalised bifurcation structures can greatly reduce the ill-posed nature of the matching process, this method can only be deployed as long as the vasculature pattern can be segmented reliably.

Various graph measures-based on minimal spanning trees (Hero et al., 2002; Ma et al., 2000; Neemuchwala et al., 2004), have been also proposed, as an alternative to histogram-based measures since they all eliminate the need of estimating the intensity densities of images. They can be extended to registration of more than two images, but require large memory and computing resources compared to histogram-based measures. Since MI and its variants have been extensively used as similarity measures for image registration, especially in the field of medical ima-
ging (Pluim et al., 2003), Section 2.4 is dedicated to critically discuss them. The next section discusses the role of optimisation in the registration process.

### 2.3.4 Optimisation

In order to find the maximum value of similarity measure for a given registration process, an optimisation problem is formulated on the basis of number of parameters used for transformation (Hajnal and Hill, 2001). The choice of the transformation is essentially dependent on the application and its geometrical complexity (i.e. degrees of freedom). Although an optimal solution is guaranteed by exhaustive search, its computational expense is proportional to size of the search space as well as the number of parameters used for transformation and, hence, becomes infeasible as they increase (Khalifa et al., 2011). Therefore, this forms the motivation to explore refined search strategies or optimisation methods which can help to find the maximum value for a given similarity measure.

A good optimisation method should be reliable and be capable of finding the best possible transformation quickly (Goshtasby, 2012). The choice of optimisation method hence requires a good understanding of the registration problem along with the constraints that can be applied and a knowledge of numerical analysis. Many optimisation methods have been introduced and adopted for the registration process, on the basis of the transformation parameters, similarity measure, time constraints and required accuracy of registration.

Optimisation methods can be formally classified in terms of function-based, such as Powell (Collignon et al., 1995; Maes et al., 1997), Downhill Simplex
Medical Image Registration (Meyer et al., 1997) methods or gradient-based, such as simple (Rueckert et al., 1999; Wells III et al., 1996) and conjugate gradient descent (Hayton et al., 1999), stochastic gradient approximation (De Craene et al., 2004; Spall, 1992).

Downhill Simplex performs well in the absence of local distortions and can be used by combining all the transformation parameters in one function (Press et al., 2007). Relaxation Matching (Price, 1985) and Dynamic Programming (Milios, 1988) gives the advantage of handling local distortion along with overall global transformation, in case of non-rigid registration. Various other prominently used methods in the literature, including Newton-Raphson (Miga, 2003), Levenberg-Marquardt (Kabus et al., 2004; Maes et al., 1999; Thevenaz et al., 1995), simulated annealing (Mazur et al., 1993), Monte Carlo (Wong, 2010) and evolutionary computing methods such as particle swarm (Rezaei et al., 2009) and genetic algorithms (Damas et al., 2011a). Hierarchical methods (Bajcsy and Kovačič, 1989; Klein et al., 2007; Lester and Arridge, 1999; Maes et al., 1999; Thevenaz and Unser, 2000) for registration of multi-resolution medical images are also used for registering low resolution versions of the image and then progressed to higher resolution, which helps enhancing the optimisation speed and robustness of registration. Using different filters, e.g. Gaussian (Bajcsy and Kovačič, 1989; Burt and Adelson, 1983; Suárez et al., 2002), Spline (Thevenaz et al., 1995, 1998; Thevenaz and Unser, 1997; Unser et al., 1993) and Wavelet-based pyramids (Allen et al., 1993; Amit, 1994; Deubler and Olivo, 1997; Gefen et al., 2003; Wu et al., 2000) are created for multi-resolution medical images (Jenkinson and Smith, 2001; Rohlfing et al., 2003; Rueckert et al., 1999; Thevenaz et al., 1995; Wirth et al., 2002; Yeung et al., 1998).
As discussed before, since the overall performance of a registration algorithm depends on the effectiveness of the optimisation method. This in turn, depends on the number of iterations needed for the algorithm to converge. Hardware-based implementations of MIR consider the choice of optimisation method on the basis that they are fully or partially parallelise (Cooper et al., 2011; Shams et al., 2010; Tsoi et al., 2009). Various comparative evaluations for the above mentioned optimisation techniques have been conducted in the past (Khalifa et al., 2011) which investigated them on the basis of different factors which impact the performance of the registration such as accuracy, computational cost and number of iterations needed. Amongst them Powell and its variants, which are zero-order methods, have shown a balanced performance and hence been popularly adopted for a variety of registration applications in the medical imaging domain (Rios and Sahinidis, 2012; van der Bom et al., 2011).

Algorithm 2.1 shows the pseudo-code of the Powell optimisation method, considering the maximisation of a sample function \( f(x) \) involving \( W \) variables. Furthermore, Powell method can be utilise with Brent line search, which combines a parabolic interpolation with the golden section algorithm (Press et al., 2007). Hence together Powell Brent is known to provide an overall fast and accurate local search performance (Pluim et al., 2003; Zitová and Flusser, 2003). In case of MIR, it can iteratively search in a given parameter space such as \((t_x, t_y, \theta, S)\) in case of similarity transformation model, to locate the highest value of the similarity measure by individually maximising one parameter at a time. Subsequent iterations then start from the maximum of these four parameters until a predefined tolerance threshold \( \epsilon \) is met.
**Algorithm 2.1** Pseudo Code for Powell optimisation method

**Input:** Function $f(x)$

**Variables:** $W$ design variables; $x_0$ initial starting point; $l_k$ starting vector; $k$ index varying from 1 to $W$; $\epsilon$ tolerance threshold.

**Output:** Maximum value of $f(x)$

1: Choose a point $x_0$ in the design space.

2: Choose the starting vectors $l_k$, $k = 1, 2, \ldots, W$. (the recommended choice is $l_k = \gamma_k$, where $\gamma_k$ is the unit vector in the $x_k$-coordinate direction).

3: START

FOR each $k$

Maximise $f(x)$ along the line through $x_{k-1}$ in the direction of $l_k$.

Let the maximum point be $x_k$.

END LOOP

$l_{W+1} \leftarrow x_0 - x_W$

Maximise $f(x)$ along the line through $x_0$ in the direction of $l_{W+1}$.

Let the maximum point be $x_{W+1}$.

IF $|x_{W+1} - x_0| < \epsilon$ THEN

GOTO STEP 4.

END IF

FOR each $k$

$l_k \leftarrow l_{k+1}$ ($l_1$ is discarded and the other vectors are reused).

END LOOP.

4: END

Hence, the iterative registration process shown in Figure 2.5 is then concluded by saving the final transformation value $\tau_{final}$ using (2.1), at which the reference and sensed image are best registered.

The next section focuses on MI and its variants used as similarity measures in MIR.
2.4 MI and its variants for multimodal MIR

While MI originated in the field of information and communication theory some decades ago, it was first introduced into the field of image registration during the mid 1990s, by two independent research groups (Collignon et al., 1995; Maes et al., 1997) and (Wells III et al., 1996). While both the versions were the same, they were formulated independently and differently. As mentioned earlier, the prime success of MI as a similarity measure is because it makes very few assumptions about the images to be registered. It only assumes statistical dependence between the images, which helps it in enhancing its performance, especially for multimodal registration, where the relationship between the intensity values of two images is non-linear.

Research related to MI began in 1980s with the start of intensity-based registration methods. These methods had an advantage over traditional registration methods, since no prior segmentation or feature extraction was needed. While initial intensity-based methods involved image registration by matching the moments (Flusser and Suk, 1994) and principal axis (Alpert et al., 1990), later correlation-based approaches (Manduchi and Mian, 1993) emerged. However, these techniques were not effective for multimodal image registration because they only used intensity information (van den Elsen et al., 1993).

In 1993, the first similarity measure for multimodal data was proposed and known as the Woods Criterion (WC) (Woods et al., 1992, 1993), which quantified the relationship between the intensity values for corresponding points for a known region. It needed however, the regions of both images to be segmented manually.
before registration. This was overcome by an adaptation from (Hill et al., 1993), which used a feature space in the form of a 2D joint histogram. At the correct alignment, the histogram is found to be more focused and dispersed during misaligned positions. Later, Hartley introduced Entropy as a similarity measure (Collignon et al., 1995), which was originally developed in communication theory (Hartley, 1928). Hartley’s entropy was extended by Shannon, which did not assume that each symbol was equally likely to occur but could be expressed as a probability distribution (Shannon, 1948). The final phase of this research movement was the independent introduction of MI by (Viola and Wells, 1995) and (Collignon et al., 1995). While Viola and Wells technique used Parzen window density estimation (Wells III et al., 1996), Collignon’s formulation used a frequency histogram technique for Kullback-Leiber or Shannon’s information (Maes et al., 1997).

Given the reference image $I_R$ and its probability $p_{I_R}(I_R)$, the entropy $H(I_R)$ is defined as:

$$H(I_R) = -\sum_{I_R} p_{I_R}(I_R) \log_2 p_{I_R}(I_R)$$

(2.7)

Correspondingly, for sensed image $I_S$, $H(I_S)$, is defined as:

$$H(I_S) = -\sum_{I_S} p_{I_S}(I_S) \log_2 p_{I_S}(I_S)$$

(2.8)

Then, the joint histogram is constructed with the entries $U_{I_R I_S}$ using $b$ number of bins. $U$ is the total count of possibilities, while $U_{I_R}$ and $U_{I_S}$ are count of possibility for $I_R$ and $I_S$ only, respectively. The marginal probability $p_{I_R}(I_R)$ and
\( p(I_S) \), and the joint probability \( p_{IRIS}(I_R, I_S) \) are determined by:

\[
p_{IR}(I_R) = \frac{U_{IR}}{U} \tag{2.9}
\]

\[
p_{IS}(I_S) = \frac{U_{IS}}{U} \tag{2.10}
\]

\[
p_{IRIS}(I_R, I_S) = \frac{U_{IRIS}}{U} \tag{2.11}
\]

Correspondingly, the joint entropy \( H(I_R, I_S) \) are defined as:

\[
H(I_R, I_S) = - \sum_{I_R, I_S} p_{IRIS}(I_R, I_S) \log_2 p_{IRIS}(I_R, I_S) \tag{2.12}
\]

The conditional entropies \( H(I_R|I_S) \) and \( H(I_S|I_R) \) are defined as:

\[
H(I_R|I_S) = - \sum_{I_R, I_S} p_{IRIS}(I_R, I_S) \log_2 \frac{p_{IRIS}(I_R, I_S)}{p_{IS}(I_S)} \tag{2.13}
\]

\[
H(I_S|I_R) = - \sum_{I_R, I_S} p_{IRIS}(I_R, I_S) \log_2 \frac{p_{IRIS}(I_R, I_S)}{p_{IR}(I_R)} \tag{2.14}
\]

The strength of dependence between \( I_R \) and \( I_S \) is measured by their MI value as:
\[ MI(I_R, I_S) = H(I_R) + H(I_S) - H(I_R, I_S) = H(I_R) - H(I_R | I_S) \] (2.15)

\[ = H(I_S) - H(I_S | I_R) \]

\[ = - \sum_{I_R, I_S} p_{I_RI_S}(I_R, I_S) \log_2 \left\{ \frac{p_{I_RI_S}(I_R, I_S)}{p_{I_R}(I_R)p_{I_S}(I_S)} \right\} \]

To illustrate how to calculate the MI, consider the following scenario of two matrices (representing an image dataset of spatial dimensions $3 \times 3$ pixels) $I_R$ and $I_S$ is as follows:

Let, $I_R = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix}$ and $I_S = \begin{bmatrix} 3 & 3 & 1 \\ 2 & 2 & 1 \\ 1 & 3 & 1 \end{bmatrix}$.

Then, the joint probability distribution can be tabulated for one to one correspondence, using $b = 3$ as:

<table>
<thead>
<tr>
<th></th>
<th>$I_R$</th>
<th>1 2 3</th>
<th>$p_{I_S}(I_S)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$I_S$</td>
<td>1</td>
<td>2/9</td>
<td>0 2/9 4/9</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>0 1/9</td>
<td>1/9 2/9</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>2/9</td>
<td>1/9 0 3/9</td>
</tr>
<tr>
<td>$p_{I_R}(I_R)$</td>
<td>4/9</td>
<td>2/9</td>
<td>3/9</td>
</tr>
</tbody>
</table>

Here, the rightmost column and bottom row consists of the marginal probabilities $p_{I_S}(I_S)$ and $p_{I_R}(I_R)$, respectively. Similarly, the first three columns and rows
show the joint probability $p_{IRIS}(IR, IS)$. For example, the joint probability value in first column and row denotes the one to one correspondence where elements of $IR$ and $IS$ are both equal to 1. Moreover, this 2D histogram table shows all the combinations of one to one correspondence possible, with elements equal to 0 highlighting that the particular combination does not exist.

Finally, the calculation of MI using (2.15) is:

$$MI(IR, IS) = \frac{2}{9} \log_2 \left( \frac{8}{9} \right) + \frac{2}{9} \log_2 \left( \frac{8}{12} \right) + \frac{1}{9} \log_2 \left( \frac{4}{3} \right) + \frac{1}{9} \log_2 \left( \frac{4}{3} \right) + \frac{2}{9} \log_2 \left( \frac{12}{9} \right) + \frac{1}{9} \log_2 \left( \frac{12}{9} \right) + \frac{1}{9} \log_2 \left( \frac{6}{4} \right) + \frac{1}{9} \log_2 \left( \frac{6}{4} \right) + \frac{2}{9} \log_2 \left( \frac{9}{8} \right) + \frac{1}{9} \log_2 \left( \frac{9}{6} \right) + \frac{1}{9} \log_2 \left( \frac{9}{4} \right) + \frac{1}{9} \log_2 \left( \frac{9}{4} \right) + \frac{2}{9} \log_2 \left( \frac{9}{8} \right) + \frac{1}{9} \log_2 \left( \frac{9}{6} \right) + \frac{1}{9} \log_2 \left( \frac{9}{4} \right) + \frac{1}{9} \log_2 \left( \frac{9}{4} \right)$$

$$\approx 0.6227.$$

Here, the MI value quantifies the spatial alignment between $IR$ and $IS$. Now considering the scenario where $IS$ is rotated $+90^\circ$, while $IR$ remains unchanged as follows:

$IR = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix} \text{ and } IS = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 3 \\ 1 & 1 & 1 \end{bmatrix}.$

Then, the joint probability distribution can be tabulated as:
2.4 MI and its variants for multimodal MIR

<table>
<thead>
<tr>
<th></th>
<th>$I_R$</th>
<th>$I_S$</th>
<th>$p_{IS}(I_S)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\frac{4}{9}$</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$I_S$</td>
<td>2</td>
<td>0</td>
<td>$\frac{2}{9}$</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
<td>$\frac{3}{9}$</td>
</tr>
<tr>
<td>$p_{IR}(I_R)$</td>
<td>$\frac{4}{9}$</td>
<td>$\frac{2}{9}$</td>
<td>$\frac{3}{9}$</td>
</tr>
</tbody>
</table>

and

$$MI(I_R, I_S) = \frac{4}{9} \log_2 \left( \frac{4/9}{16/81} \right) + \frac{2}{9} \log_2 \left( \frac{2/9}{4/81} \right) + \frac{3}{9} \log_2 \left( \frac{3/9}{9/81} \right)$$

$$= \frac{4}{9} \log_2 \left( \frac{9}{4} \right) + \frac{2}{9} \log_2 \left( \frac{9}{2} \right) + \frac{3}{9} \log_2 \left( \frac{9}{3} \right)$$

$$\approx 1.5305.$$  

Evidently, the MI value has increased since $I_R$ and $I_S$ are both now spatially registered. Although, the marginal probabilities $p_{IR}(I_R)$ and $p_{IS}(I_S)$ are same in both scenarios, the joint probability has decreased in the latter case. This is highlighted in the second scenario by the tabulated distribution being noticeably more concentrated compared with the earlier misaligned registration. Moreover, it is clear MI uses the one-to-one correspondence without assuming any neighbourhood relationships.

Since the original proposals of MI, various researchers have contributed to the performance, robustness, and other relevant issues of MI. Many research groups then demonstrated its properties using different medical imaging data (Gao et al.,
2.4 MI and its variants for multimodal MIR

2008; Holden et al., 2000; Studholme et al., 1996b, 1997; West et al., 1997) and clinical applications (HAWKES, 1998; Meyer et al., 1997). While MI has been successfully applied to a range of modalities, there are few cases, when the measure does not provide correct spatial alignment, due to the presence of local or spurious global maxima (Rodriguez-Carranza and Loew, 1998), this is particularly observed when the regions of overlaps are small (Studholme et al., 1998).

Two independent groups (Rodriguez-Carranza and Loew, 1998; Studholme et al., 1998) proposed NMI, which improved the behaviour of MI, by significantly enhancing the overlap invariance of the measure. NMI can be denoted as:

\[
\text{NMI}(I_R, I_S) = \frac{H(I_R) + H(I_S)}{H(I_R, I_S)}
\]

Later, multivariate and high dimensional versions of MI and NMI (Gan and Chung, 2005; Meyer et al., 1999; Wang and Shen, 2006; Zhang and Rangarajan, 2005) were proposed, which extended the registration to three or more images and provided a faster convergence and better final accuracy. Although these MI variants have been popular, their performance degrades for some modalities. For example, in case of retinal image registration their performance is degraded due to the presence of innately challenging characteristics (Kubecka and Jan, 2004). These similarity measures had a disadvantage that they ignore the spatial information i.e. if the 2D histograms of registering images are the same, the images were assumed to be matched, which was not true if the intensity values are rearranged in a particular way (Rueckert et al., 2000).
This gave researchers an opportunity to improve the MI by incorporating spatial information in various ways. It was proposed to combine MI with connected region labelling in the past (Studholme et al., 1996a) and then HO-MI, (Rueckert et al., 2000) using co-occurrence information matrices of neighbouring intensities for better registration results.

GMI combined the gradient information with MI (Pluim et al., 2000a) which helps to align images even at lower sampling resolutions, achieving better global maximum and have very low interpolation induced local maxima. For a given 2D reference and sensed image pair \(I_R\) and \(I_S\), it computes two partial derivatives for each sample point \(x\) and \(x'\) in \(I_R\) and \(I_S\), forming together the gradient vector, calculated by convolving the image with the appropriate first derivatives of a Gaussian kernel having scale \(\psi\). Thus the angle \(\phi_{x,x'}(\psi)\) between the gradient vectors is defined by:

\[
\phi_{x,x'}(\psi) = \arccos \frac{\nabla x(\psi) \cdot \nabla x'(\psi)}{|\nabla x(\psi)| |\nabla x'(\psi)|}
\] (2.17)

Here, \(\nabla x(\psi)\) denotes the gradient vector at point \(x\) of scale \(\psi\) and \(| . |\) denotes the magnitude. Furthermore, since multimodal images can have different gradient value at best aligned position, a weighting function \(\kappa\) is incorporated as follows:

\[
\kappa(\phi) = \frac{\cos(2\phi) + 1}{2}
\] (2.18)

Thus, GMI between \(I_R\) and \(I_S\) can be defined as:

\[
GMI(I_R, I_S) = G(I_R, I_S) \, MI(I_R, I_S)
\] (2.19)
with the gradient term $G(I_R, I_S)$ as:

$$G(I_R, I_S) = \sum_{(x, x') \in (I_R \cap I_S)} \kappa(\phi_{x,x'}(\psi)) \min(\| \nabla x(\psi) \|, \| \nabla x'(\psi) \|)$$

(2.20)

The gradient term not only seeks to align locations of high gradient magnitude, but also aims for a similar orientation of the gradients at these locations. Although GMI perform better for low-resolution images and decrease interpolation-induced local minima, it underperforms in the presence of noise. In a similar approach, Holden et al. (Holden et al., 2004) used Gaussian scale space derivatives as local image structure to integrate another channel with MI to propose Multi-channel MI (MC-MI).

Russakoff et al. (Russakoff et al., 2004) proposed a novel extension to MI called RMI, which use local regions to compute MI and considers high dimensional distributions to be normally distributed. RMI takes neighbouring pixels into account to incorporate spatial information. Essentially, for each pixel, a vector of all the local intensities is created for $I_R$ and $I_S$. It provides much greater relational information for each pixel, it also means that many intensities need to be considered. Therefore, using a joint histogram to represent this information would lead to high computational cost. To overcome this complexity problem, the samples are replaced by a covariance matrix which substantially reduces the amount of data. The covariance matrix represents the relation between the original vector elements by approximating the joint intensities by a normal distribution.

Considering a normally distributed $d$-dimensional distribution, the entropy $H_g$ can computed in terms of covariance matrix $\Sigma_d$ as:
\[ H_g(\Sigma_d) = \log((2\pi e)^\frac{d}{2} \det(\Sigma_d)^\frac{1}{2}) \]  

This is mathematically equivalent to transforming the points into a new basis, where each dimension is uncorrelated, projecting the data onto each of the \( d \) new axes, and summing the entropies of those \( d \) independent 1-dimensional distributions (Shannon, 1948). It allows RMI to offer a clear improvement over standard MI, although computing the original matrix can still be computationally demanding. This becomes even more apparent in difficult registration problems where a larger neighbourhood radius is required. Similarly, Tomazevic et al. (Tomazevic et al., 2004) contributed a Multi-feature MI (MF-MI). While they had a common advantage of approximating the distribution as normal, still the calculation burden was computationally expensive.

Various other hybrid approaches also combine different aspects of the aforementioned feature and intensity methods, as for example, by using extracted vascular structures (Chanwimaluang et al., 2006) together with spatial information in feature neighbourhood MI (FNMI) (Legg et al., 2008, 2009, 2015). These techniques use covariance matrices to reduce the data complexity instead of high-dimensional histograms, though as the spatial information increases, so commensurately does the corresponding computational cost. Other contributors have combined the spatial information in different ways but still the computation overhead is significant (Butz and Thiran, 2001; Chappelow and Madabhushi, 2010; Kang et al., 2011; Luan et al., 2005, 2008; Woo et al., 2015; Yang et al., 2006; Yi and Soatto, 2011; Zhang et al., 2011; Zheng, 2010; Zouqi et al., 2010).
Although MI variants have found numerous clinical applications, there are still many issues which pose challenges (Pluim et al., 2003). As discussed in Section 2.2, the presence of intensity variations due to variety of reasons poses a problem in multimodal imaging, which includes INU and the radio frequency inhomogeneity in MRI imaging. While MI can also be computed locally, i.e. within the image by dividing it into sub images and then performing registration, but then the significant problem is that with the decreased image size, the statistical power of MI is reduced drastically. To improve registration in such cases images are registered rigidly using global MI and then locally in sub images using local MI (Maintz et al., 1998) and interpolated finally using Thin-Plate Splines, at the cost of computational complexity. Similar results are also reported using HO-MI (Denton et al., 1999; Rueckert et al., 2000, 1999), and other approaches which improve the MI performance by using varying number of bins, based on image characteristics instead of choosing fixed number of bins (Legg et al., 2007).

MI and its spatial information-based variants have also been used for non-rigid registration, giving the flexibility of local estimation of the joint histogram, by dividing the image into small regions (Garcia-Arteaga and Kybic, 2008; Russakoff et al., 2004; Zhang et al., 2011). However, as the subdivided image parts become too small, the small samples can limit the estimation performance of the local histogram. Several adaptations have been proposed to overcome this. Likar and Pernuš (2001) combined the local and global intensity distribution and Andronache et al. presented a local intensity remapping to allow the use of cross correlation as similarity measure in the smaller sub-images (Andronache et al., 2006, 2008). Weese et al. (1999) argued that, when the image parts are sufficiently small, they
will likely contain not more than two different structures that are rather homogeneous and hence used CC straight away, without any intensity remapping but Studholme et al. (2006) uses total correlation between the intensity joint histogram and a third channel representing a spatial label. Loeckx et al. considered the spatial coordinates of sample points in the reference image as a priori known conditions and proposed conditional MI (cMI) as similarity measure (Loeckx et al., 2007, 2010), however, it is found to be ten times slower than the traditional MI measures.

The generalised weighting scheme for spatial information encoding, was extended by Zhuang et al. which proposed a new registration method called as spatially encoded MI (SE-MI) (Zhuang et al., 2011, 2009, 2010). The weighting scheme is used to vary the contribution of pixels to a set of joint histogram tables, which associate with a spatial variable. Similarly, multidimensional self-similarity-based descriptor (Heinrich et al., 2012) and an organ focused MI (OF-MI) are reported, which need pre-processed segmentation of organ to be registered (Fernandez-de Manuel et al., 2014). Recently, two of the popular MI variants, namely GMI and RMI have been extensively employed for different medical registration applications as reported in (Jin et al., 2013; Yang and Fu, 2012; Yang et al., 2015) and (Chen and Lin, 2011; Ng et al., 2015; Pradhan and Patra, 2013, 2015), respectively.

In the past, many dimensionality reduction techniques have been utilised in the registration process. For example, RMI uses a covariance matrix for the fast computation of entropies instead of a high dimensional histogram to provide a rough estimate but as neighbourhood region size increases it becomes computationally intensive. Conventional Principal Component Analysis (PCA) being a versatile
multivariate analysis tool has also been popularly deployed for faster registration (Leng et al., 2015; Lu and Chen, 2007) due to its ability to linearly compress and preserve the randomness of the higher dimensional space with fewer components. But as the dimensionality increases, its covariance-based computation is also high. This led to the development of other methods which do not involve computation of covariance matrix and can be classed as either power iteration or competitive learning methods (Roweis, 1998). The power iteration method such as Arnoldi (Saad, 2011) iteratively update their eigenvector estimates, through repeated multiplication with the matrix to be diagonalised, while competitive learning methods like Sanger (Hancock et al., 1992) find the principal subspace by setting a learning rate parameter.

