Non-linear registration of medical images

Hava Lester

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

of the

University of London.

Department of Computer Science
University College London

1999
Abstract

This thesis provides a systematic analysis of registration algorithms for application to medical images. We divide our survey into four parts, each of which concentrates on a particular aspect of the algorithms. Linear methods are reviewed in respect to their selection of corresponding features, and their application to the inter-modality case. Non-linear methods are extensively analysed to understand the transition from a conceptual physical or statistical model to the mathematical model and its implementation.

A chapter is devoted to hierarchical methods and their application to solving the local minima problem. Constructions of hierarchies are grouped as temporal variations in data complexity, in warp complexity and in model complexity. These divisions are paralleled in the classification of inhomogeneous methods, where the application of an algorithm varies spatially within the image. Thus we identify variances in relevance of data, in deformability and in chosen model type. In respect of these divisions, we have introduced a nomenclature to describe the restriction or otherwise of the deformation of selected regions in the image. We distinguish between passive and actively-deforming regions, between strongly and weakly deformable regions, and describe two specialisations of rigid regions, namely those which are motionless and those which are independently moving.

The second main contribution of this work is in presenting three inhomogeneous variants to the viscous fluid registration algorithm, one for each of the three classes of inhomogeneity an algorithm may exhibit. In particular one of the variants exhibits a varying viscosity over the image. They are all tested for their ability to restrict the deformation of a specified region independently of the information contained within it.

Finally we evaluate a selection of non-linear registration algorithms using both global and local registration metrics in a variety of tests. The dissertation concludes with three interesting suggestions for future projects.
## Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preface</td>
<td>20</td>
</tr>
<tr>
<td>Abbreviations</td>
<td>21</td>
</tr>
<tr>
<td>Notation</td>
<td>22</td>
</tr>
<tr>
<td>1 Introduction</td>
<td>23</td>
</tr>
<tr>
<td>1.1 Overview of General Image Registration</td>
<td>28</td>
</tr>
<tr>
<td>1.2 Medical Image Data</td>
<td>29</td>
</tr>
<tr>
<td>1.3 Medical Image Processing</td>
<td>34</td>
</tr>
<tr>
<td>1.4 Definitions, Statement of the Problem, Aims and Scope</td>
<td>36</td>
</tr>
<tr>
<td>1.4.1 Definitions</td>
<td>36</td>
</tr>
<tr>
<td>1.4.2 Statement of the Problem</td>
<td>38</td>
</tr>
<tr>
<td>1.4.3 Aims</td>
<td>39</td>
</tr>
<tr>
<td>1.4.4 Scope of this work</td>
<td>39</td>
</tr>
<tr>
<td>1.4.5 Original contributions</td>
<td>40</td>
</tr>
<tr>
<td>1.5 Structure of the Thesis</td>
<td>40</td>
</tr>
<tr>
<td>2 Linear Registration Methods</td>
<td>42</td>
</tr>
<tr>
<td>2.1 Introduction</td>
<td>43</td>
</tr>
<tr>
<td>2.2 Linear transformations</td>
<td>45</td>
</tr>
<tr>
<td>2.3 Principal Axes</td>
<td>47</td>
</tr>
<tr>
<td>2.4 Feature Matching Methods</td>
<td>51</td>
</tr>
<tr>
<td>2.4.1 Landmark-based</td>
<td>51</td>
</tr>
</tbody>
</table>
2.4.2 Surface matching ................................................................. 51
2.4.3 Feature matching based on crest-lines ........................................ 52
2.4.4 Feature matching based on involutions ........................................ 54
2.4.5 Discussion ............................................................................ 54

2.5 Pixel Matching Methods ............................................................. 55
2.5.1 Pair-wise pixel intensity measures ........................................... 55
2.5.2 Cross-correlation variants for adaptation to the inter-modality problem ........................................ 58
2.5.3 Pre-processing ........................................................................ 59
2.5.4 Grouping during measure ....................................................... 60
2.5.5 Mutual Information ................................................................. 60

2.6 Directed search based on optical flow ........................................... 61

2.7 Piecewise-linear methods .......................................................... 63

2.8 Summary .................................................................................. 63

3 Single Level Non-linear Registration Methods ................................. 64

3.1 Introduction ............................................................................... 65

3.2 Representations of deformations ................................................... 66
3.2.1 Deformation field vs. basis function expansions .......................... 66
3.2.2 Landmarks with interpolating functions ...................................... 67

3.3 Pixel/voxel intensity interpolation and resampling ........................... 67
3.3.1 Common interpolation models for pixel-level / micro interpolation .................................. 68
3.3.2 Interpolation for non-linear deformations .................................... 72

3.4 Registration as an optimisation problem ........................................ 76
3.4.1 Regularisation ........................................................................ 77
3.4.2 Physical Models ....................................................................... 77
3.4.3 Statistical Models and the Bayesian Approach ................................ 78
3.4.4 Solution of the optimisation problem .......................................... 82

3.5 Model One: Landmark matching with spline interpolation ................ 83
3.5.1 The Thin-Plate Spline ............................................................... 83
### Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.5.2 Extensions to the Thin-Plate Spline</td>
<td>90</td>
</tr>
<tr>
<td>3.6 Model Two: SPM and other truncated basis function expansion methods</td>
<td>95</td>
</tr>
<tr>
<td>3.6.1 Solution using spectral decomposition</td>
<td>96</td>
</tr>
<tr>
<td>3.6.2 Derivation of basis functions for bending energy smoothness constraints</td>
<td>100</td>
</tr>
<tr>
<td>3.6.3 Other basis functions</td>
<td>102</td>
</tr>
<tr>
<td>3.7 Model Three: Elastic Registration</td>
<td>102</td>
</tr>
<tr>
<td>3.7.1 Historical development of elastic registration</td>
<td>105</td>
</tr>
<tr>
<td>3.8 Model Four: Fluid Registration</td>
<td>107</td>
</tr>
<tr>
<td>3.8.1 The fluid model</td>
<td>107</td>
</tr>
<tr>
<td>3.8.2 Derivation of the fluid PDE</td>
<td>109</td>
</tr>
<tr>
<td>3.9 Summary</td>
<td>110</td>
</tr>
<tr>
<td>4 Hierarchical Methods</td>
<td>111</td>
</tr>
<tr>
<td>4.1 Introduction</td>
<td>112</td>
</tr>
<tr>
<td>4.2 Increasing Data complexity</td>
<td>116</td>
</tr>
<tr>
<td>4.3 Increasing Warp Complexity</td>
<td>125</td>
</tr>
<tr>
<td>4.3.1 Warps summarised by function coefficients</td>
<td>125</td>
</tr>
<tr>
<td>4.3.2 Warps summarised by displacements of nodes or landmarks</td>
<td>126</td>
</tr>
<tr>
<td>4.4 Increasing Model Complexity</td>
<td>127</td>
</tr>
<tr>
<td>4.4.1 Restricting repeated resampling</td>
<td>127</td>
</tr>
<tr>
<td>4.5 Summary</td>
<td>128</td>
</tr>
<tr>
<td>5 Evaluations of Nonlinear and Hierarchical Methods</td>
<td>129</td>
</tr>
<tr>
<td>5.1 Introduction</td>
<td>130</td>
</tr>
<tr>
<td>5.2 Global registration metrics: comparison of the 2-D TPS, SPM and fluid registration</td>
<td>130</td>
</tr>
<tr>
<td>5.2.1 TPS with automatic landmark extraction</td>
<td>131</td>
</tr>
<tr>
<td>5.2.2 Manual landmark extraction</td>
<td>133</td>
</tr>
<tr>
<td>5.2.3 Quantitative comparisons of the registration methods</td>
<td>133</td>
</tr>
<tr>
<td>5.3 Local and global registration metrics: comparison of scale-space methods</td>
<td>139</td>
</tr>
<tr>
<td>5.3.1 Fluid without scale space</td>
<td>139</td>
</tr>
</tbody>
</table>
6 Inhomogeneous Registration

6.1 Introduction

6.2 A classification of spatial inhomogeneities in registration algorithms

6.2.1 Variable data influence

6.2.2 Variable deformability

6.2.3 Variable model type

6.3 Review of current inhomogeneous registration algorithms

6.3.1 Three-component finite element model

6.3.2 Incorporation of rigid structures in the multiquadric (MTQ) spline

6.3.3 Elastic registration with variable elasticity

6.4 Modifications to the constant viscosity fluid algorithm

6.4.1 MF1: Inhomogeneous filter applied to the force field

6.4.2 MF2: Extra boundary conditions in the solution of the velocity fields

6.5 MF3: The inhomogeneous fluid model - fluid with spatially varying viscosity

6.5.1 Introduction

6.5.2 Theory

7 Evaluations of Inhomogeneous Registration Algorithms

7.1 Introduction

7.2 Local metrics of rigidity

7.3 2D Results

7.3.1 Synthetic labelled images

7.3.2 Pre/post-operative head images, arbitrary weakly-deformable body

7.3.3 Artificially-warped knee images

7.3.4 Pre/post-operative head images, rigid scalp

7.4 3D Results
## List of Tables

5.1 Description of location of the manually-selected landmarks for insertion onto a thin-plate spline applied to registering head3A to head3B. A: anterior, P: posterior, L: left, R: right.  133