EMPCA is another effective approach (Roweis, 1998), which iteratively estimates the dominant principal components using an expectation (e-step) and subsequent maximisation step (m-step). It avoids the computation of entire covariance matrix and effectively finds PCA components with fewer iterations than power iteration and competitive learning methods (Roweis, 1998).

Another popular approach used with PCA computation involves selecting the best subset of significant principal components instead of using all PCA components. Methods like Scree Graph, Broken–stick and their variants (Cangelosi and Goriely, 2007; Jolliffe, 2002) choose a subset by computing the cumulative variance of all components. In contrast, the Kaiser Rule (Jackson, 1993; Jolliffe, 2002) retains only those principal components whose eigenvalues are greater than one. Some surveys on the similarity measures also exist, which discuss and compare them in terms of their strengths and applications (Klein et al., 2009; Penney
et al., 1998; Rogelj, 2001; Rogelj and Kovačič, 2003; Skerl et al., 2006, 2008; Wu et al., 2009). Next section provides a discussion on the existing similarity measures and highlight their limitations.

2.5 Discussion

As discussed in the previous section, the registration performance was shown to clearly be dependent on the appropriate choices made in each step, with a complex interdependency existing between different steps of registration, in addition to other challenging factors such as multimodality, interpolation artefacts and the presence of INU and homogeneous content in clinical images (Woo et al., 2015; Zhuang et al., 2011). It is evident the similarity measure used to quantify the level of alignment is one of the prime factors influencing the overall quality of the registration.

In addition to the popular use of MI and its variants as a similarity measure in multimodal MIR, they have also found applications in many other domains including remote sensing and pattern recognition (Argyriou et al., 2015). Since MI computes the statistical relationship between different modalities, it is an appropriate choice for remote sensing image registration and assessment (Chen et al., 2003).

Recent work also shows the prominence of MI in the remote sensing domain, with (Yao et al., 2010) proposing MI for feature selection in high dimensional hyper-spectral imagery and claiming it performs better than conventional classific-
2.5 Discussion

Wang and Jin (2012) has also recently implemented a range of MI variants for analysing post-earthquake building damage using pre-event IKONOS optical images and post-event Radarsat-2 and COSMO-SkyMed synthetic aperture radar images. Similarly, MI has also been applied in the field of pattern recognition and machine learning for dimensionality reduction by feature extraction and selection (Wu and Murphy, 2010).

These aforementioned applications of MI are not exhaustive, but they do showcase its versatility as a statistical measure which is especially relevant to the multimodal image registration challenges this thesis addresses.

Table 2.1 provides a comparative summary, presented in chronological order of existing similarity measures in terms of their appropriateness for mono and multimodal registration, characteristic type and computational efficiency. It highlights their drawbacks particularly reflecting their sensitivity to illumination, segmentation quality, use of spatial information and computational cost.

Broadly, the intensity and feature-based similarity measures are sensitive to both noise and the quality of segmentation, with their performance being further degraded in clinical images containing INU and large homogeneous regions. Although, incorporating local spatial information into intensity-based similarity measures to develop hybrid solutions appears promising for multimodal registration, the associated high dimensionality leads to large computational overheads (Goshtasby, 2012). Clearly, the computational efficiency is further impacted when dealing with datasets which include the aforementioned artefacts leading to slower convergence (Oliveira and Tavares, 2014).
### Table 2.1: Comparative summary of major existing similarity measures.

Overall, these observations conclusively confirm the need for a new hybrid similarity measure which combines local regional features with MI to enable the registration of challenging multimodal clinical images while at the same time incurring low computational cost. Since iterative EMPCA offers speed benefits over conventional dimensionality reduction mechanisms, this was the motivation to investigate whether an efficient hybrid EMPCA-based similarity measure could be...
framed without compromising registration performance. This was the context for the overarching research question formulated in Section 1.4 and from which the new EMPCA-MI similarity measure framework presented in this thesis was developed as detailed in Chapters 4, 6 and 7.

Before presenting the constituent components of this framework, the next chapter describes the research methodology adopted.

### 2.6 Summary

This chapter discussed the historical background and developments in the medical imaging domain in brief. It highlighted the need of multimodal registration and an insight into the various challenges of the process. It also presents an elaborative literature review of the various steps involved in the registration process, which include different transformations, interpolations, similarity measures and optimisation methods. Since MI-based similarity measures have been most popular in the MIR domain, a review of them has been presented. Finally, a comparative summary table is presented for the limitations of major existing similarity measures. The next chapter will focus on the research methodology adopted for this thesis.
Chapter 3

Research Methodology

3.1 Introduction

This chapter presents the research methodology adopted to design, develop as well as test and critically evaluate the new hybrid EMPCA-MI similarity measure framework. The methodology is based on the generic MIR model (Oliveira and Tavares, 2014), to critically analyse the robustness of the framework in a variety of practical scenarios, including different levels of misregistration along with the presence of modality artefacts and noise. This enables an equitable performance comparison of the new framework by assessing its impact both qualitatively and quantitatively for a given reference and sensed medical image pair (Khalifa et al., 2011; Zitová and Flusser, 2003). The research methodology utilises images from a variety of mono and multimodal clinical datasets.
Medical image registration is a challenging and computationally intensive problem, which has led to its methodological design and development on both dedicated hardware-based prototype and computer-based simulations (Eklund et al., 2013; Fluck et al., 2011; Neri et al., 2008). As discussed in Section 2.3, a software implementation offers greater flexibility of rapid development/verification cycle and is more cost effective in comparison to a hardware-based solution. Hence, a software simulation-based test-bed was adopted for developmental and experimental purposes.

Figure 3.1 shows the high-level block diagram of both the methodology and test bed used. It highlights the generic MIR model which uses various inputs and parameter settings; a set of predefined performance metrics for evaluation and a final registration output of a pair of medical images which are spatially aligned and have a minimum registration error. The various phases of the adopted research methodology are summarised in following steps:

1. Critical review of the MIR literature and narrow the research focus to similarity measure in the registration process.

2. Implement an established generic MIR model which will be used as an underlying building block for the evaluation process.

3. Identify the gaps and limitations of existing similarity measures. Iteratively test and develop new ideas for new hybrid EMPCA-MI similarity measure framework which can close the gap with the objective to improve registration
performance with respect to the challenging conditions highlighted in Section 1.3. This process comprises the development cycle as shown in Figure 3.1, which includes design, development followed by testing and evaluation.

4. Based on the generic MIR model, implement the EMPCA-MI similarity
measure framework (step 3) in the simulation environment using a reference and sensed multimodal medical image pair from the image dataset and chosen set of simulation specific parameter settings.

5. Setup a new scenario experiment to validate the level of robustness of the EMPCA-MI similarity measure framework in terms of registration performance for a given level of misregistration and noise artefacts.

6. Perform the registration simulation as in Figure 2.6 and save the final registration results using (2.1).

7. Critically evaluate the registration quality using appropriate performance metrics.

8. Repeat steps 5-7 until all scenarios have been considered.

9. Analyse the new similarity measure results for a range of scenario experiments and compare with the performance of existing similarity measures.

These steps provide the necessary rigour to critically evaluate the EMPCA-MI similarity measure framework. The next section will consider the registration model set-up.

3.3 Generic MIR Model Setup

As discussed in Section 2.3, most of the literature uses a generic registration model for the iterative MIR process which comprises of various steps as shown in Figure 2.5. Since the generic model has been rigorously used in the past for a variety of
medical registration applications (Oliveira and Tavares, 2014), it was considered as
the primary choice for the deployment of the new EMPCA-MI similarity measure
framework. The transformation assumed for the registration model is two dimen-
sional rigid (Brown, 1992) having $t_x$ and $t_y$ translations along the $x$ and $y$ axes,
and rotational angle $\theta$ for all registrations, except the retinal image datasets, where
for the reasons given in Section 2.3.1, a similarity transformation with uniform
scaling $S$ (Gharabaghi et al., 2013; Legg et al., 2008, 2009) is used. For interpol-
ation, in the past various methods have been applied (Zitová and Flusser, 2003)
as discussed in Section 2.3.2. Of these, bicubic, bilinear and nearest neighbour
interpolations were chosen for the generic registration model since they have been
shown to produce minimal artefacts compared to other techniques (Acharya and
Tsai, 2007; Pluim et al., 2000b) and are computationally faster than higher-order
interpolation methods (Thevenaz et al., 2000).

As the behaviour of an interpolator is also dependent on the similarity mea-
ure, a rigorous analysis and investigation using these interpolators within the new
hybrid EMPCA-MI similarity measure will be undertaken in Chapter 4. The most
appropriate interpolation method emerging in this analysis will be then used in the
subsequent chapters. Furthermore, due to the overall faster and better convergence
performance, the Powell Brent optimisation (Press et al., 2007) has been adopted
for the registration model. Table 3.1 shows the different parameter settings and
their empirical selections used in the registration model and these remain fixed
throughout the experimental analysis unless otherwise stated.
3.4 Image Datasets

A range of mono and multimodal clinical datasets are chosen for the evaluation of the EMPCA-MI similarity measure framework for the registration experiments. Table 3.2 shows the details of the image datasets.

### 3.4.1 Mono and Multimodal BrainWeb Dataset

The *BrainWeb* Database (Montreal Neurological Institute and Hospital, nd) comprises multimodal MRI T1, T2 and PD brain images which have been synthesised from a digital phantom (Collins et al., 1998). The *BrainWeb* images are defined at a 1 mm isotropic voxel grid in Talairach space, with dimensions as shown in Table 3.2. They have been used extensively to study the performance of MIR methods (Gholipour et al., 2007; Razlighi et al., 2013) as well as in other applications (Khademi et al., 2014; Nordenskjöld et al., 2014). The main advantages of using this dataset are: i) the ground truth is known a priori, and ii) the imaging parameters can be independently controlled. Since the source for simulation of

<table>
<thead>
<tr>
<th>Generic MIR Model</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Parameter Setting</strong></td>
</tr>
<tr>
<td>Transformation</td>
</tr>
<tr>
<td>Interpolation</td>
</tr>
<tr>
<td>Optimisation</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

Table 3.1: Parameter settings for the generic MIR model.
3.4 Image Datasets

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Modality</th>
<th>Resolution (pixels)</th>
<th>INU(α)</th>
<th>Noise(β)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BrainWeb</td>
<td>MRI T1 (T1)</td>
<td>[181×217×181]</td>
<td>α_{20}=20% INU,</td>
<td>Gaussian</td>
</tr>
<tr>
<td></td>
<td>MRI T2 (T2)</td>
<td></td>
<td>α_{40}=40% INU</td>
<td>(µ=0.01,</td>
</tr>
<tr>
<td></td>
<td>MRI PD (PD)</td>
<td></td>
<td></td>
<td>σ^2=0.01)</td>
</tr>
<tr>
<td>DRIVE</td>
<td>Fundus</td>
<td>[768 × 584]</td>
<td>INU and noise are not present in these images</td>
<td>though they are characterised by having low</td>
</tr>
<tr>
<td>Brno-Mono</td>
<td>Fundus</td>
<td>[3504 × 2336]</td>
<td>contrast, non-uniformly illuminated</td>
<td>homogeneous regions &amp; different pathologies.</td>
</tr>
<tr>
<td>Brno-Multi</td>
<td>Fundus</td>
<td>[3888 × 2592]</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SLO</td>
<td>[768 × 768]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3.2: Various clinical image dataset details (Kolar et al., 2013; Kolar and Tasevsky, 2010; Montreal Neurological Institute and Hospital, 2000, nd).

the functional images is the same digital phantom, one has a systematic means of establishing a gold standard for registration and control over the level of image degradation for all modalities (Collins et al., 1998). For these reasons, this particular database was used for the experimental work and will be referred throughout as BrainWeb. Figures 3.2(a) – (c) show a sample set of MRI T1, T2 and PD images of the BrainWeb dataset. In order to test the robustness of the EMPCA-MI similarity measure framework, both INU and Gaussian noise-based artefacts were introduced into the BrainWeb dataset as will be discussed in Section 3.5.2.

3.4.2 Mono and Multimodal Retinal Datasets

As mentioned in Section 2.2, retinal images represent very challenging cases for MIR due to the presence of low contrasted homogeneous and non-uniformly illuminated regions. The following two mono and multimodal retinal databases were used for the experimental work:
3.4 Image Datasets

Figure 3.2: Example BrainWeb images of brain MRI (a) T1, (b) T2 and (c) PD from (Montreal Neurological Institute and Hospital, nd).

DRIVE Retinal Database

*Digital Retinal Images for Vessel Extraction* (DRIVE) (Image Sciences Institute, nd) is a clinical database of monomodal retinal images from a diabetic retinopathy screening program in The Netherlands (Staal et al., 2004). This database comprises of twenty colour retinal images acquired with 45° field of view, using a Canon CR5 non-mydriatic 3CCD camera. *DRIVE* database was initially established to enable comparative studies on the segmentation of blood vessels in retinal images (Fraz et al., 2012), but has recently been more widely adopted for retinal image registration (Gharabaghi et al., 2013; Patankar and Kulkarni, 2015), due to the lack of publicly accessible retinal datasets (Karnowski et al., 2013; Sim et al.,
3.4 Image Datasets

Figure 3.3: Example retinal images (a) and (b) of monomodal colour Fundus retinal images from the DRIVE dataset (Image Sciences Institute, nd) and (c) colour Fundus, (d) SLO from Brno-Multi dataset (Kolar et al., 2013; Kolar and Tasevsky, 2010).

This database is denoted as the DRIVE dataset in this thesis, with Figures 3.3(a) – (b) showing a sample set of retinal images from this dataset.

Brno Retinal Database

This retinal database is maintained by Brno University of Technology and comprises both mono and multimodal retinal datasets (Kolar et al., 2013; Kolar and Tasevsky, 2010), which are denoted throughout this thesis, as Brno-Mono and Brno-Multi respectively. The Brno-Mono dataset comprises 44 colour Fundus im-
age pairs. Each colour Fundus image was acquired with 60° field of view using Canon CF-60UV with digital camera Canon EOS 20D. The *Brno-Multi* dataset comprises 10 pairs of colour Fundus and grayscale SLO images. Each colour Fundus image of this dataset was acquired using a Canon CR-1, while the SLO images were obtained from a SLO system from Spectralis, Heidelberg Engineering. This database is not publicly available and was made available upon request. Figures 3.3(c) – (d) show a set of retinal images from the *Brno-Multi* dataset.

### 3.5 Robustness Scenarios

To validate the robustness of the registration performance for the proposed similarity measure, various scenarios were introduced. Ideally a robust similarity measure must be stable towards perturbations in the condition of input medical images. For example, different level of misalignment and the presence of noise/INU should not significantly affect the final registration, though in practice, such scenarios tend to test the limitations of a similarity measure and the registration model as a whole, depending on its clinical applicability.

Existing scenarios as in (Argyriou et al., 2015; Brock, 2013; Hajnal and Hill, 2001; Zitová and Flusser, 2003), are deployed for testing the robustness of the new EMPCA-MI similarity measure, which include: varying the level of misalignment between the reference and sensed image pair; introducing known artefacts such as INU and noise, and using image pairs which have large homogeneous non-vascular regions. The following sections will discuss each of these scenarios in detail:
3.5 Robustness Scenarios

3.5.1 Initial Misregistration

The registration process spatially aligns a reference and sensed medical image pair, which are misregistered. The level of misregistration present can depend on various reasons such as the medical imaging equipment, its acquisition times, inter/intra patient imaging and clinical application (Hajnal and Hill, 2001). Ideally, the registration model is supposed to register an image pair having any level of misregistration, but in practice the result is highly dependent on the initial misregistration since this governs the amount of initial overlap between the images (Studholme et al., 1994).

A robust similarity measure should be able to overcome this dependency by consistently providing lower registration error (RE) even when the level of misregistration in the image pair is increased (Modersitzki, 2004). For this reason, most literature makes use of a consistent methodology of introducing different levels of random, but known initial misalignment to the reference image $I_R$ before beginning the registration process (Russakoff et al., 2003; Wiest-Daesslé et al., 2007). In the experiments, the initial misregistration transformation parameters are defined as $\tau_{\text{initial}}$, and this is used to evaluate and compare the robustness of the new EMPCA-MI similarity measure with existing similarity measures. The $\tau_{\text{initial}}$ values for the registration of both BrainWeb and retinal image datasets were adopted from (Kotsas and Dodd, 2011) and (Gharabaghi et al., 2013) respectively. Since a known ground truth between $I_R$ and $I_S$ exists, the initial RE is then computed as the mean of Euclidean distances between the four corner points of $\tau_{\text{initial}}(I_R)$ and $I_S$ in pixels (Beutel, 2000; Hajnal and Hill, 2001).
The twenty initial misregistration scenarios for BrainWeb dataset as adopted from (Kotsas and Dodd, 2011) varies with initial RE=18.59 pixels (for 1st initial misregistration scenario ($-5.55, 17.74, 0.35^\circ$)) to 109.25 pixels (for 20th initial misregistration scenario ($0.55, 29.63, -44.7^\circ$)). Similarly, five initial misregistration scenarios were used from (Chen et al., 2011) and (Gharabaghi et al., 2013) for the retinal datasets. These different initial misregistration scenarios represent the initial misalignments, which occur in practice. Tables A.1 and A.2 in Appendix A show the various $\tau_{initial}$ used throughout this thesis together with the corresponding initial RE for the BrainWeb and retinal datasets respectively. A larger value of initial RE emphasises that the image pair is more misaligned and will prove to be more challenging for registration.

3.5.2 INU and Gaussian Noise

Another challenging condition which impacts the performance of registration process is the presence of INU in multimodal medical images, as already mentioned in Section 2.2. According to the radio frequency mapping theory (Insko and Bolinger, 1993; Stollberger and Wach, 1996), INU can be modelled as a multiplicative process (Ganzetti et al., 2016; Sled, 2015). The majority of previous studies including (Belaroussi et al., 2006; Dawant et al., 1993; Pham and Prince, 1999; Wells et al., 1995), have all formalised the corrupting influence of INU in MRI images as follows:

$$a(x, y) = u(x, y) \alpha(x, y) + \beta(x, y) \quad (3.1)$$
3.5 Robustness Scenarios

Figure 3.4: Example BrainWeb image of (a) uncorrupted brain MRI T1 $u$, (b) multiplicative 20% INU $\alpha$, (c) additive Gaussian noise $\beta$ and (d) measured T1 image $a$ (Collins et al., 1998; Montreal Neurological Institute and Hospital, nd; Razlighi et al., 2013).

Where, $a$ is the measured MRI image as a function of spatial location $(x, y)$, $u$ is the uncorrupted image, $\alpha$ is the multiplicative artefact, and $\beta$ is the image noise which is Gaussian distributed and independent of $u$ (Belaroussi et al., 2006; Gudbjartsson and Patz, 1995; Henkelman, 1985). Figures 3.4(a) – (d) shows the uncorrupted MRI T1 image $u$, corresponding 20% INU $\alpha$ and noise $\beta$, and the resultant MRI image $a$. In order to test the robustness of any new similarity measure against such artefacts, INU is introduced into the MRI image pairs with levels of 20% and 40%, denoted by $\alpha_{20}$ and $\alpha_{40}$ respectively as detailed in (Collins et al., 1998), (Leondes, 2005), (Yi and Soatto, 2011) and (Razlighi et al., 2013). Also Gaussian noise denoted by $\beta$ having mean $\mu$ and variance $\sigma^2$ values is used, as shown in Table 3.2.
3.5 Robustness Scenarios

To clarify the nomenclature used, T1+$\alpha_{20}$ represents a MRI T1 image with 20% INU. Correspondingly T2+$\alpha_{40}+\beta$ represents a MRI T2 image with 40% INU and Gaussian noise. Figures 3.5(a) – (c) show a selection of misregistered images from the BrainWeb dataset, with different levels of INU and Gaussian noise combinations.

3.5.3 Low Contrast and Homogeneous Regions

As highlighted in Section 2.2, retinal images are challenging to register due to their innate features (Sanchez-Galeana et al., 2001), especially low-contrast ho-
3.5 Robustness Scenarios

mogeneous and non-uniformly illuminated regions (Marrugo and Millán, 2011), in addition to different pathological conditions like haemorrhages, retinal scars and clumping of the dark pigment (Saine and Tyler, 2002). All these characteristics are known to cause deterioration of the registration performance in many existing similarity measures (Legg et al., 2009) while some fail totally (Fang et al., 2005). For these reasons, the mono and multimodal retinal images are chosen in order to test the robustness of the hybrid EMPCA-MI similarity measure framework. Furthermore, for retinal images, registration is performed using only the green channel, since it has the highest contrast compared with the red and blue channels, which are saturated and contain acquisition noise (Chanwimaluang et al., 2006; Laliberte et al., 2003; Skokan et al., 2002). Figures 3.6(a) and (b) respectively show a sample retinal image from the Brno-Mono dataset of colour Fundus together with its green channel. Clearly the visibility of vascular regions is enhanced when only the green channel is considered, while no INU or Gaussian noise is added to any of the retinal images because these artefacts are not prominent during the acquisition process (Saine and Tyler, 2002).

Figure 3.6: Example retinal images from Brno-Mono dataset (a) colour Fundus and (b) only its green channel (shown in grayscale colour) from (Kolar et al., 2013; Kolar and Tasevsky, 2010).
3.6 Performance Metrics

These are pivotal in the evaluation and validation of an image registration process (Maintz and Viergever, 1998) by measuring the success of a registration outcome along with their computational cost in both a quantitative and qualitative manner. They also provide the insight for analysing the new EMPCA-MI similarity measure framework for experiments with the scenarios in Section 3.5. Despite the importance of this comparison, before deploying image registration for clinical use, there are no established criteria or standard methods available (D. Forsberg, 2013). This hiatus and inadequate validation practice has tended to impede the widespread usage of image registration in different clinical workflows. The following sections will review those performance metrics which have previously been used for quantitative and perceptual assessment of the registration process, along with the computational cost.

3.6.1 Quantitative Assessment

Various quantitative methods to evaluate the registration performance, including the target landmarks, embedded fiducial markers, transformation error or Euclidean distance between pair of known points in the reference and sensed images (Klein et al., 2009), (Murphy et al., 2011) and (Rohlfing, 2012). The applicability of these methods depends on different factors such as availability of a ground truth and whether the fiducial markers are surgically implanted prior to image acquisition. Of these, a distance-based numerical error is adopted as the ground truth between the image pair as it is available for all the datasets used in the experiments (Goshtasby, 2012; Hajnal and Hill, 2001).
Once the new hybrid EMPCA-MI similarity measure is maximised in (2.1), the final registration transformation parameters $\tau_{\text{final}}$ are saved and the final RE is computed as the mean of Euclidean distance between the four corner points of $\tau_{\text{initial}}(I_R)$ and $\tau_{\text{final}}(I_S)$ in pixels (Bernardes et al., 2014; Dame and Marchand, 2010; Legg et al., 2007; Tang et al., 2010; Yin et al., 2010). For any modality, the lower the final RE value, the better registered the image pair.

In order to display and critically analyse the final RE values for a group of registrations, a box and whisker plot is used. This is a descriptive statistical tool (Klein et al., 2007) chosen because it represents the spread and skewness of the RE by displaying the mean, median, lower and upper quartiles and outliers. RE values outside 1.5 times interquartile range from the lower and upper quartiles are termed as outliers (Brase and Brase, 1999). In addition, the number of outliers having RE values higher than a given limit are shown as out-of-range outliers on top of the corresponding box and whisker plot (Murphy et al., 2011).

### 3.6.2 Perceptual Assessment

The registration process performance is also qualitatively evaluated using visual quality assessment (Woods, 2009). A variety of image overlaying visualisation methods can be used including contour, coloured, Canny edge detection and checkerboard overlay (Bourland, 2012). In addition, some software-based dynamic graphic user interfaces use a sliding window method to reflect the correspondence between the reference and sensed image pair (Fedorov et al., 2012; Meyer, 2007).
This thesis utilises the \textit{Canny edge detection} and \textit{checkerboard} overlay method for perceptual evaluation of the registration process as they can effectively visualise the qualitative differences and degree of physical misalignment between the registering images.

The \textit{Canny edge detection} overlay method displays the whole image pair after the edge detection process, while the \textit{checkerboard} overlay divides both the image into square regions, with the diagonal squares displaying the regions of the same image. The \textit{checkerboard} overlay specifically suits retinal images since it enables close observation of vascular lines and their continuity (Kolar et al., 2013; Wilson and Laxminarayan, 2007). Figure 3.7(a) shows the result of multimodal T1+$\alpha_{40}$+$\beta$/T2 pair registration using \textit{Canny edge detection} overlay with an inset of zoomed in region showing the central lobes. This clearly emphasises the level of misregistration between the image pair. Similarly, Figure 3.7(b) shows an example of \textit{checkerboard} overlay after monomodal retinal Fundus registration, highlighting the evident discontinuity in vascular lines at the edges of the square regions (Gharabaghi et al., 2013).

### 3.6.3 Computational Cost

Another key challenge for MIR besides attaining lower RE, is to minimise the overall computation cost of registration (Eklund et al., 2013; Shekhar et al., 2003; Zitová and Flusser, 2003). This is the total computational time incurred for each registration component as shown earlier in Figure 2.5, which is clearly dependent on the different algorithms comprising the hybrid EMPCA-MI similarity measure.
framework. As registration is an iterative process along with, where appropriate, an order of time complexity analysis, the average time to compute one cycle of iteration of the EMPCA-MI framework during registration is also considered as a key performance metric. This is calculated by averaging the computational time for all iterations in each registration and is defined as the *average runtimes* (ART) in seconds.

### 3.7 Simulation Platform

A variety of programming languages and open-source coding platforms are available for designing and developing the MIR model (Goshtasby, 2012; Modersitzki, 2004). High level languages such as *C/C++* (Wang and Lai, 2008) and Java (Lyon, 1999) are compiler-based and require development of components and libraries for
each element, which can be time consuming in most cases. In contrast, open-source platforms like Python (Chityala and Pudipeddi, 2014) and ImageJ (Pascau and Mateos, 2013) have specific libraries for image processing applications.

A popular alternative is MATLAB® (Mathworks, nd) which is well accepted by the community because of its rapid prototyping capabilities. Although it is interpreter-based commercial software and has comparatively slower execution times, it comes with rich built-in functions, lucid error reporting and has specifically developed toolboxes for image and signal processing. The slow execution speeds can be compensated by choosing a high-end personal computer (PC). The toolboxes are written and rigorously tested, making them more compact and straightforward to use compared with C/C++ and Java code. Since the main focus is on having a development platform to design, test and evaluate the new hybrid EMPCA-MI similarity measure framework, rather than designing new tools, MATLAB was the pragmatic choice for all experimental work.

The technical specifications for the MATLAB software and high-end PC are detailed in Table 3.3. These were fixed throughout all the experimentation work to ensure consistent and accurate performance comparison of the ART metric.

3.8 Software Code Validation

Various code verification and validation techniques (Oberkampf and Roy, 2010) based on software life cycle model were used throughout the design and implementation of the EMPCA-MI similarity measure framework. These included both
### Table 3.3: Simulation platform specifications and their details.

<table>
<thead>
<tr>
<th>Software Platform</th>
<th>PC Specifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>MATLAB 7.12 R2011a</td>
<td><strong>Processor</strong> Intel® Core™ 2 Duo E7500 (2.93GHz, 1066MHz FSB, 3MB L2 Cache)</td>
</tr>
<tr>
<td></td>
<td><strong>RAM</strong> 3 GB, DDR2, 800MHz</td>
</tr>
<tr>
<td></td>
<td><strong>Hard Disk</strong> 160GB, 8MB Cache, 7200 RPM</td>
</tr>
<tr>
<td></td>
<td><strong>Operating System</strong> Ubuntu Release 10.04 (lucid)</td>
</tr>
<tr>
<td></td>
<td><strong>Kernel Linux 2.6.32-67-generic</strong></td>
</tr>
<tr>
<td></td>
<td><strong>GNOME 2.30.2</strong></td>
</tr>
</tbody>
</table>

static analysis checks, such as data type, control flow and data use, as well as dynamic checks including running the model with *Shepp-Logan* head phantom test image pair (Haidekker, 2011).

An advantage of using MATLAB was the considerably lower refinement time incurred between the development and static-dynamic test cycle in Figure 3.1. In the initial development phase, coding and design errors (like the exception handling, divide by zero along with checks for integer overflow and conditional dead logics) were trouble-shot.

The transformation and interpolation methods deployed in the registration model were available as MATLAB functions in its Image Processing Toolbox (Mathworks, nd). The Powell Brent optimisation method was implemented as in (Press et al., 2007) and then validated using the inverted version of standardised *first function of De Jong’s* benchmarking test (Molga and Smutnicki, 2005). This function is convex so has a global maximum which represents the ideal image regis-
tration scenario (Modersitzki, 2004; Zitová and Flusser, 2003). Furthermore, it has the ability to extend and simulate three and four variable maximisation representing the 2D rigid and similarity transformations respectively. This benchmarking test validated the functionality of Powell Brent optimisation method implementation which was then integrated into the registration model.

The software implementation of the EMPCA-MI similarity measure framework was verified as well as validated using two approaches:

1. By independently testing the functionality of each constituent component of the framework. For example, EMPCA iteratively computes the various PCA components, so the EMPCA (Roweis, 1998) implementation was tested with the conventional PCA function in the MATLAB Statistics Toolbox (Mathworks, nd) using a multivariate Iris dataset (Fisher, 1936). Similarly the MI code was based on (Cover and Thomas, 1991) and tested using the MI Least-dependent Component Analysis Toolbox (Stögbauer et al., 2004) for fixed test data. *Kaiser Rule* (Jackson, 1993; Jolliffe, 2002) and *Wichard* (Wichard et al., 2008) approach implementations were also validated using MATLAB Statistics Toolbox. These tests provide a higher degree of confidence to the software implementation as the functionality of each individual stage is evaluated independently.

2. To verify the software code of the registration model, dynamic testing was applied to the *Shepp-Logan* head phantom test image pair (see Figure 3.8) (Haidekker, 2011) in different misregistration configurations. This is employed in (Guo and Huang, 2008; Pilutti et al., 2014) as a model of the human
3.9 Summary

This chapter has described the methodology adopted for all experimental work in this thesis. A software simulation-based test-bed is used to implement the MIR model, which uses different test scenarios for robustness validation. A set of performance metrics have been selected to evaluate the similarity measures on clinical and general datasets. The choice of MATLAB as the main simulation platform for model implementation has been justified and details of the validation procedures for the code developed have been presented. The next chapter will present the first contribution, namely the EMPCA-MI similarity measure framework.

Figure 3.8: Shepp-Logan head phantom image (Guo and Huang, 2008; Haidekker, 2011)

head for the development and testing of image registration algorithms since it has a known ground truth. These experiments confirmed the convergent nature of the registration model as alignment of this test image was always achieved. Furthermore, the existing similarity measures used as comparators were publicly available. Details of the specific individual comparators used for validation will be given in the respective contribution chapters.
Chapter 4

EMPCA-MI for Medical Image Registration

4.1 Introduction

The performance of the registration for images of some anatomies deteriorates considerably, such as MRI images having intensity variations and noise due to radio frequency-based inhomogeneity and similarly retinal images which comprise inherently large homogeneous regions (Chanwimaluang et al., 2006; Simmons et al., 1994; Zhuang et al., 2010). It was evident from the literature in Chapter 2 that the similarity measure plays a pivotal role in the registration process. While MI has been a popular similarity measure for image registration, two of its recent variants were identified in Chapter 2, namely GMI and RMI as promising similarity measures. GMI and RMI combine intensity gradient information and local pixel regions respectively with MI to improve the image registration performance in
4.1 Introduction

comparison to traditional MI. However, GMI performance degrades in the presence of artefacts and noise, while RMI uses a covariance matrix instead of a high-dimensional histogram, which leads to reduced data complexity but generates a coarser approximation in certain cases.

In this chapter, a new similarity measure framework called EMPCA-MI is proposed for MIR. EMPCA-MI utilises the concept of neighbourhood region from the RMI algorithm to extract the spatial features with the help of principal component analysis. Unlike the approximation in RMI, EMPCA-MI achieves effective dimensionality reduction by using an iterative process to determine the dominant principal component. It uses only first principal component to extract the key features from a given neighbourhood radius of the image since it always has the highest variance. This evidently leads to a significant reduction in the computation cost.

This chapter also includes a thorough empirical EMPCA-MI analysis for the choice of interpolation and neighbourhood radius on mono and multimodal images. The rest of the chapter is organised as follows: The new EMPCA-MI framework is presented in Section 4.2. In Section 4.3, experimental results are conducted to analyse and choose the interpolation and radius of neighbourhood region from three perspectives i.e. smoothness of the similarity measure traces, computational cost involved and finally the quantitative performance in terms of RE. These choices will be used in the next chapter. Finally a summary of the chapter is provided in Section 4.4.
4.2 EMPCA-MI

In Chapter 2, various power iteration and competitive learning-based dimensionality reduction methods were discussed. From them, EMPCA was identified as an attractive solution since it requires fewer number of iterations to find the number of principal components required and can be deployed with lower computational overheads (Roweis, 1998). In the past, MI has been utilised to evaluate and select features for a diverse range of applications (Battiti, 1994; Hossain et al., 2011; Peng et al., 2005; Yao et al., 2010). Hence its statistical background and ability to perform on multimodal images makes it suitable for quantifying the level of alignment between the EMPCA-extracted features (Paragios et al., 2015). The EMPCA-MI framework combines the spatial features of neighbourhood region with MI using EMPCA in three sequential steps (Reel et al., 2012a,b). Step I comprises of rearranging the image data such that the neighbourhood region features can be effectively extracted. Step II employs EMPCA to iteratively compute the first dominant principal component. Finally Step III computes the mutual information between the features of reference and sensed images.

4.2.1 Step I: Image Neighbourhood Rearrangement

Considering an interpolated sensed image $I^*_S \in \mathbb{N}^{n \times m}$ with $n$ rows and $m$ columns, a neighbourhood region-based sliding window of size $(2r + 1) \times (2r + 1)$ pixels, using a given neighbourhood radius $r$ can be rearranged as a row vector $s \in \mathbb{N}^{1 \times d}$, where $d = (2r + 1)^2$. This can be further generalised as:
Here, $D_{\infty}(h, g)$ denotes the Chebyshev distance (Fernandez-Maloigne, 2012) of a given pixel location $h$ from the centre pixel $g$ of the neighbourhood region sliding window at $i^{th}$ row and $j^{th}$ column location. Also, $r \leq i \leq (n-r)$, $r \leq j \leq (m-r)$ and $1 < k < q$ with $q = (n-2r) \times (m-2r)$. Hence, all such $q$ neighbourhood region windows for the complete $I_S^*$, can be concatenated and defined as $Q_s \in \mathbb{N}^{q \times d}$, such that:

$$Q_s \triangleq \begin{bmatrix} S_1 \\ S_2 \\ \vdots \\ S_k \\ \vdots \\ S_q \end{bmatrix}$$

(4.3)

Here, $S_k$ denotes the $k^{th}$ neighbourhood region window. Here, $Q_s$ can be considered as a matrix consisting of $q$ observed data vectors in $d$ dimensional subspace. Forming this neighbourhood region reduces the artefact effects because the intensity variations across a localised region will be significantly lower compared to across the whole image. Since neighbourhood region pixels are more likely to be locally correlated compared to pixels located further apart, each neighbourhood region is rearranged as a $d$-dimensional vector, enabling the features to be extracted along with dimensionality reduction.
Figure 4.1: The neighbourhood region selection for feature extraction in EMPCA-MI, showing the dimensionality $d$ with radius (a) $r=1$, (b) $r=2$ and (c) $r=3$ pixels.

Figure 4.1 illustrates the neighbourhood region sliding window for $I_S^*$ with radius $r$ increasing from 1 to 3. Figure 4.1(a) shows using the central pixel of the neighbourhood (shown in dark blue colour) for which the neighbourhood region (shown in light blue colour) for radius $r=1$ is rearranged as $s$, having $d=9$. It is apparent $d$ increases with $r^2$, emphasising that a larger region is considered as a higher dimensionality for the feature extraction process. For example, in Figures 4.1(a) – (c), the value of $d$ increases from 9 to 49, as $r$ is increased from 1 to 3.

The overall impact of choosing a particular value of $r$ for EMPCA-MI computation will be rigorously analysed in Section 4.3.2.
4.2.2 **Step II: EMPCA-based feature extraction**

This step of EMPCA-MI comprises of dimensionality reduction of the neighbourhood region features. Although in *Chapter 2*, various dimensionality reduction methods were discussed but EMPCA was identified as an appropriate method due to its iterative nature and lower computational overheads instead of conventional PCA which computes all the components simultaneously. As mentioned in the literature, the sum of the variances of the principal components is equal to the variance of the original variables. While the first principal component accounts for most of the variability in the data (Rencher, 1998), various stopping rules such as Scree Graph and *Kaiser Rule* (Jackson, 1993; Jolliffe, 2002) are also used for determining the best $P$ principal components to be retained, which represent the high percentage of the variation in the data.

Hence, considering that $Q_s$ is a linear transformation of a latent variable $z \in \mathbb{N}^{P \times d}$, transformed by $\Phi \in \mathbb{N}^{d \times P}$ and additive Gaussian noise $v \in \mathbb{N}^{q \times d}$ (having a covariance matrix $V$) such that:

$$Q_s = \Phi z + v \quad z \sim N(0, 1) \quad v \sim N(0, V) \quad (4.4)$$

Here, $z$ is assumed to be independent and have unity variance. Also $v$ is independent of $z$ and is normal distributed. Therefore, EMPCA algorithm (Roweis, 1998) can be used for dimensionality reduction to extract the most prominent $P$ principal components (such that $P < d$). The expectation step estimates the unknown states $\Omega_s \in \mathbb{N}^{P \times d}$ and then choose $\Phi$ and the restricted $V$ in the maximisation step so as to maximise the expected joint likelihood of the estimated $\Omega_s$ and the observed $Q_s$. This can be written as:
4.2 EMPCA-MI

\[ e\text{-}step : \quad \Omega_s = (\Phi^T \Phi)^{-1} \Phi^T Q_s \quad (4.5) \]

\[ m\text{-}step : \quad \Phi^{new} = Q_s \Omega_s^T (\Omega_s \Omega_s^T)^{-1} \quad (4.6) \]

Here, \( Q_s \) is a matrix with dimensions \( q \times d \) for all neighbourhood regions as in (4.3). In (4.5), an orientation for the principal subspace is randomly initialised. Then considering this subspace as fixed, \( Q_s \) is projected onto it to get the values of the hidden states \( \Omega_s \). Then in (4.6), keeping the values of the hidden states fixed, choose the subspace orientation which minimises the squared reconstruction errors of the given data-points.

Rowies et al (Roweis, 1998) explains the above-mentioned two steps using a simple two-dimensional physical analogy. Assuming a rod pinned at the origin which is free to rotate, consider picking an orientation for the rod. Now, holding the rod still, project every data-point onto the rod, and attach each projected point to its original point with a spring. Next release the rod and repeat. The direction of the rod represents the guess of the principal component of the dataset. The energy stored in the springs is the reconstruction error which needs to be minimised.

Hence, from the final iteration, the first \( P \) principal components can be stored in \( X_s \in \mathbb{N}^{P \times q} \) from \( \Phi \). In a similar way, \( X_R \) can be computed from \( I_R \). EMPCA offers a significant saving in the computation cost in comparison to using conventional PCA for this step of EMPCA-MI, with latter computing all the principal components at once using the covariance matrix (Jolliffe, 2002). A discussion comparing the computational complexity between conventional PCA and EMPCA will follow later in Section 4.3.2.
4.2.3 **Step III**: Finding MI for Registration

This final step of EMPCA-MI aims to find the similarity between the extracted features of reference and interpolated sensed images. As discussed in Section 2.4, MI represents the amount of information one variable contains within another and hence is computed between \( X_S \) and \( X_R \). It assumes that a statistical relationship exists between \( X_S \) and \( X_R \), which is reflected by their marginal probabilities and joint probability. Hence, employing (2.7 - 2.15), the marginal entropies \( H(X_R) \) and \( H(X_S) \), as well as joint entropy \( H(X_R, X_S) \) are:

\[
H(X_R) = -\sum_{X_R} p_{X_R}(X_R) \log_2 p_{X_R}(X_R) \\
H(X_S) = -\sum_{X_S} p_{X_S}(X_S) \log_2 p_{X_S}(X_S) \\
H(X_R, X_S) = -\sum_{X_R, X_S} p_{I_RI_S}(X_R, X_S) \log_2 p_{X_RX_S}(X_R, X_S)
\]

and

\[
p_{X_R}(X_R) = \frac{U_{X_R}}{U} \tag{4.10}
\]

\[
p_{X_S}(X_S) = \frac{U_{X_S}}{U} \tag{4.11}
\]

\[
p_{X_RX_S}(X_R, X_S) = \frac{U_{X_RX_S}}{U} \tag{4.12}
\]

Hence, the MI between \( X_R \) and \( X_S \) can be expressed as:

\[
MI(X_R, X_S) = \sum_{X_R, X_S} p_{X_RX_S}(X_R, X_S) \log_2 \left\{ \frac{p_{X_RX_S}(X_R, X_S)}{p_{X_R}(X_R)p_{X_S}(X_S)} \right\} \tag{4.13}
\]
Here, \( p_{X_S}(X_S) \) and \( p_{X_R}(X_R) \) are the marginal probabilities for \( X_S \) and \( X_R \) respectively, and \( p_{X_RX_S}(X_R, X_S) \) represents their joint probability as defined in Section 2.4. These probability distributions are computed using \( b \) number of bins. The value of MI computed between the extracted features \( X_R \) and \( X_S \) will reflect the measure of alignment between \( I_R \) and \( I_S^* \), for a given set of transformation parameters. Hence an increasing value of MI for different sets of transformation parameters will signify the better spatial alignment. This will be more evident when the EMPCA-MI similarity measure framework is incorporated into the registration test-bed described in Chapter 3, which will use the optimisation process to maximise the EMPCA-MI to achieve the best alignment between given \( I_R \) and \( I_S^* \).

The corresponding pseudo-code of EMPCA-MI framework is given in Algorithm 4.1.

**Algorithm 4.1** Pseudo Code for EMPCA-MI similarity measure framework

**Inputs:** \( I_R \) and \( I_S^* \) each with spatial resolution \( m \times n \) pixels; \( r \) – neighbourhood radius; \( P \) – number of principal components.

**Variables:** \( d \)– dimensional space; \( q \)– total no. of pixels; \( Q_R, Q_S \)– output from rearranging \( I_R \) and \( I_S^* \); \( X_R, X_S \) – first principal component using EMPCA for \( Q_R \) and \( Q_S \).

**Output:** EMPCA-MI value

1: **Initialise** \( d = (2r + 1)^2 \) and \( q = (m - 2r) \times (n - 2r). \)
2: **Calculate** \( Q_R \) using (4.1) and (4.3) from \( I_R \).
3: **Calculate** \( X_R \) using (4.4 – 4.6) from \( Q_R \) for a given \( P \).
4: **REPEAT** Steps 2 to 3 for \( I_S^* \) to give \( X_S \).
5: **Calculate** MI(\( X_R, X_S \)) using (4.7– 4.13).
6: **STOP**
Figure 4.2: Illustration showing EMPCA-MI computation steps on reference $I_R$ and interpolated sensed $I_S^*$ images.

Figure 4.2 illustrates the various steps of EMPCA-MI with the help of example images having spatial dimensions $10 \times 10$ pixels. In Figure 4.2(a), the Step I uses neighbourhood radius $r=1$ and rearrange $I_R$ and $I_S^*$ as $Q_R$ and $Q_S$ respectively. Note the colour used for pixels represents the pre-processing step and does not represent actual image data. Figure 4.2(b) shows the Step II computation of extracting the most prominent, first principal component $X_R$ and $X_S$ using EMPCA. Finally, Figure 4.2(c) shows the Step III where the MI between them is computed according to (4.13).

A numerical example illustrating the calculation of EMPCA-MI between two matrices ($I_R$ and $I_S^*$) which represent image data having spatial dimensions $6 \times 6$ pixels, is as follows:

Let,
4.2 EMPCA-MI

\[ I_R = \begin{bmatrix}
132 & 140 & 133 & 129 & 124 & 127 \\
140 & 121 & 120 & 112 & 111 & 102 \\
122 & 118 & 98 & 96 & 81 & 80 \\
108 & 92 & 78 & 58 & 59 & 51 \\
84 & 64 & 45 & 42 & 35 & 51 \\
57 & 46 & 42 & 41 & 40 & 48 \\
\end{bmatrix} \quad \text{and} \quad
I_S^* = \begin{bmatrix}
120 & 116 & 112 & 106 & 99 & 98 \\
107 & 97 & 88 & 80 & 77 & 75 \\
85 & 66 & 58 & 54 & 52 & 59 \\
53 & 43 & 37 & 43 & 54 & 49 \\
43 & 41 & 40 & 45 & 45 & 44 \\
46 & 42 & 38 & 41 & 48 & 49 \\
\end{bmatrix}.

Where \( n = m = 6 \). Therefore, considering \( r = 1, \ d = 9 \) and \( q = 16 \).

Thus, in Step I using (4.1) and (4.3):

\[ Q_R = \begin{bmatrix}
132 & 140 & 133 & 140 & 121 & 120 & 122 & 118 & 98 \\
140 & 133 & 129 & 121 & 120 & 112 & 111 & 98 & 96 \\
133 & 129 & 124 & 120 & 112 & 111 & 98 & 96 & 81 \\
129 & 124 & 127 & 112 & 111 & 102 & 96 & 81 & 80 \\
140 & 121 & 120 & 122 & 118 & 98 & 108 & 92 & 78 \\
121 & 120 & 112 & 118 & 98 & 96 & 92 & 78 & 58 \\
120 & 112 & 111 & 98 & 96 & 81 & 78 & 58 & 59 \\
112 & 111 & 102 & 96 & 98 & 81 & 80 & 58 & 59 & 51 \\
122 & 118 & 98 & 108 & 92 & 78 & 84 & 64 & 45 \\
118 & 98 & 96 & 92 & 78 & 58 & 64 & 45 & 42 \\
98 & 96 & 81 & 78 & 58 & 59 & 45 & 42 & 35 \\
96 & 81 & 80 & 58 & 59 & 51 & 42 & 35 & 51 \\
108 & 92 & 78 & 84 & 64 & 45 & 57 & 46 & 42 \\
92 & 78 & 58 & 64 & 45 & 42 & 46 & 42 & 41 \\
78 & 58 & 59 & 45 & 42 & 35 & 42 & 41 & 40 \\
58 & 59 & 51 & 42 & 35 & 51 & 41 & 40 & 48 \\
\end{bmatrix} \]
And, in Step II, the corresponding values of $X_R$ and $X_S$ assuming $P = 1$ in $(4.4 - 4.6)$ are:

$$X_R = \begin{bmatrix} -0.1500 \\ -0.1895 \\ \vdots \\ -0.1498 \\ -0.0623 \end{bmatrix}_{16 \times 1} \quad \text{and} \quad X_S = \begin{bmatrix} -0.3462 \\ -0.3820 \\ \vdots \\ 0.0011 \\ -0.0132 \end{bmatrix}_{16 \times 1}$$

Finally, in Step III, choosing $b = 256$ for $(4.7 - 4.13)$, is $\text{MI}(X_R, X_S) \approx 2.6875$.

From the literature review in Chapter 2, the interpolation used for the transformation of the image during the registration impacts upon the performance of a similarity measure (Aljabar et al., 2005; Pluim et al., 2000b; Soleimani and
Khosravifard, 2011; Tsao, 2003). Similarly, as observed in (4.1), the underlying pre-processing stage of EMPCA-MI is dependent on the neighbourhood radius. Hence in order to critically assess the performance of the new EMPCA-MI framework, it is important that key parameter choices related to the interpolation and neighbourhood radius \( r \) are appropriately made.

### 4.3 Choosing interpolation and neighbourhood radius for EMPCA-MI

This section investigates the various choices available for interpolation and neighbourhood radius parameter settings and then evaluates the EMPCA-MI performance with the help of registration experiments. These experiments are conducted on the registration test-bed using the mono and multimodal pairs of MRI T1, T2 and PD human brain images from the \textit{BrainWeb} dataset (with \( \alpha_{20} \) and \( \beta \)) which was described in research methodology \textit{Chapter 3}.

During this analysis, to compute \textit{Step II} of EMPCA-MI a fixed value of \( P=1 \) is used in (4.5) and (4.6). This value of \( P \) is chosen because as mentioned before, it represents the first principal component which has the highest variance and contains most of the prominent information (Jolliffe, 2002). The value of \( P \) will be relaxed later in \textit{Chapter 7}, where it will be shown that EMPCA-MI can iteratively determine the best \( P \) from a dataset, instead of calculating all components using a covariance matrix. Similarly, in \textit{Step III} of EMPCA-MI, a fixed value of \( b=256 \) has been empirically chosen for the number of bins in (4.13), as in (Argyriou et al.,
2015; Hahn et al., 2010; Legg et al., 2007; Maintz et al., 1998). These fixed number of bins will be relaxed from Chapter 6 onwards, where its impact on registration performance is investigated in detail. Next, to evaluate the most appropriate interpolation strategy as well as the neighbourhood radius, three issues need to be considered:

1. The smoothness of EMPCA-MI as a cost function.

2. The computational cost.

3. The preliminary registration performance.

The next two sections describe the experiments and results for the choice of interpolation and then neighbourhood radius.

### 4.3.1 Choice of Interpolation Technique

The first series of experiments involved the process of determining the most appropriate interpolation method to be used to compute the EMPCA-MI similarity measure value for an image registration pair. As discussed in Chapter 2, the choice of interpolation for estimating intensity values of the transformed image has a significant impact on the registration algorithm in determining the final transformation value. Nearest neighbour, bilinear and bicubic methods have emerged as the most commonly used interpolation methods (Acharya and Tsai, 2007). Hence, these three interpolation schemes were examined as possible choices for the EMPCA-MI framework for image registration purposes using the above three criteria for assessment.
4.3 Choosing interpolation and neighbourhood radius for EMPCA-MI

Smoothness of EMPCA-MI plots

Each EMPCA-MI trace is obtained by shifting the sensed image along \( x \) and \( y \) axis translation and \( \theta \) rotation and keeping the reference image fixed. The EMPCA-MI values between the two images are then plotted. These traces emphasise that a global maximum is attained when the reference and sensed images are fully aligned and the EMPCA-MI values gradually decrease as higher values of shift are introduced.