5.2 Global measurements of registration accuracy between the target and the image transformed by the registration method described in the left-hand column: sum of squared differences, correlation coefficient (cc) and mutual information (mi). For comparison, metrics for the comparison of the target to itself are given in the bottom row. Good registration gives high values of mutual information, and values close to 1 and 0 for the correlation coefficient and sum of squared differences respectively.  137

7.1 Bending energies of the deformation of the door grid points  173

7.2 Maximum values of deformation metrics for the constant and variable viscosity cases of the knees test.  179

7.3 Coordinates of known source and target landmarks, with final locations of positions of the source landmarks after registration using both the regular and varying-viscosity fluid algorithms.  180

7.4 Distance metrics applied to boundaries of bone regions.  183

7.5 Parameters of image acquisition for the neck images.  190

7.6 Correlation coefficients results of registering neckD to neckI (top row) and neckO to neckI (bottom row).  193

B.1 Increase in the correlation coefficient metric during progress of fluid without scale space for the test in section 5.3 for the images shown in figures 5.10 and 5.11. Looping terminated because the change in correlation coefficient was $-0.000898$, less than the termination threshold of 0.0001  227
List of Tables

B.2 The coordinates of the 10 boundary and 10 internal landmarks, shown in figure 5.12, selected for testing the local registration quality of the fluid model without scale space. The known target landmarks within head1A were inserted into a thin-plate spline together with the landmarks listed here as ‘source’ to produce the deformed image head1B. Head1B was then registered back to head1A using the fluid model in closure scale space. The resulting deformation when applied to the source landmarks produced small regions whose centres of mass are listed here as ‘f(source)’. The columns show alternately x- and y-coordinates. ..................................................................................................................................228

B.3 The increase in the correlation coefficient (cc) metric during the progress of the fluid within closure scale space. The image numbers in bold font are those shown in figure 5.14. There are many iterations in the first scale level partly because the threshold of maximum number of SOR iterations was set at 50 so the PDE was not fully solved initially; as the fluid progressed, the computed velocity field became closer to the true velocity field for a given iteration. The second reason is that only fine adjustments (small displacements) were needed for the higher resolution levels since the images were reasonably well-registered by that stage. .................................................................................................................. 229

B.4 The coordinates of the 10 boundary and 10 internal landmarks, shown in figure 5.15, selected for testing the local registration quality of the fluid model within closure scale space.230

B.5 The coordinates of the 10 boundary and 10 internal landmarks, shown in figure 5.21, selected for testing the local registration quality of the fluid model within Gaussian scale space. ................................................................................................................................................. 231

B.6 Increase in correlation coefficient during progress of fluid within Gaussian scale space (scale levels 9 to 5). Image numbers in bold are those shown in figure 5.10. .......................... 232

B.7 Increase in correlation coefficient during progress of fluid within Gaussian scale space (scale levels 4 to zero). Image numbers in bold are those shown in figure 5.11. ........................ 233
List of Figures

1.1 How do the images differ? .................................................. 23
1.2 Subtraction of one image from another giving a difference image. ......................................... 24
1.3 Subtraction of misaligned images. ................................................ 24
1.4 Image painted on a hypothetical rubber sheet which is then stretched. ................................. 25
1.5 Local minima (marked with thin arrows) and the global minimum (thick arrow) of a one-dimensional surface. ................................................................. 27
1.6 X-ray CT (left) projector and arc of detectors; (right) helical data collection. ....................... 30
1.7 X-ray CT: (left) strengths of attenuated X-rays are measured after passing through the subject. (right) back projection during estimation of tissue density at each voxel location. 30
1.8 Precession of a paramagnetic isotope. .................................................. 31
1.9 Magnetic field induced by coil of electromagnet. ................................................................. 32
1.10 T1 relaxation. A: Axis of nuclear spin aligned with magnetic field; B: after receiving RF pulse, axis at 90° to magnetic field; C,D: nucleus loses energy by emitting electromagnetic radiation and gradually re-aligns with magnetic field. .................................................. 32
1.11 T2 relaxation. A: After receiving RF pulse, the collection of nuclei in the region exhibit a net magnetic field perpendicular to that of the scanner magnet. B: Net perpendicular magnetic field becomes zero due to variable precession rates of the nuclei. .................. 33
1.12 The medical image processing pipeline .................................................. 34
2.1 Principal axes registration of a source (S) and target (T) image pair. ................................. 48
2.2 Rotation of image about its centre after translation to the origin. ........................................... 50
2.3 A 1-dimensional curve multi-valued at $x = x_1$. ................................................................. 52
2.4 Principal directions $\mathbf{e}_M$ and $\mathbf{e}_m$ lying in the tangent plane. ........................................... 53
2.5 Convex hull of a surface. ................................................................. 54
List of Figures

2.6  Cortical involution, with convex hull as dotted line, showing depth and entry/exit angles to convex hull ................................................................. 54