For example, Figures 4.3 and 4.4 show the \( x \) and \( y \) axis translation EMPCA-MI traces using different interpolation methods for monomodal T2/T2 and multimodal PD/T1 image pairs respectively. The \( x \) axis translation plots in Figures 4.3(a) and (c) show both bilinear and bicubic interpolation methods give a similar and smooth EMPCA-MI output. On the other hand, a closer inspection in the zoomed-in view in Figure 4.3(a), shows the nearest neighbour method provides a non-smooth, stepwise output. A similar trend is also evident in the corresponding \( \theta \) rotation plot included in Figure B.2 of Appendix B. Furthermore, for \( y \) axis translation plots in Figure 4.3(b) zoomed-in view it is seen that bilinear method also has unsmooth behaviour which further aggravates in Figure 4.3(d). Furthermore, in Figure 4.4(b) a similar trend for the nearest neighbour and bilinear methods is seen for multimodal PD/T1 image pair. The bilinear method performance worsens further with the presence of many local maxima on both sides of global maximum, as seen in Figure 4.4(d) for translation along \( y \) axis. For completeness, Appendix B includes the EMPCA-MI traces for other mono and multimodal T1, T2 and PD image pairs. On closely examining the behaviour of these plots, bicubic inter-
4.3 Choosing interpolation and neighbourhood radius for EMPCA-MI

- Interpolation and neighbourhood radius for EMPCA-MI

Figure 4.3: EMPCA-MI Traces with nearest neighbour, bilinear and bicubic interpolation for \( x \) and \( y \) axis translation for monomodal T2/T2 image pairs (a & b) T2+\( \alpha \)/T2, (c & d) T2+\( \beta \)/T2 with inset showing zoomed-in view.

Interpolation is found to be the most effective interpolation method in terms of cost function smoothness. The better performance of bicubic method is due to the fact that it utilises a wider weighted average of \( 4 \times 4 \) pixels window, leading to more accurate pixel representation in the transformed sensed image. In contrast the nearest neighbour method gives a non-smooth, step-like trace resulting from the output pixels assuming the value of nearest neighbouring pixels in the input image. As neighbouring pixels will generally have similar values, it results in identical EMPCA-MI values for a range of small shifts. Similarly, the bilinear method intro-
4.3 Choosing interpolation and neighbourhood radius for EMPCA-MI

Figure 4.4: EMPCA-MI Traces with nearest neighbour, bilinear and bicubic interpolation for $x$ and $y$ axis translation for multimodal PD/T1 image pairs (a & b) PD+α$_{20}$/T1, (c & d) PD+β/T1.

The results indicate that bicubic interpolation produces smoother plots compared to the other methods, which may be due to the fact that it reduces artefacts that result in many local minima. Hence, bicubic method provides an overall smoother plot for EMPCA-MI compared to both nearest neighbour and bilinear methods.

**Computational Cost**

Figure 4.5 shows the ART performance for computing the EMPCA-MI value using nearest neighbour, bilinear and bicubic interpolation methods. It is apparent that
bicubic method is approximately 15% and 25% slower than bilinear and nearest neighbour methods respectively. Although, bicubic interpolation incurs higher computation time in order to determine the gradients and cross derivative at each re-sampled pixel location (Acharya and Tsai, 2007), but offers more accurate interpolated values and especially smoother EMPCA-MI traces.

The nearest neighbour method despite of being fastest, due to its step-like behaviour requires more overall registration time since the optimisation process does not converge effectively to the best transformation value. Similarly deploying bilinear for EMPCA-MI computation for registration will not be effective, since it will have a higher likelihood of obtaining a sub-optimal solution due to the optimisation process falling into false maxima instead of the true global maximum (Wang et al., 2008).

Figure 4.5: ART performance of EMPCA-MI using nearest neighbour, bilinear and bicubic interpolation methods for monomodal T2/T2 and multimodal PD/T1 image pairs.
4.3 Choosing interpolation and neighbourhood radius for EMPCA-MI

Preliminary Registration Performance

The preliminary registration performance is evaluated in terms of the RE for nearest neighbour, bilinear and bicubic interpolation methods in the EMPCA-MI framework. For this set of experiments, monomodal T2/T2 and multimodal PD/T1 image pairs having $\alpha_{20}$ and $\beta$ artefacts are used. As explained in the research methodology Chapter 3, 20 levels of initial misregistration will be used so there are in total 240 separate registration experiments performed for 2 modality pairs, 2 INU and Gaussian noise settings, 20 initial misregistrations and 3 interpolation methods. Figure 4.6(a) shows the RE vs. iteration plot for the registration of six monomodal T2/T2 pairs, starting with 1st initial misregistration scenario ($-5.55, 17.74, 0.35^\circ$). This initial misregistration scenario was chosen as in (Kotsas and Dodd, 2011), as already discussed in Section 3.5.1. Instead of only evaluating the final RE for each registration, RE vs. iteration plots is used. They give an insight into the registration process by highlighting the converging capability of a similarity measure.

In Figure 4.6(a), all six registrations start from a common initial RE and thereon RE is computed at the end of each Powell multidimensional direction set method iteration. As discussed in Sections 2.3.4 and 3.3, each iteration comprise of Brent-based line minimisation, with the process stopping either when the maximum number of iterations is completed or else when the tolerance for the line minimisation is achieved. The RE at the end of this process is denoted as final RE. For example, in Figure 4.6(a), T2+$\alpha_{20}$/T2 registrations using nearest neighbour, bilinear and bicubic interpolation start with an initial RE=18.6 pixels and finish with a final RE=1.03, 4.05 and 0.03 pixels respectively.
4.3 Choosing interpolation and neighbourhood radius for EMPCA-MI

Similarly, Figure 4.6 shows another set of registrations using 20th initial misregistration scenario (0.55, 29.63, −44.71°) having initial RE=109.20 pixels. Both plots show that bicubic interpolation-based registration outperforms the bilinear and nearest neighbour interpolations. It is clear that bicubic interpolation consistently leads to a lower final RE for both the initial misregistration scenarios. Especially striking is the monotonic rate of convergence in Figure 4.6(b) for bicubic interpolation, in contrast to bilinear which converges after only the 7th iteration.
but to a higher RE value. These results are congruent with the EMPCA-MI results discussed above. Since bicubic interpolation has smoother EMPCA-MI trace, it will allow the registration process to converge towards actual global maxima compared with the other interpolators. Figures 4.6(c) and (d) show the multimodal PD/T1 registrations with 1st and 20th initial misregistration scenario respectively. The nearest neighbour interpolation exhibits the poorest rate of convergence with interestingly in Figure 4.6(d) for PD+$\alpha_{20}/T1$ registration, the process terminates after only the 2nd iteration because the tolerance defined for the Powell-Brent optimisation is reached. While computationally being the fastest in terms of ART, this example highlights the inconsistent RE performance of the nearest neighbour interpolator.

Figures 4.7(a) and (b) show the detailed RE summary boxplots for monomodal T2/T2 and multimodal PD/T1 registration using the three interpolators with EMPCA-MI for all 20 initial misregistration scenarios. Each boxplot shows a bounding box defining the interquartile range with the bar across representing the median and whiskers denoting the RE range. The boxplots on top also include the number of out-of-range RE values corresponding to each interpolation method. The boxplots in Figure 4.7 reveal the overall better registration performance of bicubic interpolation which consistently achieves lower mean RE for all cases. Bicubic interpolation also exhibits the lowest number of out-of-range outliers, due to the fact that it has a steady converge as seen in Figures 4.6(a) – (d). On the other hand, nearest neighbour-based registrations lead to higher mean RE as well as more number of out-of-range outliers because of the inaccurate pixel values computed during the image transformation process during registration.
4.3 Choosing interpolation and neighbourhood radius for EMPCA-MI

In summarising, on the basis of above critical evaluation of EMPCA-MI traces, ART and image RE performance, bicubic interpolation was selected for the rest of the experimental work presented in this thesis.

4.3.2 Choice of neighbourhood radius

The impact of varying the neighbourhood radius for EMPCA-MI was studied next. These experiments were applied to determine the most appropriate neighbourhood radius to be used in EMPCA-MI for registration. As discussed in Section 4.2.1, the neighbourhood region selection uses Chebyshev distance radius $r$ to determine $Q_S$ having $q \times d$ dimensions.
4.3 Choosing interpolation and neighbourhood radius for EMPCA-MI

Figure 4.8: EMPCA-MI Traces with different radius values for x and y axis translation for monomodal T2/T2 image pairs (a & b) T2+ α_{20}/T2, (c & d) T2+ β/T2.

Smoothness of EMPCA-MI plots

In order to evaluate the neighbourhood radius, EMPCA-MI traces are generated as elaborated in the previous section. In this experimentation, the value of $r$ is varied between 1, 2, 3, 5, 7 and 10, keeping the interpolation scheme as bicubic. Figures 4.8 and 4.9 show the corresponding EMPCA-MI traces for different $r$ values for monomodal T2/T2 and multimodal PD/T1 image pairs respectively. As before, $\alpha_{20}$ and $\beta$ was introduced in these image pairs. The x axis translation plots in Figures 4.8(a) and (c) have a vertical symmetry in-comparison to the asymmetric y
4.3 Choosing interpolation and neighbourhood radius for EMPCA-MI

Figure 4.9: EMPCA-MI Traces with different radius values for $x$ and $y$ axis translation for multimodal PD/T1 image pairs (a & b) PD+$\alpha_{20}$/T1, (c & d) PD+$\beta$/T1.

The $x$ axis translation plots in Figures 4.8(b) and (d), highlighting the vertical symmetry in the anatomic structure of human brain shown earlier, in Figures 2.4(a) and (b).

Furthermore, the $x$ axis translation plots in Figures 4.8(a) and (c) show smoother EMPCA-MI output at higher radius values. A similar trend is observed for the $y$ axis translation in Figures 4.8(b) and (d), and $\theta$ rotation plots included in Figure C.2 in Appendix C respectively. But on close inspection, the $y$ axis translation plot in Figures 4.8(b) and (d) highlight that while EMPCA-MI output value increases on choosing higher radius also some local maxima are generated around the global
maximum. These are due to the fact that as the $\alpha_20$ and $\beta$ is introduced, their impact on the EMPCA-MI value is higher as the neighbourhood region is larger. The occurrence of local maxima is more adverse in the $y$ axis translation plots of multimodal PD/T1 image pair as shown in Figures 4.9(b) and (d). These local maxima become more pronounced as $r$ increases, with an analogous trend being seen in the other EMPCA-MI traces for various mono and multimodal T1, T2 and PD image pairs presented in Appendix C for completeness. The presence of local maxima around the global maximum is a well-documented problem in image registration (Hajnal and Hill, 2001), which can undermine the registration performance during optimisation. As a consequence, in the experiments using a lower radius value for EMPCA-MI will give a greater likelihood of convergence to the lowest RE value in comparison to using higher $r$ values.

Computational Cost

To evaluate the impact upon computation complexity of varying the neighbourhood radius, it is vital to understand the relationship between the radius and dimensionality $d$. As mentioned in Section 4.2.1, $d$ is dependent only on the choice of $r$. Consequently, $d$ will be the same for all images irrespective of their resolution, for a given fixed neighbourhood radius. Therefore, increasing $r$ from 1 to 10, will lead to quadratic increase of $d$, from 9 to 441, which will clearly have a commensurate effect on the ART performance. Furthermore, Figure 4.10 shows a semi-log plot for number of samples $q$ as the function of $r$ for all the datasets used in the experimental analysis. It is apparent from (4.2) that $q$ is also dependent on the image resolution in addition to $r$. For example, as $r$ increases from 1 to 10, the
4.3 Choosing interpolation and neighbourhood radius for EMPCA-MI

Figure 4.10: Number of samples ($q$) with respect to different neighbourhood radii for the variety of datasets.

corresponding number of samples in the BrainWeb dataset (consisting of images with resolution $181 \times 217$ pixels) decrease by approximately 20\% (from $3.8 \times 10^4$ to $3.1 \times 10^4$), while for Brno-Multi dataset (consisting of images with resolution $3888 \times 2592$ pixels) they reduce by 11\% (from $1.0 \times 10^7$ to $0.9 \times 10^7$). Hence for a fixed dataset, $q$ decreases as $r$ increases.

Figure 4.11 shows the ART performance for EMPCA-MI computation along with the conventional PCA with MI calculation (PCA-MI), for different radii values of $r$, using a log-log plot. It highlights the faster performance of the EMPCA-MI in comparison to the conventional PCA-MI. For example when $r=2$, EMPCA-MI is approximately five times faster than conventional PCA-MI, while in terms of order of complexity, the $Q_S$ matrix in Step II of the EMPCA-MI (discussed in Section 4.2.2) has dimensions $q \times d$ so for the conventional PCA-MI, $O(qd^2)$ operations are required to compute the covariance matrix and determine the first PCA component.
In contrast, EMPCA-MI which does not need to compute the covariance matrix, so only requires $O(qd)$ operations. Hence, for different $r$ values and corresponding $d$ and $q$ values, EMPCA-MI requires lower computation time than conventional PCA-MI.

**Preliminary Registration Performance**

Now, the quantitative registration performance in terms of RE is analysed for various $r$ values with EMPCA-MI for monomodal T2/T2 and multimodal PD/T1 BrainWeb image pairs having $\alpha_{20}$ and $\beta$ artefacts. As in Section 4.3.1 where the registration experimentation for the choice of interpolation were conducted, the same 20 levels of initial misregistration are used now for these experiments. Hence, in all 480 separate registration experiments are performed, consisting of 2 modalities pairs, 2 INU and Gaussian noise settings, 20 initial misregistration and 6 neighbourhood radii settings.
4.3 Choosing interpolation and neighbourhood radius for EMPCA-MI

Figure 4.12: RE versus iterations plots for EMPCA-MI with varying neighbourhood radius for monomodal (a) T2+α20/T2 pair (using 12th initial misregistration scenario) and (b) T2+β/T2 pair (using 16th initial misregistration scenario). Multimodal (c) PD+α20/T1 pair (using 3rd initial misregistration scenario) and (d) PD+β/T1 pair (using 17th initial misregistration scenario).

Figure 4.12 consists of four RE vs. iteration plots with each one having six registration experiments using EMPCA-MI with increasing neighbourhood radius $r=1, 2, 3, 5, 7$ and $10$. Figures 4.12(a) and (b) shows monomodal results for T2+α20/T2 pair, starting with the 12th initial misregistration scenario ($-19.11, 9.08, -28.42^\circ$) and T2+β/T2 pair, starting with 16th initial misregistration scenario ($-5.65, 16.94, -36.06^\circ$) respectively.
4.3 Choosing interpolation and neighbourhood radius for EMPCA-MI

Similarly, Figures 4.12(c) and (d) show multimodal results for PD+α_{20}/T1 pairs, starting with 3rd initial misregistration scenario (10.11, 27.22, -2.67°) and PD+β/T1 pair, starting with 17th initial misregistration scenario (0.21, -16.41, -36.15°) respectively. All the above initial misregistration scenarios were chosen as in (Kotsas and Dodd, 2011) as already described in the research methodology Chapter 3.

The results confirm that r=1 has the best rate of convergence as well as a lower RE in comparison to all other neighbourhood radii for both mono and multimodal pairs. Choosing r=2 also provide good convergence and lower RE, except in case of the monomodal T2+α_{20}/T2 registration in Figure 4.12(a) where the optimisation converges at a local maxima leading to high final RE. However, as the neighbourhood radius increases, the corresponding registrations all provide higher RE allied with poorer rates of convergence, incurring more iterations to achieve their final RE value. This is evident in the cases of T2+β/T2 and PD+α_{20}/T1 registrations for r=3, 5, and 7 shown in Figures 4.12(b) and (c) respectively.

Interestingly at r=10, only multimodal pairs in Figures 4.12(c) and (d) converged but at a considerably higher RE. On the other hand, in the case of monomodal registrations in Figures 4.12(a) and (b), the iterations end prematurely before the maximum number of iterations are completed with final RE even higher than the initial RE. This happens due to the convergence of the registration process to the false local maxima present near the initial misregistration. Hence, increasing r decreases the level of granularity needed for best feature extraction, which generates local maxima as seen previously on examining the smoothness of the EMPCA-MI plots.
4.3 Choosing interpolation and neighbourhood radius for EMPCA-MI

Figure 4.13: RE summary boxplot of (a) monomodal T2/T2 registration and (b) multimodal PD/T1 registration using varying neighbourhood radius with EMPCA-MI for all 20 initial misregistration scenarios. The mean and outliers are denoted by ♦ and + shapes. The numbers on top of the plots refer to the number of out-of-range outliers with RE higher than 140 and 200 pixels respectively.

Figures 4.13(a) and (b) show the detailed RE summary boxplot results for registration using the varying $r$ values with EMPCA-MI for all 20 initial misregistration scenarios for monomodal T2/T2 and multimodal PD/T1 respectively. These boxplots confirm the improved registration performance at $r=1$ of EMPCA-MI with the lowest mean RE and no out-of-range outliers. But as $r$ increases, the
registration performance deteriorates as mean RE increasing progressively. From \( r=5 \) and beyond, the interquartile range is also prominent along with the presence of \textit{out-of-range} outliers.

These results strongly support the earlier qualitative analysis of the EMPCA-MI traces for varying neighbourhood radius, which means that choosing the lowest \( r \) in the \textit{Step I} of EMPCA-MI helps to attain the granularity needed for optimal feature extraction. As and when the higher radii are chosen, local maxima were generated due to the INU and noise artefacts as in Figures 4.9(c) and (d), which consequently lead to poor registration performance. As \( r=1 \) has smoother EMPCA-MI traces, clearly it led to better convergence to the real maxima, in comparison to others.

Thus, considering the above qualitative analysis of EMPCA-MI traces, computation complexity/cost and finally registration performance, \( r=1 \) is chosen to be used for the rest of the experimentation work in this thesis.

4.4 Summary

In this chapter, a new hybrid similarity measure framework called EMPCA-MI has been introduced for MIR, with the aim of providing superior performance in the presence of artefacts and noise in both mono and multimodal images. EMPCA-MI utilises the concept of neighbourhood region from the RMI to extract the spatial features via principal component analysis, but unlike RMI, EMPCA-MI provides dimensionality reduction by iteratively determining the dominant principal com-
ponents. The computation of EMPCA-MI in this chapter was based on choosing one dominant principal component for EMPCA and fixed number of bins for MI computation, though these assumptions will be relaxed in subsequent chapters.

A critical analysis and evaluation of various system choices and parameters relating to interpolation and neighbourhood radius for the EMPCA-MI framework has been presented based on their impact on three factors, namely smoothness of the EMPCA-MI traces, computational complexity/cost involved and RE performance. This analysis showed how the choices of these parameters influence the overall registration performance, with bicubic interpolation and unity radius for neighbourhood regions emerging as the best choices for EMPCA-MI framework. While bicubic interpolation is computationally expensive, it still provided a smoother EMPCA-MI trace and lowest RE during preliminary registration. Similarly, $r=1$ for neighbourhood regions for EMPCA-MI gives consistently smoother traces, lower RE and minimal computational cost.

The next chapter will further investigate the quantitative as well as qualitative registration performance of EMPCA-MI framework on different mono and multimodal medical images. Furthermore, it will include a comparative analysis of EMPCA-MI framework with other existing similarity measures from the literature.
Chapter 5

EMPCA-MI-based Registration

Results

5.1 Introduction

This chapter investigates the performance of the proposed EMPCA-MI similarity measure framework for the registration of medical image datasets. As evident from the rigorous analysis in the previous chapter, the bicubic interpolation and neighbourhood region with $r=1$ were chosen as the most appropriate parameter choices for EMPCA-MI, so they are deployed through all the experiments of this chapter. Other than testing the EMPCA-MI framework, the experiments will also draw a comparison with MI, GMI, RMI, GFM and RIR-BS which were identified as comparators in Chapter 3. This comparative study will examine them from the perspective of both quantitative and qualitative registration performance as well as the computational cost involved.
5.2 Experimental Setup

<table>
<thead>
<tr>
<th>EMPCA-MI Framework</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Parameter Setting</strong></td>
</tr>
<tr>
<td>Neighbourhood radius ($r$)</td>
</tr>
<tr>
<td>Principal components ($P$)</td>
</tr>
<tr>
<td>Number of bins ($b$)</td>
</tr>
</tbody>
</table>

Table 5.1: Parameter settings for EMPCA-MI framework.

Different initial misregistration scenarios along with various INU and noise levels were used during the experiments to test the robustness of EMPCA-MI framework as well as other comparators.

5.2 Experimental Setup

To critically assess the performance of the new EMPCA-MI framework, various experimental steps are employed. Firstly, EMPCA-MI framework is integrated within the MIR test bed. Then, the test trials using Shepp-Logan head phantom test image pair are used to validate and verify the implementation, as already elaborated in the research methodology Chapter 3. The generic MIR model discussed in Section 3.3 is employed for experiments along with the parameter settings of the EMPCA-MI framework shown in Table 5.1.

For the EMPCA-MI framework fixed values of $P$ and $b$ are chosen. This assumption for $P$ and $b$ values will be relaxed and investigated in the next chapters. Registration experiments were conducted in two sets namely Set 1 and Set 2 as detailed in Table 5.2.
5.2 Experimental Setup

<table>
<thead>
<tr>
<th>Experiment Set</th>
<th>Set 1</th>
<th>Set 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dataset Name</td>
<td>BrainWeb</td>
<td>DRIVE</td>
</tr>
<tr>
<td>Dataset Modality</td>
<td><strong>Monomodal</strong></td>
<td><strong>Multimodal</strong></td>
</tr>
<tr>
<td>Dataset Images</td>
<td>T1/T1, T2/T2, PD/PD</td>
<td>T1/T2, T2/PD, PD/T1</td>
</tr>
<tr>
<td>Added INU &amp; Noise</td>
<td>5 levels (No INU &amp; noise, $\alpha_{20}$, $\alpha_{40}$, $\beta$ and $\alpha_{40} + \beta$)</td>
<td>N.A (inherently low contrast and homogeneous regions present)</td>
</tr>
<tr>
<td>Initial Misregistration Scenarios</td>
<td>20 (See Section 3.5.1)</td>
<td>4</td>
</tr>
<tr>
<td>Transformation</td>
<td>Rigid</td>
<td>Similarity</td>
</tr>
<tr>
<td>Comparators</td>
<td>MI, GMI, RMI</td>
<td>GFM</td>
</tr>
<tr>
<td>Total Registrations</td>
<td>2400</td>
<td>160</td>
</tr>
</tbody>
</table>

| Table 5.2: Experiments and their relevant parameter details. |

The **Set 1** of registration experiments uses the mono and multimodal pairs of MRI T1, T2 and PD human brain images of the *BrainWeb* dataset. It comprises of all 2400 separate registration experiments, which include 3 monomodal image pairs, 3 multimodal image pairs, 5 levels of INU and Gaussian noise settings, 20 initial misregistration scenarios and 4 similarity measures (including the EMPCA-MI framework). Also, MI, GMI and RMI have emerged as prominent intensity-based similarity measures and were used as the comparators. The **Set 2** of registration experiments is performed on the mono and multimodal pairs of retinal images from the *DRIVE* and *Brno-Multi* datasets respectively. The monomodal retinal Fundus registration uses GFM (Gharabaghi et al., 2013) for comparison. This subset consists of 160 registrations, which include 20 monomodal image pairs, 4 initial misregistration scenarios and 2 methods (GFM and EMPCA-MI framework) (Reel et al., 2013c). Similarly, the multimodal registration of Fundus and SLO images...
uses the recent RIR-BS (Chen et al., 2015, 2011) as a comparator. This subset consists of 60 registrations, which include 10 multimodal pairs, 3 initial misregistration scenarios and 2 methods (RIR-BS and EMPCA-MI framework) (Reel et al., 2013b). The quantitative final RE and the qualitative checkerboard overlay are used to compare all the registration results.

5.3 Experimental Results Discussion

5.3.1 Set 1: BrainWeb Dataset-based Experiments

The first set of experimental results is for the Set 1 of registration experiments which included both mono and multimodal BrainWeb dataset.

Monomodal Results

Figure 5.1 consists of five RE vs. iteration plots for the monomodal subset for five levels of INU and noise levels in Table 5.2. It shows the 18th initial misregistration scenario \((-23.6, -7.72, -37.6^\circ)\) with initial RE=92.12 pixels. Each of these five plots of Figures 5.1(a) – (e) shows four registration experiments using EMPCA-MI alongside the comparators MI, GMI and RMI. They are plotted with a broken y axis, in order to highlight and zoom on the lower RE value as the registration process progresses to higher iteration values. It is clear from Figures 5.1(a) and (b) that all four similarity measures converge to a lower RE for monomodal T1/T1 registration both when no INU and noise are present and when \(\alpha_{20}\) INU is introduced.
5.3 Experimental Results Discussion

Figure 5.1: RE versus iteration plots for MI, GMI, RMI and EMPCA-MI using 18th initial mis-registration scenario for monomodal registration of (a) T1/T1, (b) T1+$\alpha_{20}$/T1, (c) T1+$\alpha_{40}$/T1, (d) T1+$\beta$/T1 and (e) T1+$\alpha_{40}$+$\beta$/T1 pairs.
But as INU is increased to $\alpha_{40}$ as observed in Figure 5.1(c) only EMPCA-MI provides a lower RE and better rate of convergence with less number of iterations in comparison to others.

Furthermore, as $\beta$ (Figure 5.1(d)) and then finally $\alpha_{40} + \beta$ (Figure 5.1(e)) are used, EMPCA-MI consistently provides lower RE throughout. This is due to the fact that the neighbourhood region (with $r=1$) pre-processing used for extraction of first principal component with EMPCA provides accurate information for the computation of MI, even when high levels of INU and Gaussian noise are introduced. In addition, using bicubic interpolation during the registration helped EMPCA-MI to achieve smoother EMPCA-MI traces, which leads to its better convergence. On the other hand, GMI used gradient information for its computation. Hence, in cases with higher INU and Gaussian noise levels, more noise is introduced into the gradient information, which leads to its inferior convergence (Figures 5.1(d) and (e)). For example, in Figure 5.1(d) GMI failed to register the monomodal pair and gave final RE even higher than the initial misregistration value. While EMPCA-MI is able to accurately quantify the spatial alignment between the image pairs with lower RE, RMI using a rough approximation of MI shows a degrading performance with increasing levels of INU and noise conditions (Figures 5.1(c) – (e)).

To have a qualitative assessment of the registration process outcome, Figures 5.2(a) – (f) show the *Canny edge detection* overlay results corresponding to the monomodal T1+$\alpha_{40}$+$\beta$/T1 pair registration shown in Figure 5.1(e). Figure 5.2(a) shows the initial misregistration for 18th initial misregistration scenario and the ground truth in Figure 5.2(b).
Figure 5.2: Qualitative results of monomodal T1+\(\alpha\)+\(\beta\)/T1 pair using 18\textsuperscript{th} initial misregistration scenario showing (a) initial misregistration, (b) ground truth and (c) MI, (d) GMI, (e) RMI, (f) EMPCA-MI output with inset showing zoomed-in central lobes.
Figures 5.2(c) – (f) show the registration output for MI, GMI, RMI and EMPCA-MI respectively. The yellow colour in these images represents the correct alignment of the region, while red and green colour shows the misalignment. The quality of registration is apparent, particularly from the alignment of the two central lobes in the insets of Figures 5.2(c) – (f) of the registration output. Figure 5.2(f) clearly shows that EMPCA-MI provides best alignment in comparison to MI, GMI and RMI output and is nearest to the ground truth shown in Figure 5.2(b). These results are consistent with the earlier qualitative results and highlight the superior registration performance of EMPCA-MI in comparison to others.

Similar results were also observed for monomodal T2/T2 and PD/PD pair registrations for the same initial misregistration scenario. For completeness, Figures D.1(a) and (b) in Appendix D, shows the RE versus iterations plots for MI, GMI, RMI and EMPCA-MI for T2+$\alpha_{40}+\beta$/$T2$ and PD+$\alpha_{40}+\beta$/$PD$ pairs respectively. They were predominantly chosen for display since they show the results for the images with highest level of INU and noise present. Similarly, Figures D.2 and D.3 show the qualitative results of the above two mentioned pairs with their corresponding ground truths.

**Multimodal Results**

Figure 5.3 shows the RE vs. iteration plots for the multimodal $T1+$+$\alpha_{40}+\beta$/$T2$ pair using $15^{th}$ initial misregistration scenario $(−5.23, 2.31, 34.8°)$ with initial RE=84.16 pixels. It is again evident that EMPCA-MI provides a better rate of convergence and leading to the lowest RE, followed by GMI and MI.
5.3 Experimental Results Discussion

![Figure 5.3: RE versus iterations plots for MI, GMI, RMI and EMPCA-MI using 15th initial misregistration scenario for multimodal registration of (a) T1+α40+β/T2 pair.](image)

Even at higher level of INU and Gaussian noise, EMPCA-MI converges to a lower RE from 4th iteration onwards due to its ability to accurately represent the neighbourhood region features and quantify the spatial alignment between them. On the contrary, MI and GMI provide much higher RE, while RMI exhibits poor rate of convergence which locks into a local maximum too far from the global maximum.

To have a qualitative assessment of this particular multimodal T1+α40+β/T2 pair, Figures 5.4(a) – (f) shows the Canny edge detection overlaid corresponding results. Figure 5.4(a) shows the initial misregistration for 15th initial misregistration scenario and the corresponding ground truth in Figure 5.4(b). While Figure 5.4(c) – (f) show the registration output for MI, GMI, RMI and EMPCA-MI respectively. It is apparent, particularly from the alignment of the two central lobes in the insets of Figures 5.4(c) – (f) showing the registration output. Figure 5.4(f) clearly shows that EMPCA-MI provides best alignment followed by GMI output and is the closest to the ground truth in Figure 5.4(b).
Figure 5.4: Qualitative results of multimodal T1+α₄₀⁺β/T2 pair using 15ᵗʰ initial misregistration scenario showing (a) initial misregistration, (b) ground truth and (c) MI, (d) GMI, (e) RMI, (f) EMPCA-MI output with inset showing zoomed-in central lobes.
In addition, similar trends are seen in the results for the registration of the other four INU and Gaussian noise levels of this multimodal pair, included in Figures D.4(a) – (d) of Appendix D for completeness. These plots show the subsequent RE vs. iteration plots for the T1/T2, T1+α20/T2, T1+α40/T2 and T1+β/T2 pairs. It is clear from them that MI, GMI and RMI perform well when no or low INU is present, but as high INU and Gaussian noise are introduced, their performance worsens significantly with a lower rate of convergence. This can be further verified from the multimodal T2/PD and PD/T1 pair registrations for the same initial misregistration scenario. Figures D.4(e) and (f) in Appendix D, show the RE versus iterations plots for T2+α40+β/PD and PD+α40+β/T1 pairs. They were chosen because they represent the results for the most challenging pairs with highest level of INU and Gaussian noise. Similarly, Figures D.5 and D.6 of Appendix D show the qualitative results of both these pairs with their corresponding ground truths.

Results Summary

Finally, the overall RE performance for the Set 1 experiments are summarised. Figures 5.5(a) and (b) show the detailed summary boxplot for all INU and noise levels, over the 20 initial misregistration scenarios for monomodal T1/T1 and multimodal T1/T2 registration experiments respectively. Each boxplot shows a bounding box defining the interquartile range with the bar across representing the median and whiskers denoting the RE range. The boxplots on top also include the number of out-of-range RE values corresponding to each similarity measure used.
Both boxplots in Figure 5.5 confirm the improved registration performance of EMPCA-MI with the lowest mean RE and few number of out-of-range outliers in comparison to the other three comparative similarity measures. For monomodal T1/T1 pair in Figure 5.5(a), EMPCA-MI consistently achieve better RE than all other similarity measure, especially in the latter two cases (β and $\alpha_{40} + \beta$).
For example in case of the monomodal T1+$\alpha_{40}+\beta$/T1 pair, EMPCA-MI maintains a lower RE and interquartile range with single *out-of-range* outlier, on the contrary the *mean* RE for RMI overshoots due to the large RE values for two *out-of-range* outliers. Similarly for multimodal T1/T2 pair in Figure 5.5(b), EMPCA-MI has lowest *mean* RE for all five INU and noise settings, while GMI performance degrades in the last two challenging cases ($\alpha_{40}$ and $\alpha_{40}+\beta$). Furthermore, the *mean* RE and *interquartile* range for RMI is also significantly large as higher INU and noise levels are introduced.

These results strongly support the individual RE vs. iteration plots for mono and multimodal registration experiments. Similar observations can be deduced from the boxplots for the mono (T2/T2 and PD/PD) and multimodal (T2/PD and PD/T1) registration experiments, which are included as Figures D.7 and D.8 of Appendix D for completeness. In summary, considering all these boxplots for all the set of registrations for BrainWeb dataset, EMPCA-MI is able to reduce the RE on an average by 45%, 42% and 61% for MI, GMI and RMI respectively. These results clearly help to draw the conclusion that EMPCA-MI using $r=1$ for neighbourhood region pre-processing with $P=1$, to select the most prominent feature to calculate the MI, while other comparators perform well when no or low INU is present, at higher levels of INU and Gaussian noise their performance degrades.

On closer examination however, it is also evident from these results that there are few cases of EMPCA-MI-based registrations, which gave higher RE. For example, the outlier (with *final* RE value of 44.80 pixels) shown in the EMPCA-MI boxplot for monomodal T1+$\alpha_{40}+\beta$/T1 in Figure 5.5(a). A few other results with
higher RE for EMPCA-MI can also be seen in Figure D.7 and D.8 of Appendix D. Interestingly, all these cases are mostly observed where high INU and noise levels are present. The reason of EMPCA-MI underperformance particularly in these cases will be discussed later in Section 5.3.3 of this chapter.

To evaluate the computational cost involved in the iterative registration process, an ART analysis was undertaken of all the similarity measures. Figure 5.6 shows the corresponding ART performance of all the four similarity measures considered for the mono and multimodal BrainWeb dataset registration experiments of Set 1. Here, EMPCA-MI requires around 15% less computation time in comparison to RMI and despite of being slower than other two comparators (MI and GMI) it is able to afford more robust and accurate registration performance both quantitatively and qualitatively. This is due to the fact that it utilises the iterative EMPCA mechanism to compute the first principal component from the neighbour region features, which helps in overcoming the artefacts due to the high INU and noise levels. While, MI takes the minimum computation time among all the sim-
ilarity measures since it bins the intensity values of the image into histograms to compute the individual and joint probability distribution, it was observed to have degraded registration performance. Also in case of GMI, the computation time increases by $25\%$ with respect to MI, which reflects the added overhead of gradient information computation needed in addition to the MI calculation. Although RMI deploys a covariance matrix-based approximation to compute the individual and joint entropies, it consequently leads to doubling of the computation time with respect to GMI.

Overall, it is comprehensible that while EMPCA-MI has higher ART in contrast with other two similarity measure (MI and GMI), it is still able to provide better registration performance in terms of quantitative RE and qualitative assessment.

5.3.2 Set 2: DRIVE and Brno-Multi Dataset Experiments

This section discusses the experimental results for EMPCA-MI on Set 2 of monomodal DRIVE dataset followed by those of the multimodal Brno-Multi dataset.

Monomodal Results

For the monomodal registrations, GFM was chosen as a comparator since it has emerged as one of the recent methods developed specifically for monomodal Fundus registration and has been used upon the same DRIVE dataset. In these experiments four initial misregistration scenarios were used. The first three initial misregistration scenarios were chosen as in (Gharabaghi et al., 2013) in order to enable a direct comparison.
Moreover, in order to further investigate the performance of EMPCA-MI, an additional 4th initial misregistration scenario was introduced. In the 1st and 2nd initial misregistration scenarios, only one registration parameter varies i.e. rotation $\theta=5^\circ$ and $\theta=60^\circ$ with the other parameters being fixed. While in 3rd and 4th initial misregistration scenario all four registration parameters are varied. This includes $S$, which simulates that Fundus image have been acquired using different equipment magnifications.

Figure 5.7 consists of RE summary bar graph of 20 monomodal Fundus image pairs using EMPCA-MI and GFM for the 3rd initial misregistration scenario ($5, 5, 20^\circ, 2.8$). This bar graph are plotted with twice broken y axis in order to accommodate all the RE values on the same plot as well as to able to show the lower RE values. For example, the RE for Image Pair 5 using EMPCA –MI is almost 5 times lower than GFM. Image Pair 5 highlights the inherent characteristics present in all retinal images which lead to challenges in their registration. The EMPCA-MI framework is able to register with lower RE since by using a neighbourhood region-
based extraction of the first principal component helps in accurately compute MI despite of homogeneous regions and low contrast present, whereas GFM being a point-based method provided a false match for the registration and had significant RE for this Image Pair.

To have a qualitative assessment of the registration outcome in this subset, Figures 5.8(a) – (e) show the checkerboard overlay results of Image Pair 5 using 3rd initial misregistration scenario. Figure 5.8(a) shows the initial misregistration for 3rd initial misregistration scenario and the ground truth with three marked regions in Figure 5.8(b). Also, Figures 5.8(c) and (d) show the registration output for GFM and EMPCA-MI with corresponding marked regions respectively. In Figure 5.8(e) the zoomed three marked regions of the ground truth and two registration outputs are shown. It is clear from Figure 5.8(e) that EMPCA-MI provides best vascular structure continuity in comparison to GFM and is also nearer to the ground truth in terms of alignment. Overall, these results confirm that EMPCA-MI afforded better registration performance compared to GFM in variable contrast and illumination conditions.

It is especially notable to highlight that EMPCA-MI successfully registered Image Pair 9 as shown in Figures 5.9(a) – (e), whereas GFM failed to converge (Gharabaghi et al., 2013). This is due to the fact that EMPCA-MI does not rely on point-based feature extraction, but uses intensity-based feature extraction and hence it was more robust in being able to successfully registering Image Pair 9. In contrast, GFM extracts the closed boundary regions from Image Pair 9 but since they have similar shapes and their corresponding affine moment invariant descriptors were very close to each other, it led to its failure in registering them.
Figure 5.8: Qualitative results of monomodal Fundus retinal Image Pair 5 using 3\textsuperscript{rd} initial misregistration scenario, showing (a) initial misregistration, (b) ground truth, (c) GFM and (d) EMPCA-MI output. Also (e) shows the zoomed-in regions of ground truth (1\textsuperscript{st} row), GFM (2\textsuperscript{nd} row) and EMPCA-MI output (3\textsuperscript{rd} row) respectively.
5.3 Experimental Results Discussion

Figure 5.9: Qualitative results of monomodal Fundus retinal Image Pair 9 using 3\textsuperscript{rd} initial misregistration scenario, showing (a) initial misregistration, (b) ground truth, (c) EMPCA-MI output, with (d) zoomed-in ground truth and (e) EMPCA-MI regions. N.B. GFM failed to register this image pair.

The zoomed in qualitative results for Image Pair 9 using EMPCA-MI along with the corresponding ground truth are shown in Figures 5.9(e) and (d) respectively.

Finally, Figure 5.10 summarises the first subset of Set 2 registration experiments. It shows the RE summary boxplot of all monomodal Fundus retinal image pairs using GFM and EMPCA-MI for the 4 test initial misregistration scenarios. The boxplots reveal the overall better registration performance of EMPCA-MI.
5.3 Experimental Results Discussion

![Boxplot of RE summary](image)

Figure 5.10: RE summary boxplot of 20 monomodal Fundus retinal image pairs using GFM and EMPCA-MI for all 4 initial misregistration scenarios. The mean and outliers are denoted by ♦ and + shapes. The numbers on top of the plot refer to the number of out-of-range outliers with RE higher than 18 pixels. N.B. GFM results for scenario no. 4 were unavailable.

over GFM. In all the initial misregistration scenarios EMPCA-MI achieves a lower mean RE as well the interquartile range, except for 2nd initial misregistration scenario where its mean RE is marginally higher than GFM, despite of having a lower interquartile range. This was due to the one out-of-range outlier for which the final RE was much higher and was challenging for both EMPCA-MI and GFM to register. Such particular cases will be discussed in detail in Section 5.3.3.

Multimodal Results

Now the registration experiment results for the second subset of Set 2 using multimodal Fundus/SLO Brno-Multi dataset are discussed. To draw a comparison, RIR-BS was chosen as a comparator for these multimodal registrations, as it was recently developed specifically for multimodal retinal registration. As discussed earlier in Section 2.3.3, it employs a bifurcation structure matching which comprises of a master bifurcation point and three connected neighbours (Chen et al., 2015, 2011).
5.3 Experimental Results Discussion

In these registration experiments, three initial misregistration scenarios were used. In the 1\textsuperscript{st} initial misregistration scenario, only the rotation parameter was varied to $\theta=60^\circ$ with the other parameters being fixed. The 2\textsuperscript{nd} and 3\textsuperscript{rd} misregistration represents more challenging scenarios with all the four registration parameters being varied. The value of $S$ as in previous subset reflects the multimodal pair has been acquired with different equipment magnifications while the translations are pragmatically kept low as these motions are generally minimal during the image acquisition process.

In Figure 5.11, the RE summary bar graph of 10 multimodal Fundus/SLO image pairs using EMPCA-MI and RIR-BS for the 2\textsuperscript{nd} initial misregistration scenario $(10, 10, 30^\circ, 2)$ are shown. For example, for Image Pair 3 and 8, EMPCA–MI provides 25% and 50% less RE respectively, in comparison to RIR-BS. EMPCA-MI is in fact able to attain a lower RE since it utilises the local neighbourhood region using $r=1$ to extract the most prominent feature from the largely homogeneous and variable contrasts multimodal image pairs. Other than evaluating the quantitative RE performance, in order to appreciate the superiority of the
EMPCA-MI results, a qualitative assessment of the registration outcome in also conducted.

Figures 5.12(a) – (e) show the checkerboard overlay result of Image Pair 3 using 2\textsuperscript{nd} initial misregistration scenario. Figures 5.12(a) and (b) show the initial misregistration for 2\textsuperscript{nd} initial misregistration scenario and the ground truth with three marked regions respectively. Whereas, Figures 5.12(c) and (d) show the registration output for RIR-BS and EMPCA-MI with corresponding marked regions respectively. In Figure 5.12(e) the zoomed three marked regions of the ground truth and two registration outputs are shown. It is evidently clear that EMPCA-MI provides much better structural continuity in comparison to RIR-BS output. Overall, it is apparent that EMPCA-MI provided lower RE for almost all image pairs, except Image Pair 4 and 6.

Interestingly, notable exceptions for EMPCA-MI are Image Pair 4 and 6 where RIR-BS has a marginally lower RE. Similar higher RE for EMPCA-MI was again observed for the 3\textsuperscript{rd} initial misregistration scenario for Image Pair 4. This is due to the fact that two image pairs, especially Image Pair 4 is one of the most challenging pairs in the Brno-Multi dataset as it exhibits pathology alongside low contrast and large homogeneous regions, which compromises EMPCA-MI performance when only one principal component is used. In contrast, since RIR-BS method uses segmentation, its registration quality is influenced much more by low contrast in the images rather than the presence of large homogeneous regions in this pair. The higher RE of EMPCA-MI highlights a limitation of the EMPCA-MI in the case of such challenging Image Pair, which will be analysed in the next section.
5.3 Experimental Results Discussion

Figure 5.12: Qualitative results of multimodal Fundus/SLO retinal Image Pair 3 using \(2^{nd}\) initial misregistration scenario, showing (a) initial misregistration, (b) ground truth, (c) RIR-BS and (d) EMPCA-MI output. Also (e) shows the zoomed-in regions of ground truth (1\(^{st}\) row), RIR-BS (2\(^{nd}\) row) and EMPCA-MI output (3\(^{rd}\) row) respectively.
Finally, Figure 5.13 summarises the second subset of Set 2 registration experiments. It shows the RE summary boxplot of all multimodal Fundus/SLO retinal image pairs using RIR-BS and EMPCA-MI for the three initial misregistration scenarios. The boxplots reveal the overall superior registration performance of EMPCA-MI in terms of both mean RE and interquartile range over RIR-BS in all scenarios.

The boxplot for the 3rd initial misregistration scenario highlights the better RE performance achieved by EMPCA-MI with lower mean RE, smaller interquartile range and no out-of-range outlier RE value. Here, the neighbourhood region-based pre-processing followed by the feature extraction for MI computation enables the EMPCA-MI to perform better multimodal registration. On the contrary, despite of using a point-based approach, RIR-BS does not avail from any vessel width-based refinement (Chen et al., 2011) for its matching, which leads to an overall higher mean RE in these multimodal experiments. Hence, these results confirm...
that overall EMPCA-MI provided better registration performance compared to RIR-BS especially in the presence of low contrast and large homogeneous non-vascular regions.

Moreover, the summary boxplots in Figures 5.10 and 5.13 help to conclude that EMPCA-MI is able to overall reduce the RE on an average by 30% for mono and multimodal retinal image registrations with respect to GFM and RIR-BS.

5.3.3 EMPCA-MI Framework Robustness

The previous sections have investigated the performance of EMPCA-MI framework along with other comparators for two sets of mono and multimodal registration experiments. Although overall EMPCA-MI achieved lower RE in comparison to others, as shown in Figures 5.5, 5.10 and 5.13, there were cases when EMPCA-MI performance is not as accurate as the comparators. To critically evaluate the reasons for this, two worst case scenarios are considered.

In Section 5.3.1, there were few individual cases in the mono and multimodal registration of Set 1, having high levels of INU and noise (for example $\alpha_{40}$ and $\alpha_{40}+\beta$) where EMPCA-MI gave higher RE values. For example in Figure 5.5(a) the EMPCA-MI boxplot for monomodal T1+$\alpha_{40}+\beta$/T1 shows one such outlier (with RE value of 44.80 pixels) for 11$^{th}$ initial misregistration scenario ($24.21, -25.2, -25.8^\circ$). Few similar cases can also be observed for multimodal registrations of Set 1. The EMPCA-MI boxplot for multimodal T2+$\alpha_{40}+\beta$/PD in Figure D.8(a) of Appendix D shows two such outliers (with RE values of 29.48 and 49.35 pixels) representing result of 7$^{th}$ and 20$^{th}$ initial misregistration scenarios.
5.3 Experimental Results Discussion

Figure 5.14: Qualitative results of multimodal $T1+\alpha_{40}+\beta/T1$ pair using 11$^{th}$ initial misregistration scenario showing (a) initial misregistration with (b) ground truth and (c) EMPCA-MI output with inset showing zoomed-in central lobes.

To analyse the impact on the quality of registration, Figures 5.14(a) – (c) show the qualitative result for EMPCA-MI of monomodal $T1+\alpha_{40}+\beta/T1$ registration with 11$^{th}$ initial misregistration scenario. The initial misregistration for this monomodal pair is shown in Figure 5.14(a). It is clear from the misaligned central lobes shown in the zoomed-in inset in Figure 5.14(c) that the EMPCA-MI registration performance is inferior in comparison to the ground truth in Figure 5.14(b). It can be argued that the deterioration in the EMPCA-MI performance specifically in the presence of high levels of INU and noise is due to the design of Step I of the EMPCA-MI framework in Figure 4.2 (which shows an example of the EMPCA-MI framework).
However, in closely observing Figure 4.2, there is a noticeable repetitive pattern in $Q_R$ and $Q_S$ which correspond to the rearranged neighbourhood region information for the reference and sensed image respectively. This $1^{st}$ order representation of the neighbourhood region does not include any relative information about neighbourhood pixels, so at higher levels of INU and noise, the $Step I$ of the EMPCA-MI framework may not always be able to represent the neighbourhood information accurately. This could indeed lead to inaccurate EMPCA-based extraction of features, which are used for the MI computation.

Moreover, as mentioned in Legg et al. (2013), the number of bins for MI computation has been shown to be a critical parameter for accurate MI computation. Therefore, the empirically fixed number of bins in the $Step III$ of the EMPCA-MI framework may lead to inaccurate MI calculation especially when the image intensity values are corrupted by noise due to high levels of INU. Hence, the role of the neighbourhood region rearrangement (in $Step I$) and number of bins for MI (in $Step III$) in EMPCA-MI framework will be separately investigated in the next chapter, with an aim to enhance the EMPCA-MI framework as described in Figure 1.1.

Furthermore in $Section 5.3.2$, which discussed the $Set 2$ of registration experiments, EMPCA-MI provided high RE for the $Image Pair 4$, which is one of the most challenging pairs in the multimodal $Brno-Multi$ dataset. To understand the reasons behind this degraded performance for this particular case, further experimentation was conducted by choosing two principal components i.e. $P=2$ manually in the $Step II$ of EMPCA-MI framework, while keeping all other parameters fixed for both the $2^{nd}$ and $3^{rd}$ initial misregistration scenarios.
Figure 5.15: Qualitative results of multimodal Fundus/SLO retinal Image Pair 4 using 3\textsuperscript{rd} initial misregistration scenario, showing (a) EMPCA-MI output (with $P=2$) and (b) RIR-BS output. Also (c) shows the three zoomed-in regions of EMPCA-MI output (with $P=2$) (1\textsuperscript{st} row) and RIR-BS (2\textsuperscript{nd} row) respectively.

Figures 5.15(a) – (c) show the qualitative results for this particular multimodal experiment for the 3\textsuperscript{rd} initial misregistration scenario. It is clear, that using EMPCA-MI with $P=2$ provided improved registration results in comparison to RIR-BS, which is attributed to the empirically chosen $P=2$, which represents the overall variance of the image pair, but now with two orthogonal principal components. This also highlights the versatility of the EMPCA-MI framework to iteratively determine more principal components and this feature can be exploited.
in such cases. This provided the motivation to investigate this feature of EMPCA-MI framework in more detail. Hence, Chapter 7 of this thesis will investigate a mechanism, which has the ability, in Step II of EMPCA-MI framework to automatically select the number of principal components adaptively instead of manually as shown in Figure 1.1.