3.1  Nearest neighbour resampling at a boundary region. .............................................. 69

3.2  Linear interpolation function ...................................................................................... 69

3.3  1D bicubic spline function ............................................................................................. 70

3.4  Sinc function (left); pulse or rect function (right) ............................................................. 70

3.5  Fourier transform of truncated sinc, an approximation to the pulse/rect. ......................... 71

3.6  Appodised sinc .................................................................................................................. 71

3.7  First 3 convolutions of rectangle, giving the nearest-neighbour, linear, quadratic and cubic interpolation functions. Circles indicate pixel locations. .................... 72

3.8  'Backwards' interpolation. (right) The transformed source image is constructed by scanning the output lattice and mapping each pixel to the position from which it originated in the pre-transformed source. This is generally an interstitial location (centre). Its intensity value is obtained by a weighted sum of the values of neighbouring lattice pixels in the undeformed source (left). The weights and choice of neighbours is determined by the interpolation model. ......................................................... 73

3.9  Sinc interpolation after a non-linear deformation ............................................................. 75

3.10 Derivation of the mathematical model ............................................................................. 76

3.11 Prior, likelihood and MAP parameter estimates ............................................................... 81

3.12 distances between point $P_i$ and other points in the set ................................................. 86

3.13 Vertical displacements upwards of landmarks originally at $(0, \pm 1)$ in the plane $z = 0$ and downwards of those originally at $(\pm 1, 0)$ ........................................ 88

3.14 Radial symmetry of basis functions $U(r)$ ...................................................................... 92

3.15 Circular paraboloid ($\alpha = 2$) and circular hyperboloid ($\alpha = 1$). ................................. 93

3.16 Effect of altering the parameters $C$ and $w$ in the quadrics ........................................... 93

3.17 The Gaussian basis functions .......................................................................................... 94

3.18 Variants in the derivation and solution of the mathematical model for elastic registration 104

4.1  Data set of (left) head1A, MRI head scan and (right) head1B, a warped version of head1A 112

4.2  Data set of MRI head scans of two different subjects, (left) head2A, (right) head2B .......... 113
List of Figures

4.3 Misregistration of brain to scalp in the fluid deformation of head1B to head1A, showing growth of new 'scalp' from existing vertical boundary in top of scalp. .................. 113

4.4 Subtraction images of the first, second and third deformed images in figure 4.3, with (left) source image, showing the cause of misregistration: the overlap of different structures with similar contextual contrast. ......................... 114

4.5 The first 8 timesteps in the fluid registration of test image head2B to head2A showing misregistration of scalp to cortex. ................................. 114

4.6 Difference image of head2A and head2B showing overlapping cortex-skull and scalp-background boundaries, the source of the confusion causing misregistration. .... 115

4.7 Scale space stack generated from an input image. ................................. 117

4.8 Slices of the isotropic Gaussian scale space of head1A, at scales of (left–right) $\sigma = 2, 6, 15$ and 20 pixel units. .......................... 117

4.9 Application of anisotropic blur to head1A with conductance given by equation (4.2); (left–right): timesteps $t = 1, 4, 7$ and 10. ......................... 119

4.10 Application of anisotropic blur to head1A with conductance given by equation (4.3); (left–right): timesteps $t = 1, 4, 7$ and 10. ......................... 120

4.11 Application of inhomogeneous blur with conductance given by equation (4.3); timestep 20 with an edge region marked by a box .................. 121

4.12 Continued application of inhomogeneous blur (4.3) in the box region marked in figure 4.11, timesteps 20–24, showing leakage across edges. ........ 121

4.13 Closure on test image head1A. (left) using a structuring element of radius 20 identifies two objects: scalp and ‘inside-scalp’. (right) a structuring element of radius 5 is less successful at identifying 3 objects: scalp, skull and brain. False structures shaped by the structuring element are evident. .............................. 122

4.14 Fluid registration of head1B to head1A in the first level of the Gaussian pyramid. .......................... 123

5.1 Data set of (left) head3A and (right) head3B ........................................ 131

5.2 Propagation of initial landmarks by the fluid algorithm. The top row illustrates the source image deforming over time; the bottom row shows the concurrent deformation of an image of a bright dot at a landmark location. ......................... 132

5.3 Diagram showing locations of manually selected landmarks in head3 images. Landmark number-labels are those given in the first column of table 5.1. ......................... 134
5.4 (left) Landmarks automatically selected in the source image (head3A); (right) new positions of landmarks after fluid registration of source to target; points have spread to small regions.

5.5 Head3A registered to head3B using: (top, left to right) linear; TPS (18 manually-selected landmarks); TPS (170 landmarks derived from fluid); (bottom, left-right): SPM (4 basis fns); SPM (8 basis fns); SPM (12 basis fns); fluid registration; target.