5.4 Summary

This chapter investigated the registration performance of EMPCA-MI framework using human brain *BrainWeb* dataset followed by monomodal retinal *DRIVE* dataset and multimodal *Brno-Multi* dataset. It also conducted a comparative performance study using recent suitable comparators from the literature to enable a critical and thorough analysis of registration experiments using RE vs. iteration plot, quantitative and qualitative RE and ART as the performance metrics. This chapter showed how the EMPCA-MI framework is able to outperform its comparators by reducing the RE and providing a robust registration performance for the various mono and multimodal datasets, even when high levels of INU, low contrast and large homogeneous regions, which are known to make the registration process challenging. Finally, this chapter also gave an insight into the shortfalls of EMPCA-MI framework, while discussing two such cases in detail, where the registration performance degraded. The next chapter will propose and investigate enhancements based on the rearrangement of neighbourhood region and number of bins for MI computation in the EMPCA-MI framework, with an aim to further increase the robustness and achieve lower RE.
Chapter 6

EMPCA-MI Framework

Enhancements

6.1 Introduction

In the previous chapter, the experimental results of proposed EMPCA-MI framework provided consistent and robust registration performance for both mono and multimodal medical datasets. However, in a few registration experiments, the performance of EMPCA-MI framework was observed to be compromised, especially where high INU noise and large homogeneous regions were present. This was associated to the 1st order-based rearrangement of the neighbourhood region and choosing fixed number of bins for the MI computation for the EMPCA-MI framework. Hence, this chapter proposes two novel modifications to the EMPCA-MI framework, primarily in its pre-processing and MI computation steps. Firstly, it utilises the 2nd order representation of the image intensity values for direct
(4-pixel) and indirect (8-pixel) neighbourhood connectivity, instead of just re-arranging the pixels in the pre-processing stage. Secondly, it explores the effect of introducing either fixed or adaptive number of bins selection strategy for the MI computation stage. Both these enhancements have been rigorously investigated using separate experiments to provide the numerical and qualitative registration results which confirm the improved robustness and show their corresponding computational cost. Figure 6.1(a) shows the original EMPCA-MI framework, which utilised the 1st order indirect (8-pixel) neighbourhood region connectivity for the pre-processing stage and fixed number of bins for MI computation. Section 6.2 describes the introduction of 2nd order direct (4-pixel) and indirect (8-pixel) neighbourhood connectivity for the pre-processing of the reference and sensed images.
Section 6.3 explores the use of different fixed and adaptive number of bins selection approaches for the MI computation step. Figures 6.1(b) and (c) show the various approaches explored as Enhancement 1 and 2, which will be discussed in Sections 6.2 and 6.3 respectively.

6.2 EMPCA-MI Framework Enhancement 1

As detailed in Chapter 4, the pre-processing step of EMPCA-MI framework rearranges the neighbourhood region grayscale data values into a vector form to preserve both the spatial and intensity information of the images. Figure 6.2(a) shows the illustration of EMPCA-MI framework with Step I in detail, which reorganises the image grayscale values using 1st order-based (8-pixel) neighbourhood region to incorporate the spatial information. This provided noteworthy performance in the registration experiments in Chapter 5, particularly when there is neither INU nor noise present. However, when there was high levels of INU and noise, the corresponding registration performance dropped.

This limitation of EMPCA-MI framework was due to the 1st order-based neighbourhood region representation, which considers each pixel independently and without cognisance of any neighbourhood relations. This was reflected in the repetitive patterns in the columns of $Q_R$ and $Q_S$ matrix for the neighbouring position of the sliding window, which can be seen in Figure 6.2(b). This provides the motivation for modifying this step of EMPCA-MI framework such that the neighbouring relations within a given neighbourhood region are preserved for other steps in the EMPCA-MI framework.
**6.2 EMPCA-MI Framework Enhancement 1**

**EMPCA-MI framework**

*1st order 8-pixel connectivity*

**Proposed modified EMPCA-MI framework**

*2nd order 8-pixel connectivity*

*2nd order 4-pixel connectivity*

**Figure 6.2:** Illustration of the original EMPCA-MI framework with 1st order (8-pixel) and the proposed 2nd order (8-pixel and 4-pixel) region connectivity-based pre-processing in Step I, for an image pair with size of 10 × 10 pixels.
Hence, in order to achieve more robust registration between the reference and sensed images, this rationale during the pre-processing step a neighbourhood region can be more precisely characterised by a 2\textsuperscript{nd} order representation, where the relationship between pixels can be exploited instead of just pixel values.

### 6.2.1 Neighbourhood Region Selection

This section illustrates the modified pre-processing step for EMPCA-MI similarity measure framework (Reel et al., 2013a). Considering a sliding window \( B \) with radius \( r \) having a \((2r + 1) \times (2r + 1)\) pixels neighbourhood region. Then the resulting single row vector \( B^\ast \) will have length \( c+1 \), and can be represented as:

\[
B^\ast_i = \begin{cases} 
B_i - B_{\frac{c}{2}+1} & i \in [1, c+1]; i \neq \frac{c}{2} + 1 \\
B_{\frac{c}{2}+1} & i = \frac{c}{2} + 1
\end{cases} \tag{6.1}
\]

Each row vector \( B^\ast \) now represents the differential value of \( c \) connected pixels with respect to the centre pixel \( B_{\frac{c}{2}+1} \). Figure 6.2(a) gives an example using an image pair with size of 10 \( \times \) 10 pixels showing the proposed 2\textsuperscript{nd} order 4-pixel (direct neighbours) and 8-pixel (indirect neighbours) connectivity alongside the original EMPCA-MI framework. Here, \( B \) is having a 3 \( \times \) 3 pixel neighbourhood region when \( r=1 \) for both \( c=4 \) and 8 (shown with yellow colour in Figure 6.2(a)) and the corresponding row vector \( B^\ast \) with centre pixel \( B_3 \) and \( B_5 \) (shown with orange colour in Figure 6.2(b)) respectively. Here, the modified EMPCA-MI pre-processing no longer generates repetitive pattern as in original EMPCA-MI, but instead provides unique relative intensity values (See \( B^\ast \) in Figure 6.2) for its next
computational steps. Hence this enhancement helps in representing the spatial information for effectively and thus now leading to more accurate EMPCA-MI value in comparison to the original EMPCA-MI.

Although both 2nd order (8-pixel and 4-pixel) neighbourhood connectivity will provide better spatial information in comparison to the 1st order (8-pixel) neighbourhood, amongst the latter both, the 2nd order (8-pixel) neighbourhood is known to have degraded performance (Demirkaya et al., 2008). Since, 8-pixel connectivity assumes all the 8 pixels to be connected to the centre pixel, it gives an equal weighting to them but in fact the diagonal pixels have a weaker linkage to the centre pixel in comparison to the horizontal and vertical neighbouring pixels. Also as observed from Figure 6.2(b), the dimensionality $d$ for 2nd order (8-pixel and 4-pixel) neighbourhood connectivity will be 9 and 5, respectively.

### 6.2.2 Experiment Setup

To assess the performance of the EMPCA-MI framework Enhancement 1, registration experiments on mono and multimodal BrainWeb dataset are performed while keeping the remaining two processing steps (Step II and Step III) of EMPCA-MI framework the same in Chapter 4. All the parameter settings of the generic registration model and the EMPCA-MI framework for the experiments are same as mentioned in Tables 3.1 and 5.1, respectively, except the neighbourhood region enhancement in the EMPCA-MI framework.

Furthermore, the mono and multimodal BrainWeb dataset was chosen with all experimental settings same as in the previous chapter, as shown in Table 5.2.
In all 1800 separate registration experiments are performed in this set, which include, 3 monomodal image pairs, 3 multimodal image pairs, 5 levels of INU and Gaussian noise settings, 20 initial misregistration scenarios and 3 similarity measures (including the 2nd order (8-pixel and 4-pixel) modifications along with the original EMPCA-MI framework).

### 6.2.3 Experimental Results Discussion

The first set of experiments included both mono and multimodal BrainWeb dataset in order to evaluate the impact of modifying the Step I of EMPCA-MI framework on the registration performance, especially in the presence of high INU and noise levels. The three neighbourhood region connectivity considered for these experiments were the 2nd order (8-pixel and 4-pixel) based enhancements as well as the original EMPCA-MI similarity measure with 1st order (8-pixel) connectivity.

Figures 6.3(a) and (b) show the detailed summary boxplot for all INU and noise levels, over the 20 initial misregistration scenarios for monomodal T1/T1 and multimodal T1/T2 registration experiments respectively. Both boxplots in Figure 6.3 confirm the improved registration performance of EMPCA-MI framework using 2nd order (4-pixel) neighbourhood region connectivity with the lowest mean RE and fewer out-of-range outliers in comparison to the other two neighbourhood region connectivity schemes.

For example, in Figure 6.3(a) which shows the monomodal T1/T1 pair registrations, 2nd order (4-pixel) neighbourhood region connectivity consistently have lower mean RE than the original EMPCA-MI (1st order (8-pixel)) as well as 2nd
Figure 6.3: RE summary boxplot of (a) monomodal T1/T1 pair and (b) multimodal T1/T2 pair using 1st order (8-pixel), 2nd order (8-pixel) and 2nd order (4-pixel) neighbourhood region connectivity in EMPCA-MI framework for all 20 initial misregistration scenarios. The mean and outliers are denoted by ♦ and + shapes. The numbers on top of the plots refer to the number of out-of-range outliers with RE higher than 14 and 20 pixels respectively.

order (8-pixel) due to the more accurate representation of relative spatial information.

On the other hand, the performance of 1st order (8-pixel) and 2nd order (8-pixel) neighbourhood connectivity worsens in the later three cases (α40, β and
\(\alpha_{40} + \beta\) due to the presence of high INU and noise. The 2\textsuperscript{nd} order (4-pixel) neighbourhood region connectivity still provides consistent lower mean RE performance, with an exception of one out-of-range outlier in case of monomodal T1+\(\alpha_{40}/T1\) pair. Both 2\textsuperscript{nd} order (8-pixel) and original EMPCA-MI (1\textsuperscript{st} order (8-pixel)) have more out-of-range outliers as well as larger interquartile range as the level of INU and noise increases.

These results show that using 2\textsuperscript{nd} order (4-pixel) neighbourhood region connectivity in the Step I of EMPCA-MI framework reflects neighbourhood spatial information more accurately by considering a 2\textsuperscript{nd} order representation of the directly related pixel values with respect to the centre pixel of the neighbourhood region.

In the case of multimodal T1/T2 pair, Figure 6.3(b) shows that both 2\textsuperscript{nd} order (8-pixel and 4-pixel) neighbourhood region connectivity perform better than original EMPCA-MI (1\textsuperscript{st} order (8-pixel)) for all INU and noise levels. Since, the 2\textsuperscript{nd} order (4-pixel) neighbourhood region connectivity exploits the strong pixel relations providing more relevant spatial information about a region’s local neighbourhood for subsequent EMPCA and MI computation. Whereas, the 2\textsuperscript{nd} order (8-pixel) neighbourhood region connectivity has higher number of out-of-range outliers, which emphasises its poor rate of convergence due to the weaker indirect pixel neighbours, especially when higher INU and noise levels are present. Similar trends can also be observed from the summary boxplots of the remaining mono (T2/T2 and PD/PD) and multimodal (T2/PD and PD/T1) registration experiments, which are included as Figures E.1 and E.2 of Appendix E for completeness.
6.2 EMPCA-MI Framework *Enhancement 1*

In summary, considering all these boxplots for all the set of registrations for *BrainWeb* dataset, EMPCA-MI framework using 2nd order (4-pixel) is able to reduce the average RE by 79% and 71% with respect to original EMPCA-MI (1st order (8-pixel)) and 2nd order (8-pixel), respectively. Hence, it helps to draw the conclusion that 2nd order (4-pixel) provides more accurate features of the reference and the sensed images in comparison to both the original EMPCA-MI (1st order (8-pixel)) and 2nd order (8-pixel), especially when higher INU and noise levels were present.

In order to evaluate the computational cost of *Enhancement 1*, an ART analysis for the registration experiments is considered. Figure 6.4 shows the ART performance of all the neighbourhood region connectivity considered for the mono and multimodal *BrainWeb* dataset registration experiments. As illustrated in Figure 6.2, since the data dimensionality of EMPCA-MI framework using 2nd order (4-pixel) neighbourhood region connectivity is reduced to 5 from the fixed value of 9 for both original EMPCA-MI framework using 1st order (8-pixel) and 2nd
order (8-pixel), the corresponding ART value is considerably lower, i.e., 95 ms compared to 155 ms for 2\textsuperscript{nd} order (8-pixel) and 160 ms for 1\textsuperscript{st} order (8-pixel) to determine only the first principal component. Hence, evidently EMPCA-MI framework using 2\textsuperscript{nd} order (4-pixel) neighbourhood connectivity is able to provide 40\% improvement in computational efficiency along with the better registration performance in comparison to others.

In Section 5.3.1, there were few individual cases in both mono and multimodal registration of BrainWeb dataset, having high levels of INU and noise (for example $\alpha_{40}$ and $\alpha_{40}+\beta$) where the standard EMPCA-MI framework (using 1\textsuperscript{st} order (8-pixel) neighbourhood region connectivity) provided higher RE. One particular example of deteriorated registration performance was monomodal T1+$\alpha_{40}+\beta$/T1 registration (with RE value of 44.80 pixels) with 11\textsuperscript{th} initial misregistration scenario ($24.21, -25.2, -25.8^\circ$), shown in Figure 5.14.

The EMPCA-MI framework using 2\textsuperscript{nd} order (4-pixel) gave lowest RE (RE value of 0.39 pixels) in comparison to both the 2\textsuperscript{nd} order (8-pixel) (RE value of 10.30 pixels) and original EMPCA-MI framework (1\textsuperscript{st} order (8-pixel)) (RE value of 44.80 pixels) respectively. In order to observe the corresponding quality of registrations, Figure 6.5 shows the qualitative results.

Figures 6.5(a) and (b) display the initial misregistration of 11\textsuperscript{th} initial misregistration scenario and the ground truth respectively. Also Figures 6.5(c) – (e) consist of Canny edge detection overlay results for EMPCA-MI framework using original EMPCA-MI framework (1\textsuperscript{st} order (8-pixel)) and 2\textsuperscript{nd} order (8-pixel and 4-pixel) neighbourhood region connectivity respectively.
Figure 6.5: Qualitative results of monomodal T1+α40+β/T1 pair using 11th initial misregistration scenario showing (a) initial misregistration and (b) ground truth followed results for (c) original EMPCA-MI framework (1st order (8-pixel)), then EMPCA-MI framework with (d) 2nd order (8-pixel) and (e) 2nd order (4-pixel) neighbourhood region connectivity enhancements (with inset showing the zoomed-in central lobes).

Clearly, the EMPCA-MI framework using 2nd order (4-pixel) neighbourhood region connectivity outperformed the other two connectivity schemes. This is evid-
ent from the aligned central lobes in Figure 6.5(e), which is in correspondence to
the ground truth illustrated in Figure 6.5(b). The other two connectivity schemes
have significant misalignments, as seen in Figures 6.5(c) and (d).

Hence, the EMPCA-MI framework using 2\textsuperscript{nd} order (4-pixel) is consistently
able to provide better registration performance in terms of RE and qualitative
assessment with also the lowest computational cost in comparison to both the 2\textsuperscript{nd}
order (8-pixel) and original 1\textsuperscript{st} order (8-pixel) connectivity schemes.

The aforementioned connectivity schemes explored as enhancement for EMPCA-
MI framework are not exhaustive and can be deployed in many more shapes as well
as sizes. In order to explore these possibilities and investigating them, they are
considered as one of the major future works which are discussed later in Chapter
8.

6.3 EMPCA-MI Framework Enhancement 2

The standard EMPCA-MI framework computes MI between the principal com-
ponents $X_R$ and $X_S$ to determine the final EMPCA-MI value using fixed number
of bins value $b=256$. While, the registration results in Chapter 5 showed the
framework was robust and efficient in its performance with lower computational
overheads, due to the challenging characteristics of some medical modalities, such
as retinal images, the EMPCA-MI framework performance can deteriorate. For
example the inherently large homogeneous regions in retinal images mean that
choosing a high number of bins for probability distribution estimation can produce
many empty bins and more inaccurate MI value. This provided the motivation to develop a formal mechanism for the best number of bins selection in the Step III of EMPCA-MI framework, which is based on image statistics.

6.3.1 Number of bins for MI computation

In the MI literature, as noted in (Legg et al., 2013), the role of number of bins selection has often been overlooked with initial works (Collignon et al., 1995; Pluim et al., 2003) not specifically stating the chosen choices for number of bins. Though later on (Maes et al., 1997) have primarily relied on experimental selection rather than any statistical basis, they still highlight that an appropriate choice of number of bins for MI computation can significantly improve the overall registration performance (Breitner et al., 2008; Nam et al., 2009; Ritter et al., 1999). These optimal number of bins selection methods, used for probability distribution estimation to compute MI, were also successfully employed for registration applications in remote sensing domain (Moigne et al., 2011). While, some methods employ empirically testing of fixed number of bins in the range from 256 to 4 bins, in a binary-reducing approach (Kang et al., 2011; Tsao, 2003; Zhu and Cochoff, 2002), others use different statistical analysis (Birgé and Rozenholc, 2006; Davies et al., 2009, 2007; Parmehr et al., 2013) such as skewness and kurtosis factors to choose the most appropriate number of bins for MI computation.

Hence, different fixed and adaptive number of bins selection approaches were considered to be incorporated for individual and joint probability estimation for more accurate MI computation between $X_R$ and $X_S$ in Step III of EMPCA-MI
framework (Reel et al., 2014a). The first set of approaches investigated empirically reducing the fixed number of bins from 256 in the original EMPCA-MI framework to 128, 64, 32, 16, 8 and 4 bins respectively and analysing the corresponding impact on registration performance. Intuitively, reducing the number of bins will improve the probability distribution estimation, since the data is now distributed across fewer sparse bins. Conversely, radically reducing the number of bins will lead to a loss of the unique features, which could lead to inaccurate MI between them. Furthermore, in terms of computation cost, the approach of fixed number of bins is faster since the number of bins are predefined and not iteratively computed, though more bins require more computation time for calculating the individual and joint probabilities.

An alternative number of bins selection strategy was to investigate the various statistical methods which are adaptive in nature, as they are dependent on the statistical characteristics of the given data $Z$. Sturges (1926) proposed a rule which utilised the properties of $Z$ assuming it to be normal, then its histogram can be approximated with a binomial distribution. Hence the Sturges Rule defined the number of bins as:

$$\frac{R}{(1 + \log_2(w))}$$

(6.2)

Here, $R$ and $w$ represent the range and number of elements in $Z$ respectively.

Since Sturges Rule assumes $Z$ to be normally distributed, it will give inaccurate results when this condition is not met. Meanwhile, Scott (Scott, 1979) proposed a rule similar to Sturges Rule-based on the standard deviation of $Z$. It defined the
number of bins as:

\[ R \frac{3.49\sigma_{w^{-\frac{1}{3}}}}{w} \]  

(6.3)

Here, \( \sigma \) represents the standard deviation of \( Z \).

While Scott Rule also assumed \( Z \) to be normal, it would provide number of bins based on the actual values of \( Z \). Various closely related variants of Scott and Sturges Rules also exist which have change the number of bins based on similar properties (Devroye and Gyorfi, 1985; Izenman, 1991; Kanazawa, 1993; Taylor, 1987). Later, Scott proposed a new extension to its own previous work by multiplying by a skewness factor (Scott, 1992), defined as:

\[ \text{skewness factor} = \frac{2^{\frac{1}{2}} \sigma}{e^{2 \frac{\sigma^2}{4}}(\sigma^2 + 2)^{\frac{1}{2}}(e^{\sigma^2} - 1)^{\frac{1}{2}}} \]  

(6.4)

The amount of skewness allowed the Scott’s New Rule (denoted as Scott1) to recommend slightly higher number of bins, which improved the registration performance. Similarly, the peak or flat nature of the distribution in comparison to normal distribution can also be estimated using kurtosis. While, higher kurtosis signifies the steep peak close to the mean of \( Z \), a lower kurtosis denotes that \( Z \) has a much flatter distribution. This was deployed by (Wichard et al., 2008), which adapted the Sturges Rule to incorporate kurtosis such that the number of bins is:

\[ R \frac{\log_2(n) + 1 + \log_2(1 + K(\frac{n}{k})^{\frac{1}{2}}))}{log_2(n) + 1 + \log_2(1 + K(\frac{n}{k})^{\frac{1}{2}})} \]  

(6.5)
Here, $\bar{Z}$ and $K$ represent the mean and kurtosis of $Z$ respectively.

Here, including the higher-order moments (skewness and kurtosis) of $Z$ will accurately model their flat or peak distribution nature, which will be better represented in the number of bins selection strategy. Though, inclusion of higher-order moments will impact on the computational times for Scott and Wichard, it will be marginally lower for Sturges and Scott. Overall, adaptive number of bins selection approaches will incur higher computational costs because of their iterative nature, as the number of bins will be computed for every iteration.

These different approaches for Step III of EMPCA-MI framework will provide a mechanism for determining the best number of bins for accurate MI calculation between $X_R$ and $X_S$. In comparison to the fixed 256 bins used for original EMPCA-MI framework, these approaches will introduce a new degree-of-freedom based on the statistical characteristics of the reference and sensed images. Hence, a fixed number of bins approach (along with the 256 bins) and four adaptive selection methods (Scott, 1979, 1992; Sturges, 1926; Wichard et al., 2008) will be incorporated into the MI computation (Step III) of EMPCA-MI framework as shown in Figures 6.1(a) and (c), and a comparative analysis will be undertaken to evaluate their impact on the overall registration performance.
6.3.2 Experimental Setup

In order to evaluate the performance of the EMPCA-MI framework with the adaptive number of bins enhancement, various registrations experiments were conducted using the generic registration model and EMPCA-MI with parameter settings same as mentioned in Tables 3.1 and 5.1, respectively, except the various fixed and adaptive number of bins for MI computation in the EMPCA-MI framework. These registration experiments were especially performed on the monomodal pairs of retinal images from the Brno-Mono dataset (details in Section 3.4.2) for the analysis since they are considered challenging for registration due to the presence of different pathologies including haemorrhages, retinal scars and clumping of the dark pigment which all accentuated the registration. Unlike in the previous chapter where four initial misregistration scenarios were used, this set of experiment only uses one initial misregistration scenario \((100, 100, 45^\circ, 0.8)\), which had emerged as the most challenging. Therefore, this experiment consisted of 484 registrations which included 44 monomodal image pairs, 1 initial misregistration scenario and 11 methods (including the 7 fixed number of bins settings along with 256 for original EMPCA-MI framework and the 4 adaptive number of bins strategies).

6.3.3 Experimental Results and Discussion

Now the results for registration experiments corresponding to the adaptive number of bins enhancement of Step III of EMPCA-MI framework are presented. The detailed boxplot in Figure 6.6 summarises the RE results of the retinal registration experiments using different fixed and adaptive techniques integrated with the original EMPCA-MI framework.

It is evident from Figure 6.6 that in terms of registration performance, that the Wichard adaptive approach (Wichard et al., 2008) has the lowest mean RE of 33.65 pixels since it iteratively computes the best number of bins based upon the characteristics of $X_R$ and $X_S$ for each monomodal retinal pair. In terms of the empirically fixed number of bins approaches, 32 bins performed best with a mean RE of 51.82 pixels. The boxplot reveals a clear trend the mean RE decreases as the fixed number of bins is reduced from the standard 256 bins of the original EMPCA-MI framework to 32, before it then starts increasing again when the number of bins are further reduced. The reason for this is that when large numbers of bins are used in the MI computation in Step III, there is a tendency to have more sparsely-populated bins within the joint histogram which leads to poorly estimated entropy. Conversely, for fewer number of bins (16, 8 and 4), unique features will tend to be assigned to the same bin, leading to a corresponding degradation in retinal registration performance.
Hence for this particular clinical dataset, 32 number of bins is most appropriate within the fixed bin approaches. This may vary for different retinal datasets depending upon the precise characteristics of the retinal dataset, which highlights one of the limitations of adopting a fixed number of bins approach. For adaptive number of bins selection, the results show that Wichard approach (Wichard et al., 2008) achieved the lowest mean RE followed by Scott1 adaptive strategy (Scott, 1992). These techniques include higher-order moments of kurtosis and skewness of the data distribution which assists to better model the retinal characteristics. In contrast, the performance of both Sturges (Sturges, 1926) and Scott (Scott, 1979) is much lower which is due to the fact that their underlying assumption is that the retinal data is normally distributed which leads to inaccurate MI computation and higher mean RE. In summary, considering the mean RE values, it is clear from Figure 6.6 that the EMPCA-MI framework using Wichard adaptive approach is able to reduce the RE by up to 68% in comparison to the original EMPCA-MI using 256 fixed bins.

Moreover, Figure 6.7 shows the computational overheads in terms of ART for EMPCA-MI framework-based retinal registration, when integrated with different fixed and adaptive number of bins selection strategies. It is clear from Figure 6.7 that the ART decreases for fewer fixed number of bins e.g., 1.32 s to 0.86 s for 256 to 4 bins respectively, due to the lower individual and joint probability computational times incurred for fewer number of bins. While, higher ART of 3.98 s and 3.14 s are respectively observed for the two adaptive approaches of Wichard and Scott1 since they required the calculation of higher-order moments in their number of bins selections for determining the respective individual and joint probabilities to
compute MI and provide much lower mean RE in comparison to other strategies.

Finally, Figure 6.8 shows the qualitative results for Image Pair 22, using the checkerboard overlay, with reference and sensed retinal images in light and dark respectively. This particular retinal pair was especially challenging as it includes pathologies along with low contrast and large homogeneous regions. Hence it was chosen for qualitative comparison. Figure 6.8(a) and (b) shows the initial misregistration and ground truth respectively. Furthermore, Figure 6.8(c) consists of the zoomed in ground truth for a given region and the corresponding results for different fixed and adaptive number of bins selection strategies. The superior continuity of the vessel structures is evident for the Wichard adaptive approach and fixed 32 bins. It also validates their effective qualitative performance in contrast to employing either 256 bins-based standard EMPCA-MI framework or minimum number of bins (4 bins) in EMPCA-MI. Clearly, using 32 bins gives both the best numerical and qualitative results amongst fixed number of bins considered.
Figure 6.8: Qualitative results of monomodal Fundus retinal Image Pair 22, showing (a) initial misregistration, (b) ground truth and (c) the zoomed-in regions of ground truth and different fixed and adaptive number of bins selection strategies with EMPCA-MI framework.
because it represents the probability distribution most accurately, leading to an overall retinal performance improvement. But as mentioned earlier, this choice of fixed number of bins was empirical in nature and hence provided better registration performance on this particular retinal dataset.

Hence, overall, it is confirmed that using Wichard adaptive approach in Step III of the EMPCA-MI framework gave better registration performance in terms of quantitative RE and qualitative assessment in comparison to other with a trade-off in terms of computational cost.

6.4 Summary

This chapter proposed and investigated Enhancement 1 and 2 in the Step I and III of EMPCA-MI framework respectively. Enhancement 1 uses the 2nd order neighbourhood region connectivity for the pre-processing, while Enhancement 2 proposes to use different fixed and adaptive number of bins selection strategies for the MI computation in EMPCA-MI framework. The experimental results for Enhancement 1 of EMPCA-MI framework showed that 2nd order (4-pixel) neighbourhood region connectivity provided robust registration performance for BrainWeb dataset and was also computationally more efficient in comparison to the 2nd order (8-pixel) as well as the original EMPCA-MI using 1st order (8-pixel) scheme. Similarly, the experimentation for Enhancement 2 of EMPCA-MI framework showed that fixed 32 bins and Wichard adaptive number of bins selection strategy provided lower RE performances for the retinal Brno-Mono dataset amongst all the fixed and adaptive approaches analysed. Although Wichard approach provided lowest
overall RE, it still had a trade-off in terms of higher computational cost in comparison to others. Furthermore, all the experiments for these enhancements used the underlying assumption of choosing only one principal component \( (P=1) \) in the Step II of the EMPCA-MI framework. Therefore, the next chapter will propose a mechanism for relaxing this assumption and allow adaptively choosing the number of principal components within the EMPCA-MI framework. In addition, it will also integrate the enhancements presented in this chapter to make all the steps of EMPCA-MI framework more effective.
Chapter 7

Adaptive EMPCA-MI Framework

7.1 Introduction

In the previous chapter, Steps I and III of EMPCA-MI framework in Figure 6.1, were modified by respectively utilising Enhancement 1 and 2, both of which incurred corresponding computational costs. The experimental results for these two enhancements were independently evaluated and both provided improved and robust registration performance for different medical datasets. These datasets are difficult to register due to their inherent characteristics. However, the EMPCA-MI framework still makes the same assumption as in Chapter 4, of calculating only the first principal component using EMPCA in Step II.

Interestingly, in Section 5.3.3 it was shown that by empirically increasing the number of principal components (from $P=1$ to $P=2$), the framework provided considerably improved performance for multimodal retinal registration of Fundus/SLO
image pair from the *Brno-Multi* dataset. This was due to the fact that two orthogonal principal components provided more variance than a single component for this retinal image pair. This provided the impetus to formalise the new EMPCA-MI framework so that it is able to adaptively determine the most appropriate number of principal components to represent prominent neighbourhood region features for a particular dataset.

This chapter proposes an *adaptive* EMPCA-MI framework for MIR. It exploits the *Kaiser Rule* (Jackson, 1993; Jolliffe, 2002) to iteratively find the best number of principal components to be computed by the EMPCA framework in Step II of the registration process. In addition, it also employs both the *Enhancement 1* and *2* proposed in *Chapter 6* for Steps I and III.

Two sets of experiments using human brain *BrainWeb* and retinal *DRIVE* and *Brno-Multi* datasets confirm that the *adaptive* EMPCA-MI framework provides consistently lower RE and improved robustness with a corresponding trade-off in computational overheads.

### 7.2 *Adaptive* EMPCA-MI framework

This section describes the *adaptive* EMPCA-MI framework (Reel et al., 2014b). *Section 7.2.1* briefly discusses the concept of stopping rules for PCA, before *Section 7.2.2* explores the use of applying the *Kaiser Rule* (Jackson, 1993; Jolliffe, 2002) to find the most appropriate number of principal components for EMPCA.
7.2.1 Stopping Rules for PCA

In PCA, selecting the best number $P$ of principal components is a major challenge, because while retaining too many components can potentially lead to overfitting by including noise in the data, choosing too few components can equally mean discarding valuable information (Jackson, 1993). Commonly the first principal component is empirically chosen since it accounts for the majority of variability in the data (Rencher, 1998), and so was used for the EMPCA-MI framework. But as mentioned in Chapter 2, various approaches for selecting the best subset of significant principal components have been proposed including Scree Graph, Broken-stick and their variants (Cangelosi and Goriely, 2007; Jolliffe, 2002), which select a subset by computing the cumulative variance for all components. In contrast, the Kaiser Rule (Jackson, 1993; Jolliffe, 2002) retains all principal components whose eigenvalues are greater than one. It assumes that image data variables are independent. Hence their principal components are same as variable, having unity variance in case of correlation matrix. Hence, any component with an eigenvalue less than one is assumed to not contain valuable information (Ravi and Palaniappan, 2005). This important attribute of the Kaiser Rule is image dependent, so it can automatically choose the number of significant principal components required to represent image features for a given accuracy. Furthermore, it is suited to the EMPCA-MI framework since it iteratively determines the $P$ principal components in descending order of eigenvalues. These two distinct and relevant characteristics provided the motivation to investigate incorporating the Kaiser Rule into the Step II of the EMPCA-MI framework to determine the best $P$ value choice for a particular dataset.
7.2 Adaptive EMPCA-MI framework

This section describes the adaptive mechanism for choosing $P$ for the EMPCA-MI framework. The three primary computation steps include the pre-processing of neighbourhood region information for both $I_R$ and $I_S^*$ followed by adaptively determining the $P$ principal components using EMPCA. Finally, the MI output is computed. Figure 7.1 shows the flowchart diagram of the adaptive EMPCA-MI framework with these three steps indicated.

Step I of adaptive EMPCA-MI pre-processes the various image grayscale values by using the Enhancement 1-based 2$^{nd}$ order (4-pixel) neighbourhood region connectivity between them as described in Section 6.2. This retains the local spatial neighbourhood relationship rather than using pixel values as in 1$^{st}$ order (8-pixel)
neighbourhood relations. Hence, by using (6.1), $Q_R$ and $Q_S$ are computed using $I_R$ and $I^*_S$ respectively.

*Step II* of the *adaptive* EMPCA-MI proposes the dimensionality reduction for $Q_R$ and $Q_S$ using *Kaiser Rule* (Jackson, 1993; Jolliffe, 2002) to adaptively find the most significant $P$ principal components using EMPCA. The flowchart in Figure 7.1 shows how the *adaptive* EMPCA computes the first $P$ principal components from $Q_R$ having eigenvalue $\lambda _P > 1$. Similarly the first $P'$ principal components from $Q_S$ are computed with these then concatenated as $X_R$ and $X_S$ respectively. Crucially, unlike $P=1$ assumption considered in earlier chapters, this adaptive approach retains accurate image features which will be reflected in the corresponding registration performance.

Finally, *Step III* of *adaptive* EMPCA-MI determines the MI between the highest number of common principal components in both $X_R$ and $X_S$ of the reference and sensed images. This step adopts the *Enhancement 2* proposed in Section 6.3, with the *Wichard* (Wichard et al., 2008) approach chosen to select the best number of bins to calculate the individual and joint probabilities, since it exploits the *kurtosis* measure of the data distribution to facilitate a more accurate MI value. Next section discusses the experimental setup used for the registration experiments.

### 7.3 Experimental Setup

In order to assess the performance of the *adaptive* EMPCA-MI framework, various registration experiments were conducted. The nomenclature used in this chapter
## 7.3 Experimental Setup

### Adaptive EMPCA-MI framework

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Selection</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neighbourhood radius ($r$)</td>
<td>1 (using Enhancement 1)</td>
</tr>
<tr>
<td>Number of Principal Components ($P$)</td>
<td>Adaptively (using Kaiser Rule)</td>
</tr>
<tr>
<td>Number of bins ($b$)</td>
<td>Variable (using Enhancement 2)</td>
</tr>
</tbody>
</table>

**Table 7.1:** The different parameter selections for the adaptive EMPCA-MI framework.

is adaptive EMPCA-MI \{P\}, where $P$ denotes the number of principal components determined by the Kaiser Rule in Section 7.2.2. Similarly, the EMPCA-MI framework with $P=1$ in Chapter 4, 5 and 6 will be denoted as EMPCA-MI \{1\}. These notations allow a clear distinction to be made between both the original and adaptive EMPCA-MI frameworks in the subsequent result analysis. To evaluate adaptive EMPCA-MI \{P\}, the registration experiments were categorised into two sets (Set 1 and Set 2) as in Table 5.2, with the parameter selections for the adaptive EMPCA-MI \{P\} and EMPCA-MI \{1\} models being given in Tables 7.1 and 5.1, respectively. The MIR model used for experiments employ the same parameter settings as in Table 3.1.

The Set 1 of registration experiments used the mono and multimodal pairs of MRI T1, T2 and PD human brain images of the BrainWeb dataset, with all experimental settings as in the previous chapters. In all 1200 individual registration experiments were performed in this set, which included: 3 monomodal pairs, 3 multimodal pairs, 5 levels of INU and Gaussian noise settings, 20 initial misregistration scenarios and 2 similarity measures, including both the adaptive EMPCA-MI \{P\} and EMPCA-MI \{1\} frameworks.
The Set 2 of registration experiments were performed on the mono and multimodal pairs of retinal images from the DRIVE and Brno-Multi datasets respectively. A detailed description of these datasets was discussed earlier in Section 3.4.2. The monomodal subset using Fundus retinal images of DRIVE dataset consists of 160 registrations which include 20 monomodal pairs, 4 initial misregistration scenarios and 2 similarity measures including the adaptive EMPCA-MI \{P\} and EMPCA-MI \{I\} frameworks. Similarly for the multimodal subset using Fundus and SLO retinal images from the Brno-Multi dataset consists of 60 registrations which include 10 multimodal pairs, 3 initial misregistration scenarios and 2 similarity measures including the adaptive EMPCA-MI \{P\} and EMPCA-MI \{I\} frameworks.

7.4 Experimental Results Discussion

7.4.1 Set 1: BrainWeb Dataset Registration Experiments

This section discusses the registration experiments results for adaptive EMPCA-MI \{P\} on Set 1 of the BrainWeb dataset, with EMPCA-MI \{I\} used as the comparator. Figure 7.2(a) and (b) show the detailed summary boxplot for all INU and noise levels of 20 initial misregistration scenarios for the monomodal T1/T1 and multimodal T1/T2 registration experiments respectively. Each boxplot shows a bounding box defining the interquartile range with the bar across representing the median and whiskers denoting the RE range. The boxplots also include the number of out-of-range RE values corresponding to either the EMPCA-MI \{I\} or
adaptive EMPCA-MI \( \{ P \} \) frameworks.

The two boxplots in Figure 7.2 confirm the superior registration quality of adaptive EMPCA-MI \( \{ P \} \) with both lower mean RE and smaller interquartile ranges compared with EMPCA-MI \( \{ 1 \} \). For the T1/T1 pairs in Figure 7.2(a), no out-of-range outliers with high RE are observed for adaptive EMPCA-MI \( \{ P \} \) with the exception of the monomodal T1+\( \alpha_{40} \)+\( \beta \)/T1 pair. In contrast the RE monotonically increases in the case of EMPCA-MI \( \{ 1 \} \) reflecting its poor convergence in achieving the best alignment between the reference and sensed images. This trend is also manifest in Figure 7.2(b) for the multimodal T1/T2 pairs, with again the improved registration performance of adaptive EMPCA-MI \( \{ P \} \) clearly evident.

Similar conclusions can be deduced from the boxplots for the remaining mono (T2/T2 and PD/PD) and multimodal (T2/PD and PD/T1) registration experiments which for completeness, are included in Figures F.1 and F.2 respectively in Appendix F.

In analysing all the BrainWeb dataset results for different registrations, adaptive EMPCA-MI \( \{ P \} \) has been able to reduce RE by an average of 84\% in comparison to the EMPCA-MI \( \{ 1 \} \). Similarly, comparing the adaptive EMPCA-MI \( \{ P \} \) with MI, GMI and RMI (See Section 5.3.1) reveals an overall RE reduction of by an average of 93\%. These results confirm that adaptive EMPCA-MI \( \{ P \} \) is able to more accurately extract features from the reference and sensed images enabling a consistently superior quality of registration, even in the presence of significant INU and noise levels. In contrast, EMPCA-MI \( \{ 1 \} \) performs satisfactorily only when little or no INU is present in the images and as the levels of INU and Gaussian noise increase, its performance deteriorates.
In order to demonstrate the performance of adaptive EMPCA-MI \{P\} especially in presence of higher INU and noise, one of the individual registrations of monomodal T1+0.40/β/T1 pair shown in Figure 7.2(a) is analysed. Registration of the T1+0.40/β/T1 pair with the 14th initial misregistration scenario (−26.32,
0.55, 33.23°) has an initial RE of 82.27 pixels. The adaptive EMPCA-MI \{3\} for this registration computed with Kaiser Rule in combination with Enhancement 1 and 2, provides a significantly lower final RE value of just 1.59 pixels. Conversely, the EMPCA-MI \{1\} provides a final RE value of 98.91 pixels and therefore it is demarcated as an out-of-range outlier in Figure 7.2(a). Interestingly, reflecting on the registration performance of the T1+α_{40}+β/T1 pair with same initial misregistration scenario analysed in Chapter 6, EMPCA-MI \{1\} using Enhancement 1 provided final RE of 9.01 pixels, which is seen as an outlier in Figure 6.3(a). This highlights the improved performance attained by the adaptive EMPCA-MI \{3\} by choosing higher numbers of principal components and more accurate MI computation, to facilitate lower RE and better convergence, even for registration cases which have higher INU and noise levels.

Figure 7.3 shows the qualitative performance for the above mentioned monomodal T1+α_{40}+β/T1 pair registration. The initial misregistration and the corresponding Canny edge detection ground truth are displayed in Figures 7.3(a) and (b) respectively, with Figures 7.3(c) and (d) showing the corresponding registered images using the EMPCA-MI \{1\} and adaptive EMPCA-MI \{3\} frameworks. The latter clearly outperforms the EMPCA-MI \{1\} in terms of its alignment with the ground truth, as the three zoomed-in areas confirm.

Analogous findings are observed for the multimodal registrations. For example, the multimodal T1+α_{40}+β/T2 pair shown in Figure 7.2(b) for the 13th initial misregistration scenario (−6.45, −20.83, −32.64°) has an initial RE of 80.60 pixels. The adaptive EMPCA-MI \{2\} for this registration computed with Kaiser Rule and using Enhancement 1 and 2, affords a final RE value of only 0.67 pixels. In con-
7.4 Experimental Results Discussion

Figure 7.3: Qualitative results of monomodal T1+\(\alpha\_\text{ref}+\beta\)/T1 pair using 14\textsuperscript{th} initial misregistration scenario showing (a) initial misregistration, (b) ground truth and (c) EMPCA-MI \{1\} and (d) adaptive EMPCA-MI \{3\} output (with inset showing the zoomed-in central lobes).

Contrast, EMPCA-MI \{1\} provided a final RE value of 21.85 pixels and so has been demarcated as an out-of-range outlier in Figure 7.2(b). Similarly, EMPCA-MI \{1\} using Enhancement 1 gave a final RE of 17.18 pixels, and was also designated as an outlier in Figure 6.3(b).

Figure 7.4 shows the corresponding qualitative performance for this particular multimodal T1+\(\alpha\_\text{ref}+\beta\)/T2 pair registration, which has high INU and Gaussian noise. Figures 7.4(a) and (b) display the initial misregistration and the corresponding Canny edge detection ground truth. While the registered images using the EMPCA-MI \{1\} and adaptive EMPCA-MI \{2\} framework are shown in Figures
7.4 Experimental Results Discussion

Figure 7.4: Qualitative results of multimodal T1+$\alpha_{40}+\beta$/T2 pair using 13th initial misregistration scenario showing (a) initial misregistration, (b) ground truth and (c) EMPCA-MI \{1\} and (d) adaptive EMPCA-MI \{2\} output (with inset showing the zoomed-in central lobes).

7.4(c) and (d) respectively, with strikingly adaptive EMPCA-MI \{2\} providing a better perceptual outcome than EMPCA-MI \{1\} when compared with the ground truth shown in Figure 7.4(b).

These mono and multimodal results confirm that the adaptive EMPCA-MI \{P\} framework is able to achieve more consistent improved registration performance for the BrainWeb dataset, even in the presence of high INU and noise levels.
7.4 Experimental Results Discussion

178

Registration Error (RE) of EMPCA-MI $\{1\}$ and Adaptive EMPCA-MI $\{P\}$ for Image Pairs 1 to 20.

Figure 7.5: RE summary bar graph of 20 monomodal Fundus retinal image pairs using EMPCA-MI $\{1\}$ and adaptive EMPCA-MI $\{P\}$ for the 3rd initial misregistration scenario.

7.4.2 Set 2: DRIVE and Brno-Multi Dataset Experiments

This section analyses the registration results for adaptive EMPCA-MI $\{P\}$ on Set 2 of monomodal retinal DRIVE dataset and multimodal retinal Brno-Multi dataset. To draw a comparison as in earlier section, EMPCA-MI $\{1\}$ was used as a comparator.

The first subset of registration experiments of DRIVE dataset used four initial misregistration scenarios, the same as in Chapter 5. In the first two initial misregistration scenarios, only one registration parameter (namely rotation $\theta$) was varied, while in the last two scenarios all the four registration parameters are varied. Figure 7.5 shows the RE summary bar graph of 20 monomodal Fundus image pairs using EMPCA-MI $\{1\}$ and adaptive EMPCA-MI $\{P\}$ for the 3rd initial misregistration scenario ($5, 5, 20^\circ, 2.8$). This bar graph is plotted with twice broken $y$ axis in order to accommodate all the RE values on the same plot as well as to equitably show the lower RE values. The RE results confirm adaptive EMPCA-MI $\{P\}$ provided superior registration performance with consistently lower RE compared
with the EMPCA-MI \{I\} in challenging contrast and illumination conditions.

For example, the RE for Image Pair 20 using adaptive EMPCA-MI \{P\} is nearly half that of EMPCA-MI \{I\}. It was stressed in Chapter 5, that for this particular Image Pair, EMPCA-MI \{I\} provided poorer registration performance than GFM as shown earlier in Figure 5.7, because EMPCA-MI \{I\} uses only one principal component, which is insufficient for the inherent challenging homogeneous regions and leads to its underperformance. On the other hand, adaptive EMPCA-MI \{3\} with Kaiser Rule computation and using Enhancement 1 and 2 for this particular Image Pair gives a better registration performance with RE reduction of 54% and 52% respectively compared with EMPCA-MI \{I\} and GFM.

The qualitative assessment of the registration for Image Pair 20 can be observed in Figure 7.6. It shows the checkerboard overlay results with the initial misregistration in Figure 7.6(a) and the ground truth with three marked regions for visual inspection in Figure 7.6(b). Figures 7.6(c) and (d) show the registration output for EMPCA-MI \{I\} and adaptive EMPCA-MI \{3\} frameworks, with the correspondingly marked regions. Moreover, Figure 7.6(e) shows the zoomed three marked regions of the ground truth and two registration outputs. It is clear from the Figure 7.6(e) results that adaptive EMPCA-MI \{3\} gives better vascular structure continuity in comparison to EMPCA-MI \{I\}, and in terms of alignment, is also much closer to the ground truth by virtue of adaptively choosing more principal components to better represent the localised features in Image Pair 20.

Finally, Figure 7.7 summarises the first subset of Set 2 registration experiments. It shows the RE summary boxplot of all monomodal Fundus retinal image pairs.
Figure 7.6: Qualitative results of monomodal Fundus retinal Image Pair 20 using the $3^{rd}$ initial misregistration scenario, showing (a) initial misregistration, (b) ground truth, (c) EMPCA-MI $\{I\}$ and (d) adaptive EMPCA-MI $\{3\}$ output. Also (e) shows the zoomed-in regions of ground truth ($1^{st}$ row), EMPCA-MI $\{I\}$ ($2^{nd}$ row) and adaptive EMPCA-MI $\{3\}$ ($3^{rd}$ row).
7.4 Experimental Results Discussion

Figure 7.7: RE summary boxplot of 20 monomodal Fundus retinal image pairs using EMPCA-MI \{1\} and adaptive EMPCA-MI \{P\} for all 4 initial misregistration scenarios. The mean and outliers are denoted by ♦ and + shapes. The numbers on top of the plot refer to the number of out-of-range outliers with RE higher than 18 pixels.

Now considering the results for second subset of Set 2 using multimodal Fundus/SLO Brno-Multi dataset, as in Chapter 5, this subset utilises three separate initial misregistration scenarios. In the first, just one registration parameter ($\theta$) is varied, while in the other two scenarios all the four registration parameters are varied. Figure 7.8 shows the corresponding RE summary bar graph of 10 multimodal Fundus/SLO image pairs using adaptive EMPCA-MI \{P\} and EMPCA-MI \{1\} for the 2\textsuperscript{nd} initial misregistration scenario ($10, 10, 30^\circ, 2$). It is apparent adaptive
7.4 Experimental Results Discussion

**Figure 7.8:** RE summary bar graph of 10 multimodal Fundus/SLO retinal image pairs using EMPCA-MI \{1\} and adaptive EMPCA-MI \{P\} for the 2\textsuperscript{nd} initial misregistration scenario.

EMPCA-MI \{P\} again consistently provides lower RE than EMPCA-MI \{1\} for all Image Pairs. It is especially notable to scrutinise the results for the challenging Image Pair 6, which previously in Figure 5.11, produced a higher RE for EMPCA-MI \{1\} in comparison to RIR-BS. Using adaptive EMPCA-MI \{2\} with Kaiser Rule, together with Enhancement 1 and 2 now gives better registration with an average RE reduction of 35\% and 33\% respectively compared to EMPCA-MI \{1\} and RIR-BS.

To affirm this superior performance, Figure 7.9 shows the matching qualitative results for Image Pair 6 for the 2\textsuperscript{nd} initial misregistration scenario, with three zoomed-in regions shown in Figure 7.9(e). Figures 7.9(a) and (b) show the initial misregistration for the 2\textsuperscript{nd} scenario and the ground truth respectively with three marked regions. Also, Figures 7.9(c) and (d) show the registration output for EMPCA-MI \{1\} and adaptive EMPCA-MI \{P\} with corresponding marked regions respectively. Figure 7.9(e) shows the zoomed three marked regions of the ground truth and two registration outputs. Note, in all three marked regions, adaptive EMPCA-MI \{2\} provides enhanced structural continuity com-
7.4 Experimental Results Discussion

Figure 7.9: Qualitative results of multimodal Fundus/SLO retinal Image Pair 6 using the 2\textsuperscript{nd} initial misregistration scenario, showing (a) initial misregistration, (b) ground truth, (c) EMPCA-MI \{1\} and (d) adaptive EMPCA-MI \{2\} output. Also (e) shows the zoomed-in regions of ground truth (1\textsuperscript{st} row), EMPCA-MI \{1\} (2\textsuperscript{nd} row) and adaptive EMPCA-MI \{2\} (3\textsuperscript{rd} row).
pared with EMPCA-MI \{I\} by adaptively determining $P$ directly from the images and thereby leading to consistently superior and more robust registration, even for challenging Image Pairs.

Figure 7.10 summarises the second subset of Set 2 registration experiments for the multimodal Fundus/SLO retinal image pairs using EMPCA-MI \{I\} and adaptive EMPCA-MI \{P\} for the three initial misregistration scenarios. The results reveal the superior registration performance of adaptive EMPCA-MI \{P\} in terms of both lower mean RE and interquartile range over EMPCA-MI \{I\} in all the scenarios. Moreover, the underperformance of EMPCA-MI \{I\} is evident from the presence of one out-of-range outlier in last two scenarios, while in contrast adaptive EMPCA-MI \{P\} is able to consistently afford lower RE with no out-of-range outliers. Overall these multimodal results of adaptive EMPCA-MI \{P\} provide a respective average reduction in RE by 59% and 39% in comparison to those of
RIR-BS and EMPCA-MI \{1\} in Section 5.3.2.

In summary, these findings confirm that adaptive EMPCA-MI \{P\} provides better image registration performance compared to EMPCA-MI \{1\} with an overall 45% reduction for the DRIVE and Brno-Multi datasets, which are especially characterised by the presence of low contrast and large homogeneous non-vascular regions.

### 7.4.3 Computational Cost of Adaptive EMPCA-MI Framework

As shown earlier in Section 4.3.2, the order of computational complexity for the EMPCA \{1\} was $O(qd)$. Similarly for the adaptive EMPCA \{P\}, the order of complexity is $O(Pqd)$, but as $q >> P$, $d$ the order will be same as EMPCA \{1\}, though the iterative nature of the Kaiser Rule to determine the best $P$ means the overall time cost will tend to be slightly higher. Adaptive EMPCA-MI \{P\} can thus be viewed as a generalisation of the EMPCA-MI framework, with EMPCA-MI \{1\} being a special case.