5.6 Differences between head3B and deformed head3A with registration (top left to right): none; linear; TPS (18 manually selected landmarks); TPS (170 landmarks derived from fluid); (bottom, left-right): SPM (4 basis functions); SPM (8 basis functions); SPM (12 basis functions); fluid registration.

5.7 Grid image (left) deformed during fluid registration of head3A to head3B and (centre) warped by the thin-plate spline using the 170 derived landmarks. The fluid deformations are localised at each image point whereas the spline smoothly interpolates the deformations between landmarks. (right) grid deformed by the manual TPS.

5.8 Plot of metrics given in table 5.2. The correlation coefficient is plotted to the 8\textsuperscript{th} power and the mutual information of target/target has been reduced to 2 for a clearer plot. For good registration, the sum of squared differences is reduced, whereas the correlation coefficient and mutual information are increased.

5.9 Diagram showing locations of original (known) target landmarks in test image head1A: (left) 'boundary' and (right) 'inner'. Original source landmarks are in the equivalent topological locations in head1B.

5.10 Progress of the fluid registration of head1B to head1A without scale space, showing iterations 15, 30, 45; left-right: Top row shows the deforming source, the lower row shows the same deformation applied to a regular grid.

5.11 Progress of the fluid registration of head1B to head1A without scale space, showing the 60th, 75th and final (83rd) iterations; left-right: Top row shows the deforming blurred source, the lower row shows the same deformation applied to a regular grid.

5.12 (left-right) Locations of original source landmarks; regions of deformed source landmarks at termination of the fluid registration without scale space; locations of centres of mass of these regions; locations of known target landmarks.

5.13 Closure with a disc structuring element of radii (left-right) 20, 4, 3 pixels applied to the 256 × 256 images head1A (top row) and head1B (bottom row).
5.14 Progress of the fluid registration of head1B to head1A at the end of each level in closure scale space; left–right: end of scale 3 (radius 20), 88th iteration; end of scale 2 (radius 4), 117th iteration; end of scale 1 (radius 3), 121st iteration; end of scale 0 (no closure), 136th iteration. Top row shows the deforming source in closure scale space, the lower rows show the same deformation applied to a regular grid and to the original-resolution source image. ................................................................. 143

5.15 (left–right) Locations of original source landmarks; regions of deformed source landmarks at termination of fluid in closure scale space; locations of centres of mass of these regions; locations of original target landmarks. ................................................................. 144

5.16 Gaussian blurs of spatial standard deviation (left–right) 18, 16, 14, 12 applied to the 256 x 256 images head1A (top row) and head1B (bottom row). ................................................................. 145

5.17 Gaussian blurs of spatial standard deviation (left–right) 18, 16, 14, 12 applied to the 256 x 256 images head1A (top row) and head1B (bottom row). ................................................................. 146

5.18 Progress of the fluid registration of head1B to head1A at the end of the first four levels in gaussian scale space; left–right: Top row shows the deforming blurred source, the lower rows show the same deformation applied to a regular grid and to the original-resolution source image. ................................................................. 147

5.19 Progress of the fluid registration of head1B to head1A at the end of levels 5 to 2 in gaussian scale space; left–right: Top row shows the deforming blurred source, the lower rows show the same deformation applied to a regular grid and to the original-resolution source image. ................................................................. 148

5.20 Progress of the fluid registration of head1B to head1A at the end of the final two levels in gaussian scale space; left–right: Top row shows the deforming blurred source, the lower rows show the same deformation applied to a regular grid and to the original-resolution source image. ................................................................. 149

5.21 (left–right) Locations of original source landmarks; regions of deformed source landmarks at termination of fluid in gaussian scale space; locations of centres of mass of these regions; locations of original target landmarks. ................................................................. 149

5.22 Comparison of progress of the fluid registration of head1B to head1A in closure and Gaussian scale spaces and without scale space. Correlation coefficients are between the blurred target and deforming source images; at each new scale level the correlation coefficient drops as more detail is re-introduced into the images, which needs matching. Within one scale level (or in the no-scale-space case) the correlation coefficient is monotonically increasing to a horizontal asymptote (approximates a negative exponential). ................................................................. 150
5.23 Plots of landmarks listed in tables B.2, B.4 and B.5: starting locations in the source image head1B, true target locations in head1A, and locations in deformed source produced by fluid registration without scale space and in Gaussian and closure scale spaces. 151

6.1 Multiquadric deformation of knee image incorporating two rigid bodies (bones) 158

6.2 Effect of the distance functions on the combined warp 160

6.3 First three stages in Davatzikos' registration scheme. A: deformable surface is fitted to each segmented cortical volume. B: parametrisation of deformable surfaces. C: deformable surface pairs are matched by their curvatures. 161

6.4 Weighting function to apply to the EDT, here with a = 2, b = 10 163

6.5 Application of 4SED to windows in house image 163

7.1 Artificial images: (1-r) clown, house, clown2, house2 168

7.2 Constant viscosity, house deformed to clown. 170

7.3 Constant viscosity, house2 deformed to clown. 170

7.4 MF3 (Varying viscosity) house2 deformed to clown, with weakly-deformable door 171

7.5 MF3 (Varying viscosity) house2 deformed to clown2, with weakly-deformable door 171

7.6 MF1, house2 deformed to clown2, with door defined as a passive region 171

7.7 MF2, house2 deformed to clown2 with door region defined as motionless 172

7.8 Regular fluid (constant viscosity), house deformed to clown: deformed house subtracted from clown showing areas where registration was incomplete. 172

7.9 Subtraction images from deformed house2 images from (top) clown/clown2 and (bottom) original house2; (1-r) regular fluid, MF3 (weakly-deformable door; registered to clown), MF3 (weakly-deformable door; registered to clown2), MF1 (passive door), MF2 (motionless door) 172

7.10 The deformed door regions of the final grid images of each of figures (1-r) 7.3, 7.4, 7.5, 7.6 and 7.7 173

7.11 Data set of (left) head4A and (right) head4B 174

7.12 Source head4B with arbitrary weakly-deformable region highlighted and outlined 174

7.13 Deformed source images after 40 timesteps of fluid registration to the target using the (left) constant and (right) variable viscosity models 175
7.14 Magnitude of displacements: (left) constant viscosity model; (right) variable viscosity. High values are coloured black. ................................. 175

7.15 Grid images showing the deformation of the arbitrary weakly-deformable region; (centre) constant viscosity model, (right) variable viscosity model. ................................. 175

7.16 The elastic deformation metric: (left) constant viscosity, (right) varying viscosity (high values coloured white). Both clearly show regions of high area growth, indicating locations of abnormality in the source image. ................................. 176

7.17 The Laplacian bending metric: (left) constant viscosity, (right) varying viscosity. The weakly-deformable region shows little to no bending in the varying viscosity model, in contrast to the constant viscosity model. ................................. 176

7.18 The thin-plate bending metric: (left) constant viscosity, (right) varying viscosity, showing more clearly than the Laplacian the difference in bending between the constant and variable viscosity models. ................................. 177

7.19 Source knee image set within a margin. ................................. 178

7.20 Locations of the six landmarks and boundaries of three rigid bodies (labelled) in the source knee image. ................................. 179

7.21 Scale space for source (top) and target (bottom) knee images, consisting of disc structuring elements of radius (left) 20, (centre) 4 and (right) 3 pixels ................................. 180

7.22 Registration of knee image to its warped version, using (top) constant viscosity, (bottom) variable viscosity fluid models: (l-r) deformed source; corresponding grid; difference image of deformed source from target ................................. 181

7.23 Metrics of deformation of kneeB to kneeA, (l-r) Laplacian; elastic; bending; transformation, for (top) the constant viscosity and (bottom) variable viscosity cases. ................................. 182

7.24 Knee images with landmarks and bone boundaries marked: (l-r) source; source deformed using constant-viscosity fluid; source deformed by MF3 with viscosity set at 1.0 within bones and 0.1 elsewhere; target ................................. 182

7.25 Source, known target and deformed source landmarks and boundaries of bone regions for the comparison of the regular fluid and MF3 on the knee test image. ................................. 184

7.26 Constant viscosity fluid, post-op registered to pre-op. (l-r): deformed post-op; deformation applied to grid; difference image - deformed post-op subtracted from target (pre-op) showing areas where registration was incomplete, or remaining differences with the atlas; difference image - deformed post-op subtracted from post-op highlighting differences due to the deformation. ................................. 187
7.27 Constant viscosity fluid, post-op registered to pre-op, summarised/approximated by TPS. 
(l-r) TPS-deformed post-op; TPS deformation applied to grid; TPS-deformed post-op subtracted from pre-op; TPS-deformed post-op subtracted from post-op. 187

7.28 Varying viscosity fluid (MF3), post-op registered to pre-op. (l-r): deformed post-op; deformation applied to grid; deformed post-op subtracted from pre-op; deformed post-op subtracted from post-op. 187

7.29 Varying viscosity fluid, post-op registered to pre-op, summarised/approximated by TPS. 
(l-r) TPS-deformed post-op; TPS deformation applied to grid; TPS-deformed post-op subtracted from pre-op; TPS-deformed post-op subtracted from post-op. 188