While the adaptive EMPCA-MI \{P\} framework more effectively registers the mono and multimodal BrainWeb, DRIVE and Brno-Multi medical datasets, it does incur a commensurate increase in computational time due to its inherently iterative nature. While the order of computational complexity remains the same as EMPCA-MI \{1\}, it is insightful to undertake an ART analysis to evaluate the extra cost of using the Kaiser Rule to determine the best $P$ value.
7.4 Experimental Results Discussion

Figure 7.11: ART performance for EMPCA-MI \(\{1\}\) and adaptive EMPCA-MI \(\{P\}\) frameworks for Set 1 and Set 2 registration experiments.

Figure 7.11 shows the ART performance of adaptive EMPCA-MI \(\{P\}\) for registration of the Sets 1 and 2 clinical datasets in comparison with EMPCA-MI \(\{1\}\). Clearly the adaptive EMPCA-MI \(\{P\}\) computation time for both Set 1 and 2 registration results is higher, with in case of Set 1, the increase being 25% with respect to EMPCA-MI \(\{1\}\) and 35% for Set 2. Note, the reason for the overall higher Set 2 registration ART values is they reflect the greater resolution of the Fundus/SLO retinal images.

While the Enhancement 1 introduced in Chapter 6 reduces data dimensionality by using 2\textsuperscript{nd} order region connectivity, overall the adaptive EMPCA-MI \(\{P\}\) framework does incur a higher computational cost because of the inclusion of the adaptive Kaiser Rule to find \(P\) and then use Enhancement 2 to determine the best number of bins for MI computation. This pragmatically, however, is more than offset by the significant improvements achieved by the framework in superior and more robust registration performance, particularly in the processing of multimodal
image datasets characterised by high INU and noise, and large homogeneous non-vascular regions.

7.5 Summary

This chapter has presented an *adaptive* EMPCA-MI framework for MIR. It exploits one of the popular stopping rules for principal components namely the *Kaiser Rule* to adaptively find the number of principal components to be computed by the iterative EMPCA. In addition, it also employs *Enhancement 1* and *2* proposed in Chapter 6. Two sets (*Set 1* and *2*) of rigorous experiments using mono and multimodal human brain *BrainWeb* and retinal *DRIVE* and *Brno-Multi* datasets proved that *adaptive* EMPCA-MI provides consistently better registration performance with a corresponding trade-off in computational cost. Overall, the numerical and qualitative results for *adaptive* EMPCA-MI show improved robustness and consistently lower RE in comparison to EMPCA-MI using only one principal component. The next chapter presents some ideas for the future to extend the findings from this thesis.
Chapter 8

Future Work

There are a number of potential opportunities to extend the EMPCA-MI similarity measure framework presented in this thesis, as well as exploring other domains for its applicability. Some of these prospective research avenues will now be reviewed.

1. The presented work had the overarching objective of achieving effective robust registration of 2D medical images without incurring high computational cost. As most advanced medical imaging modalities nowadays generate 3D images, one promising area of investigation would be to extend the framework to 3D image registration. A new 3D region connectivity (using the 6, 18 or 26 neighbourhood) based sliding window could be considered as the pre-processing stage for the framework (J. Toriwaki and H. Yoshida, 2009; J. Toriwaki and T. Yonekura, 2002; Noël and Biri, 2015), exploiting the proven enhanced neighbourhood region relations shown in Section 6.2, while leaving the other two processing blocks of the EMPCA-MI similarity measure framework as is. In addition, it would be interesting to assess the impact
Future Work

of using different interpolations and higher neighbourhood radius for the EMPCA-MI similarity measure in the above mentioned extension, with an overall objective to reduce the computational costs.

2. The EMPCA-MI similarity measure framework processes local neighbourhood regions in order to extract spatial features. The thesis has focused upon the direct (4-pixel) and indirect (8-pixel) connectivity-based 1st/2nd order neighbourhood relationships, so the underlying idea could be extended to higher-order patterns and object shapes. More neighbourhood patterns could be inspired from the Local Binary Pattern (LBP) scheme and its recent variants (Liu et al., 2012; Pietikäinen et al., 2011), which have gained popularity in the texture classification domain due to its ability to deal with illumination variations and computational simplicity. It would be insightful to study the impact of integrating a LBP-based approach into the pre-processing stage of the EMPCA-MI similarity measure framework which can help to robustly extract the spatial information from the neighbourhood region leading to more effective registration.

3. The EMPCA-MI similarity measure framework could be applied to other domains in addition to medical imaging. While the framework determines the statistical relationship between different medical modalities, it is feasible to use it in other domains such as remote sensing image registration (Chen et al., 2003) and high dimensional hyper-spectral imagery (Yao et al., 2010). Furthermore, since it integrates local spatial information along with mutual information, it can be used for quantitative assessment of the visual quality of generic images (Li et al., 2012).
4. Although 3D registration is used in both medical and non-medical applications, the long run times involved in the convergence of these registrations is an obstacle to their deployment in time-critical applications (Argyriou et al., 2015). Hence in order to develop time sensitive registration process for such applications, it is desirable to utilise graphic processing unit-based computing. This requires granularity level visualisation of the various blocks of registration process in terms of memory and mathematical operations (Shi et al., 2012). Another extension to the EMPCA-MI similarity measure framework would be to examine graphic processing unit-based implementations for accelerated registration. The iterative and parallel process flows of EMPCA-MI could be leveraged from multi-core processor architectures, especially when deployed together with multi-resolution approaches for significant speed-ups.

5. While the EMPCA-MI similarity measure framework has been developed for rigid transform-based registration of medical images, it can be generalised to non-rigid registration scenarios for different imaging modalities as well as being utilised for sequential rigid followed by non-rigid registration in advanced biomedical applications. For example, EMPCA-MI similarity measure framework could be used for non-rigid B-Spline registration experiments on 2D mouse embryonic micro-CT data using a multi-resolution approach. This could facilitate the high-throughput mouse embryonic micro-CT data phenotyping which aims to map each mouse gene to its corresponding physiological functionality in line with International Mouse Phenotyping Consortium (Roy et al., 2013).
Chapter 9

Conclusion

Medical imaging is increasingly used for diagnosis and treatment planning applications in healthcare as well as an innovative tool for medical research. The ability to establish correspondences between locations in medical images is ubiquitous in most applications. Even though research into MIR has progressed rapidly in the last three decades, there are still many challenges. Previously, various feature and intensity-based similarity measures have been proposed for registration purposes, with MI establishing itself as a popular similarity measure due to its information-theoretic properties and ability to register multimodal medical images. Some modalities include inherent characteristics like high INU, noise and large non-vascular homogeneous regions, which makes the registration process especially challenging. Recent efforts at the robust registration of multimodal images have used hybrid similarity measures utilising spatial information from local regions along with MI, though they either incur a significant computational cost or use covariance-based approximation leading to poorer registration quality.
A new flexible, robust hybrid EMPCA-MI similarity measure framework has been presented which comprises a suite of innovative algorithms to improve the registration performance by extracting the local features from the neighbourhood region using iterative EMPCA in combination with MI to lower the computational time complexity. The new EMPCA-MI framework has been critically evaluated using different mono and multimodal clinical datasets, and makes three original contributions to the rigid image registration domain, which are summarised as follows:

1. The most significant contribution is the adaptive EMPCA-MI similarity measure framework which has the capability to flexibly choose the number of principal components from the neighbourhood region, based on the image characteristics and then to compute the MI thereby affording a key computational trade-off mechanism. The framework uses a stopping rule for guiding the iterative EMPCA to ascertain the most appropriate number of features, so EMPCA-MI becomes image dependent with the features reflecting the image information. Rigorous mono and multimodal registration experiments confirm the improved quantitative performance for the challenging brain MRI and retinal images, with the average registration error being reduced by up to 93% and 60% respectively, compared to existing similarity measures.

2. The underlying EMPCA-MI similarity measure framework also utilises local neighbourhood region information of the image, though it only employs the first iteratively computed principal component in EMPCA when computing the MI. This means it is computationally more efficient than the adaptive
EMPCA-MI similarity measure framework and crucially provides the user with the flexibility to trade between speed and registration quality. Nonetheless it has been conclusively shown to afford superior mono and multimodal registration performance compared with current techniques, with registration errors reduced by on average, 61% and 30% for the brain MRI and retinal datasets respectively.

3. Finally, a series of novel enhancements has been embedded into the EMPCA-MI similarity measure framework to further improve its performance. Enhancement 1 exploits the neighbourhood region relationships to remove information redundancy, while Enhancement 2 employs variable number of bins selection strategies to enable more accurate MI computation. Experimental results confirm both enhancements collectively provide reductions in the average registration error of up to 79% and 68% respectively, compared with the original EMPCA-MI similarity measure model.

In summarising, this new hybrid EMPCA-MI similarity measure framework makes a notable contribution to the registration performance of multimodal clinical images characterised by high INU, noise and non-vascular homogeneous regions. Most importantly, it offers a flexible and robust solution for 2D rigid registration which is able to, not only be extended to non-rigid registration, but also the 3D registration domain in the future.
References


Montreal Neurological Institute and Hospital (2000). *BrainWeb: Simulated Brain Database*.


Murphy, K., van Ginneken, B., Reinhardt, J. M., Kabus, S., Ding, K., Deng, X.,
Cao, K., Du, K., Christensen, G. E., Garcia, V., Vercauteren, T., Ayache, N.,
Commowick, O., Malandain, G., Glocker, B., Paragios, N., Navab, N., Gorbun-
ova, V., Sporring, J., de Bruijne, M., Han, X., Heinrich, M. P., Schnabel, J. A.,
Jenkinson, M., Lorenz, C., Modat, M., McClelland, J. R., Ourselin, S., Muen-
zing, S. E. A., Viergever, M. A., De Nigris, D., Collins, D. L., Arbel, T., Per-
oni, M., Li, R., Sharp, G. C., Schmidt-Richberg, A., Ehrhardt, J., Werner, R.,
Smeets, D., Loeckx, D., Song, G., Tustison, N., Avants, B., Gee, J. C., Staring,
M., Klein, S., Stoel, B. C., Urschler, M., Werlberger, M., Vandemeulebroucke,
tion methods on thoracic CT: the EMPIRE10 challenge. *IEEE transactions on*

inter-modality image registration using normalized mutual information with
coarse-binned histograms. *Communications in Numerical Methods in Engin-

matching using entropic graphs. In *IEEE International Symposium on Biomedical

Neri, E., Caramella, D., Bartolozzi, C., Baert, A. L., Knauth, M., and Sartor,
Berlin Heidelberg.


References


References


Appendices
Appendix A

Initial Misregistration Scenarios

This Appendix includes the details of the initial misregistration scenarios ($\tau_{\text{initial}}$) mentioned in Section 3.5.1 and used for mono and multimodal registration throughout this thesis. Table A.1 displays the details of twenty initial misregistration scenarios for BrainWeb dataset as adopted from (Kotsas and Dodd, 2011). Similarly, the five initial misregistration scenarios were used for DRIVE, Brno-Mono and Brno-Multi retinal datasets as adopted from (Gharabaghi et al., 2013).
### Table A.1: Initial misregistration scenarios for the BrainWeb (Collins et al., 1998) dataset as in (Kotsas and Dodd, 2011).

<table>
<thead>
<tr>
<th>Scenario No</th>
<th>Initial Misregistration ($\tau_{\text{initial}}$)</th>
<th>Initial RE (in pixels)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$x_{\text{initial}}$</td>
<td>$y_{\text{initial}}$</td>
</tr>
<tr>
<td>1</td>
<td>$-5.55$</td>
<td>$17.74$</td>
</tr>
<tr>
<td>2</td>
<td>$12.88$</td>
<td>$-8.11$</td>
</tr>
<tr>
<td>3</td>
<td>$10.11$</td>
<td>$27.22$</td>
</tr>
<tr>
<td>4</td>
<td>$26.23$</td>
<td>$-13.42$</td>
</tr>
<tr>
<td>5</td>
<td>$-17.08$</td>
<td>$12.91$</td>
</tr>
<tr>
<td>6</td>
<td>$-10.88$</td>
<td>$-11.62$</td>
</tr>
<tr>
<td>7</td>
<td>$13.62$</td>
<td>$16.41$</td>
</tr>
<tr>
<td>8</td>
<td>$16.71$</td>
<td>$-23.55$</td>
</tr>
<tr>
<td>9</td>
<td>$-9.41$</td>
<td>$-19.11$</td>
</tr>
<tr>
<td>10</td>
<td>$29.72$</td>
<td>$2.83$</td>
</tr>
<tr>
<td>11</td>
<td>$24.21$</td>
<td>$-25.20$</td>
</tr>
<tr>
<td>12</td>
<td>$-19.11$</td>
<td>$9.08$</td>
</tr>
<tr>
<td>13</td>
<td>$-6.45$</td>
<td>$-20.83$</td>
</tr>
<tr>
<td>14</td>
<td>$-26.32$</td>
<td>$0.55$</td>
</tr>
<tr>
<td>15</td>
<td>$-5.23$</td>
<td>$2.31$</td>
</tr>
<tr>
<td>16</td>
<td>$-5.65$</td>
<td>$16.94$</td>
</tr>
<tr>
<td>17</td>
<td>$0.21$</td>
<td>$-16.41$</td>
</tr>
<tr>
<td>18</td>
<td>$-23.64$</td>
<td>$-7.72$</td>
</tr>
<tr>
<td>19</td>
<td>$7.20$</td>
<td>$0.23$</td>
</tr>
<tr>
<td>20</td>
<td>$0.55$</td>
<td>$29.63$</td>
</tr>
</tbody>
</table>
### Initial Misregistration Scenarios

Table A.2: Initial misregistration scenarios for the retinal datasets as in (Gharabaghi et al., 2013).

<table>
<thead>
<tr>
<th>Scenario No</th>
<th>Initial Misregistration ($\tau_{initial}$)</th>
<th>Initial RE (in pixels)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$x_{initial}$</td>
<td>$y_{initial}$</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>5</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>
Appendix B

EMPCA-MI Results for Interpolation

This Appendix includes the detailed $x$ axis translation, $y$ axis translation and $\theta$ rotation plots of EMPCA-MI traces as mentioned in Section 4.3.1. These plots show the results for nearest neighbour, bilinear and bicubic interpolation methods for mono and multimodal T1, T2 and PD image pairs from BrainWeb dataset having five INU and noise levels (No INU and noise, $\alpha_{20}$, $\alpha_{40}$, $\beta$ and $\alpha_{40}+\beta$) as discussed in Section 3.5.2.

Figures B.1, B.2 and B.3 show the EMPCA-MI traces for monomodal T1/T1, T2/T2 and PD/PD image pairs respectively. Similarly, Figures B.4, B.5 and B.6 show the EMPCA-MI traces for multimodal T1/T2, T2/PD and PD/T1 image pairs respectively.
Figure B.1: EMPCA-MI Traces for monomodal T1/T1, T1+α20/T1, T1+α40/T1, T1+β/T1 and T1+α40+β/T1 image pairs for x, y translation and θ rotation.
Figure B.2: EMPCA-MI Traces for monomodal $\frac{T2}{T2}$, $\frac{T2+\alpha_{40}}{T2}$, $\frac{T2+\alpha_{40}}{T2}$, $\frac{T2+\alpha_{40}}{T2}$, $\frac{T2+\beta}{T2}$ and $\frac{T2+\alpha_{40}+\beta}{T2}$ image pairs for $x$, $y$ translation and $\theta$ rotation.
Figure B.3: EMPCA-MI Traces for monomodal PD/PD, PD+α20/PD, PD+α40/PD, PD+β/PD and PD+α40+β/PD image pairs for x, y translation and θ rotation.
Figure B.4: EMPCA-MI Traces for multimodal $T1/T2$, $T1+\alpha_{20}/T2$, $T1+\alpha_{40}/T2$, $T1+\beta/T2$ and $T1+\alpha_{40}+\beta/T2$ image pairs for $x$, $y$ translation and $\theta$ rotation.
Figure B.5: EMPCA-MI Traces for multimodal T2/PD, T2+α20/PD, T2+α40/PD, T2+β/PD and T2+α40+β/PD image pairs for x, y translation and θ rotation.
Figure B.6: EMPCA-MI Traces for multimodal PD/T1, PD+α₂₀/T₁, PD+α₄₀/T₁, PD+β/T₁ and PD+α₄₀+β/T₁ image pairs for x, y translation and θ rotation.
Appendix C

EMPCA-MI Results for Neighbourhood Radius

This Appendix includes the detailed $x$ axis translation, $y$ axis translation and $\theta$ rotation plots of EMPCA-MI traces as mentioned in Section 4.3.2. These plots show the results for value of neighbourhood radius $r$ varied between 1, 2, 3, 5, 7 and 10, (keeping the interpolation scheme as bicubic) for mono and multimodal T1, T2 and PD image pairs from BrainWeb dataset having five INU and noise levels (No INU and noise, $\alpha_{20}$, $\alpha_{40}$, $\beta$ and $\alpha_{40}+\beta$) as discussed in Section 3.5.2.

Figures C.1, C.2 and C.3 show the EMPCA-MI traces for monomodal T1/T1, T2/T2 and PD/PD image pairs respectively. Similarly, Figures C.4, C.5 and C.6 show the EMPCA-MI traces for multimodal T1/T2, T2/PD and PD/T1 image pairs respectively.
Appendix C: EMPCA-MI Results for Neighbourhood Radius

Figure C.1: EMPCA-MI Traces for monomodal T1/T1, T1+α20/T1, T1+α40/T1, T1+β/T1 and T1+α40+β/T1 image pairs for x, y translation and θ rotation.
Appendix C: EMPCA-MI Results for Neighbourhood Radius

![EMPCA-MI Results for Neighbourhood Radius](image)

**Figure C.2:** EMPCA-MI Traces for monomodal $T_2/T_2$, $T_2+\alpha_{20}/T_2$, $T_2+\alpha_{40}/T_2$, $T_2+\beta/T_2$ and $T_2+\alpha_{40}+\beta/T_2$ image pairs for $x$, $y$ translation and $\theta$ rotation.
Figure C.3: EMPCA-MI Traces for monomodal PD/PD, PD+α20/PD, PD+α40/PD, PD+β/PD and PD+α40+β/PD image pairs for x, y translation and θ rotation.
Appendix C: EMPCA-MI Results for Neighbourhood Radius

EMPCA-MI Traces for multimodal $T_1/T_2$, $T_1+\alpha_{20}/T_2$, $T_1+\alpha_{40}/T_2$, $T_1+\beta/T_2$ and $T_1+\alpha_{40}+\beta/T_2$ image pairs for $x$, $y$ translation and $\theta$ rotation.

Figure C.4: EMPCA-MI Traces for multimodal $T_1/T_2$, $T_1+\alpha_{20}/T_2$, $T_1+\alpha_{40}/T_2$, $T_1+\beta/T_2$ and $T_1+\alpha_{40}+\beta/T_2$ image pairs for $x$, $y$ translation and $\theta$ rotation.
Figure C.5: EMPCA-MI Traces for multimodal T2/PD, T2+/β, and T2+/α image pairs for x, y translation and θ rotation.
Figure C.6: EMPCA-MI Traces for multimodal PD/T1, PD+$\alpha_{20}/T1$, PD+$\alpha_{40}/T1$, PD+$\beta/T1$ and PD+$\alpha_{40}+\beta/T1$ image pairs for $x$, $y$ translation and $\theta$ rotation.
Appendix D

Supplementary Results for

Chapter 5

This Appendix includes supplementary results for Chapter 5. Figures D.1 (a) and (b) show the RE versus iterations plots for MI, GMI, RMI and EMPCA-MI using 18\textsuperscript{th} initial misregistration scenario for the monomodal registration of T2/T2 and PD/PD image pairs. Figures D.2 and D.3 display the corresponding qualitative results for T2+α_{40}+β/T2 and PD+α_{40}+β/PD respectively. Next, Figure D.4 shows the RE versus iterations plots for MI, GMI, RMI and EMPCA-MI using 15\textsuperscript{th} initial misregistration scenario for multimodal registration of T1/T2, T2/PD and PD/T1 pairs with different INU levels and noise. They are followed by Figures D.5 and D.6 which display the corresponding qualitative results for T2+α_{40}+β/PD and PD+α_{40}+β/T1 respectively.
Finally, Figures D.7 and D.8 show the RE summary boxplot of mono (T2/T2 and PD/PD) and multimodal (T2/PD and PD/T1) pairs respectively using MI, GMI, RMI and EMPCA-MI for all 20 initial misregistration scenarios (having five INU and noise levels).

![Graph showing RE versus iterations plots for MI, GMI, RMI and EMPCA-MI using 18th initial misregistration scenario for monomodal registration of (a) T2 + α40 + β/T2 and (b) PD + α40 + β/PD pairs.](image)

**Figure D.1:** RE versus iterations plots for MI, GMI, RMI and EMPCA-MI using 18th initial misregistration scenario for monomodal registration of (a) T2 + α40 + β/T2 and (b) PD + α40 + β/PD pairs.
Figure D.2: Qualitative results for T2+α40+β/T2 with inset showing zoomed-in central lobes for 18th initial misregistration scenario.
Figure D.3: Qualitative results for PD+α40+β/PD with inset showing zoomed-in central lobes for 18th initial misregistration scenario.
Figure D.4: RE versus iterations plots for MI, GMI, RMI and EMPCA-MI using 15th initial misregistration scenario for multimodal registration of (a) T1/T2, (b) T1+α20/T2, (c) T1+α40/T2 and (d) T1+β/T2 (e) T2+α40+β/PD and (f) PD+α40+β/T1 pairs.
Figure D.5: Qualitative results for $T_2+\alpha_{40}+\beta/PD$ with inset showing zoomed-in central lobes for 15th initial misregistration scenario.
Figure D.6: Qualitative results for PD+α_{40}+β/T1 with inset showing zoomed-in central lobes for 15th initial misregistration scenario.
Appendix D: Supplementary Results for Chapter 5

Figure D.7: RE summary boxplot of monomodal (a) T2/T2 and (b) PD/PD pairs using MI, GMI, RMI and EMPCA-MI for all 20 initial misregistration scenarios. The mean and outliers are denoted by ♦ and + shapes. The numbers on top of the plots refer to the number of out-of-range outliers with RE higher than 65 and 85 pixels respectively.
Figure D.8: RE summary boxplot of multimodal (a) T2/PD and (b) PD/T1 pairs using MI, GMI, RMI and EMPCA-MI for all 20 initial misregistration scenarios. The mean and outliers are denoted by ♦ and + shapes. The numbers on top of the plots refer to the number of out-of-range outliers with RE higher than 60 and 95 pixels respectively.
Appendix E

Supplementary Results for

*Chapter 6*

This Appendix includes supplementary results for *Chapter 6*. Figures E.1 and E.2 show the RE summary boxplot of mono (T2/T2 and PD/PD) and multimodal (T2/PD and PD/T1) pairs respectively using the 1st order (8-pixel), 2nd order (8-pixel) and 2nd order (4-pixel) neighbourhood region connectivity in EMPCA-MI framework for all 20 initial misregistration scenarios (having five INU and noise levels).
Appendix E: Supplementary Results for Chapter 6

Figure E.1: RE summary boxplot of monomodal (a) T2/T2 and (b) PD/PD pairs using 1\textsuperscript{st} order (8-pixel), 2\textsuperscript{nd} order (8-pixel) and 2\textsuperscript{nd} order (4-pixel) neighbourhood region connectivity in EMPCA-MI framework for all 20 initial misregistration scenarios. The mean and outliers are denoted by ♦ and + shapes. The numbers on top of the plots refer to the number of out-of-range outliers with RE higher than 35 and 15 pixels respectively.
**Figure E.2:** RE summary boxplot of multimodal (a) T2/PD and (b) PD/T1 pairs using 1st order (8-pixel), 2nd order (8-pixel) and 2nd order (4-pixel) neighbourhood region connectivity in EMPCA-MI framework for all 20 initial misregistration scenarios. The *mean* and *outliers* are denoted by ♦ and + shapes. The numbers on top of the plots refer to the number of *out-of-range* outliers with RE higher than 21 and 16 pixels respectively.
Appendix F

Supplementary Results for

Chapter 7

This Appendix includes supplementary results for Chapter 7. Figures F.1 and F.2 show the RE summary boxplot of mono (T2/T2 and PD/PD) and multimodal (T2/PD and PD/T1) pairs respectively using the EMPCA-MI \{1\} and adaptive EMPCA-MI \{P\} for all 20 initial misregistration scenarios (having five INU and noise levels).
Appendix F: Supplementary Results for Chapter 7

Figure F.1: RE summary boxplot of monomodal (a) T2/T2 and (b) PD/PD pairs using the EMPCA-MI $\{I\}$ and adaptive EMPCA-MI $\{P\}$ for all 20 initial misregistration scenarios. The mean and outliers are denoted by ♦ and + shapes. The numbers on top of the plots refer to the number of out-of-range outliers with RE higher than 11 and 15 pixels respectively.
Figure F.2: RE summary boxplot of multimodal (a) T2/PD and (b) PD/T1 pairs using the EMPCA-MI \{1\} and adaptive EMPCA-MI \{P\} for all 20 initial misregistration scenarios. The mean and outliers are denoted by ♦ and + shapes. The numbers on top of the plots refer to the number of out-of-range outliers with RE higher than 17 and 16 pixels respectively.