7.30 Varying viscosity fluid, post-op registered to pre-op, summarised/approximated by rigid-MTQ. (l-r) rigid-MTQ-deformed post-op; rigid-MTQ deformation applied to grid; rigid-MTQ-deformed post-op subtracted from pre-op; rigid-MTQ-deformed post-op subtracted from post-op. 188

7.31 Boundary conditions (MF2): zero velocities on scalp. (l-r): deformed post-op, deformation shown on grid; deformed post-op subtracted from pre-op; deformed post-op subtracted from post-op. 188

7.32 Constant viscosity fluid, post-op registered to pre-op. (l-r) rigid body and selected markers (features image); fluid-deformed features image subtracted from original; TPS-warped features image subtracted from original; rigid/MTQ-warped features image subtracted from original. 188

7.33 Constant viscosity fluid, post-op registered to pre-op, deformation metrics (l-r): log of the Laplacian, bending, elastic, magnitude of transformation. 189

7.34 Varying viscosity fluid (MF3), post-op registered to pre-op, deformation metrics (l-r): log of the Laplacian, bending, elastic, magnitude of transformation. 189

7.35 Pre-processing stages shown on neckD. Top left: slice 19 of the original 3D 256 \times 256 \times 38 volume; right: correction after division pixelwise by the same image blurred with a Gaussian of spatial standard deviation \( \sigma = 5 \); bottom left: masked with the aid of the automatic contouring and manual alteration tool in xdispunc; right: downsampled by a half. 191

7.36 Segmentation of vertebrae slice-by-slice from neckD. Top left: vertebrae outlined manually using the xdispunc regions tool (slice 32); right: resulting binary image after masking each vertebra and colour processing in xv; bottom left: the same spine slice downsampled by a half; slice 19 of the zero-padded and downsampled 128 \times 128 \times 32 spine volume. 192

7.37 Slice 16 (of the 32-slice downsampled volumes) of the images, left to right: D; fluid D/I; variable-viscous fluid D/I; SPM D/I; I. 194
7.38 Slice 16 (of the 32-slice downsampled volumes) of the difference images, left to right:
D/I; fluid D/I; variable-viscous fluid D/I; SPM D/I ............................................................... 195

7.39 Central slices of full resolution neckD (left) and neckI (right). Upsampled displacement
fields obtained from (centre-left) constant-viscosity and (centre-right) varying-viscosity
fluid registration of downsampled neckD to downsampled neckI. ............................................ 195

7.40 Slice 16 (of the 32-slice downsampled volumes) of the images, left to right: O/I; fluid O/I;
variable-viscous fluid O/I; SPM O/I; I ............................................................................... 195

7.41 Slice 16 (of the 32-slice downsampled volumes) of the difference images, left to right:
O/I; fluid O/I; variable-viscous fluid O/I; SPM O/I ............................................................... 195

7.42 (Top to bottom) NeckD; fluid registration of neckD to neckI; SPM registration of neckD
to neckI; variable-viscosity fluid registration of neckD to neckI; neckI. ............................. 196

7.43 (Top to bottom) NeckO; fluid registration of neckO to neckI; SPM registration of neckO
to neckI; variable-viscosity fluid registration of neckO to neckI; neckI. Rendered voxel
gradient shading, using the Analyze package, with light source at vertical angle 30°, hor­
izontal angle 25° ............................................................................................................................ 197

7.44 (top) spineD segmented from neckD; (upper middle) spineD with regular fluid deforma­
tion of neckD to neckI; (lower middle) spineD with variable-viscosity fluid deformation
of neckD to neckI with vertebrae 'rigid'. For comparison, spineI is shown segmented
from neckI (bottom). The 3D images were generated from sets of 2D slices by voxel gra­
dient shading volume rendering using the Analyze package. ..................................................... 199

7.45 (left-right) central slice of binary spine image segmented from neckD; central slice of de­
formed spine segmented from neckD computed by applying upsampled deformation field
from constant viscosity fluid registration of neckD to neckI; and from variable viscosity
fluid registration; for reference only: binary spine image segmented roughly from neckI
(this segmentation took no part in the fluid registrations) ..................................................... 200

7.46 (top) spineO segmented from neckO; (upper middle) spineO with regular fluid deforma­
tion of neckO to neckI; (lower middle) spineO with variable-viscosity fluid deformation
of neckO to neckI with vertebrae ‘rigid’. For comparison, spineI is shown segmented
from neckI (bottom). The 3D images were generated from sets of 2D slices by voxel gra­
dient shading volume rendering using the Analyze package. ..................................................... 201

7.47 Slice 16 (of the 32-slice downsampled volumes) of the metric images, D to I for (top
row) constant viscosity fluid, left to right: log of Laplacian; log of elasticity energies;
magnitude of transformations. Bottom row shows the equivalent for the varying viscosity
case. .................................................................................................................................................. 202
7.48 Slice 16 (of the 32-slice downsampled volumes) of the metric images, O to I for (top row) constant viscosity fluid, left to right: log of Laplacian; log of elasticity energies; magnitude of transformations. Bottom row shows the equivalent for the varying viscosity case. ................................................................. 202

7.49 (Top to bottom) Neckl; registration of neckl to neckD by Graeme Penney using rigid vertebrae inside 3D multiquadric; neckD. Rendered voxel gradient shading, using the Analyze package, with light source at vertical angle 30°, horizontal angle 25°. ................. 203

7.50 Registration of I to D by Graeme Penney. Shown are (1-r) slices 14, 16, 19 and 32 of the $256 \times 256 \times 64$ full-resolution volume. Slices 14 and 19 show parts of the volume outside of the envelope bounding the control points, which have undergone deformations extrapolating from the registration within the envelope, constrained only by the bending energy of the spline but not constrained by any similarity criteria. .................. 204

7.51 Registration of I to D by variable-viscous fluid following registration by Graeme Penney (fig. 7.49. Shown are (1-r) the central slice (slice 16) of the downsampled registered volume; the same central slice (slice 32) at the highest resolution using the upsampled displacement field; slice 34 of the highest resolution registered source volume; slice 32 of the full-resolution volume D (target). ......................................................... 205

8.1 Summary of review of registration algorithms ................................................................. 212

A.1 chamfer 3-4 and chamfer 5-7-11 distance transform masks ........................................... 214
A.2 Point and volume element in deformed and original target image ................................. 215
A.3 Extensional and shear strain ......................................................................................... 215
A.4 Displacements $\vec{u}$ and decompositions of line elements $d\vec{R}$ and $d\vec{r}$ (following Christensen) 216
A.5 Internal forces $\delta F$ acting on elemental area $\delta A$ ....................................................... 217
A.6 Digital representation of the Jacobi iteration matrix for finite-difference solution of the elasticity PDE ................................................................................. 221
Preface

A host of genuinely interested volunteers must be acknowledged for their many contributions. First and foremost I would like to give my deepest thanks to my supervisor Dr Simon Arridge for his unending enthusiasm, encouragement, patience and advice and without whom this work would neither have begun nor been seen to completion. Many many thanks also go to my mentor and friend, Dr Julia Schnabel, who wrote my first Makefile and \TeX template and who was always available with an abundance of cheery technical support. I would like to thank Dr David Barber and Dr Paul Tofts who have agreed to be my external and internal examiners respectively. Three people at the Institute of Neurology have shown interest in my work throughout the course of the project: John Ashburner spent many hours in discussion and provided loan of disk space to use Analyze and explanations and application of SPM; Louis Lemieux and Professor Robert Turner provided further comments and guidance. Thank you to Dr Kalvis Jansons for his derivation of the variable-viscous fluid PDE and together with Drs Mitch Berger and Robb McDonald (UCL Maths Dept) for their insights into the behaviour of fluids. Thank you also to Angela Oatridge and especially to Professor Joe Hajnal of the Hammersmith Hospital for a long Thursday evening devoted to personally providing three full neck volume images, and to Graeme Penney of UMDS for painstakingly applying their multiquadric registration. Thank you to our PhD tutor Professor Mel Slater and my second supervisor Dr Søren Aksel-Sørensen for their advice and for reading the thesis. I have been rescued many times from the vagaries of elvin my workstation by Systems group and the guys who look after our hardware; their support considerably oiled the progress of my work. EPSRC funded me for three years and the Graduate School and the Department all funded my attendance at conferences. Tina Yu had a big smile and our weekly meetings were great fun; Norman Smith provided advice and knee images and read the thesis, as did Olivier Coulon. Marion Hermes - thank you for patiently listening to all my crazy ideas and being a great flatmate. Todah, merci and danke to my aunt Dr Ruth Morris for her on-line overseas support and encouragement - with love, tn! Todah rabah grandpa for the loan of your 'puter which enabled me to work through the small hours of the night. And finally my gratitude to my parents, the list is too long.
Abbreviations

bc  boundary conditions
CSF  cerebral spinal fluid
CT  computed tomography
DT  distance transform
EBS  elastic body spline
EDT  Euclidean distance transform
FFT  fast Fourier transform
fMRI  functional MRI
MAP  maximum a posteriori
MCMC  Markov chain Monte Carlo
MF1  modified fluid 1 (forces filter)
MF2  modified fluid 2 (extra boundary conditions on velocities)
MF3  modified fluid 3 (varying viscosity)
MI  mutual information
MMSE  minimum mean sum of errors
MR(I)  magnetic resonance (imaging)
MTQ  multiquadric
PA  principal axes
PDE  partial differential equation
PET  proton emission tomography
SOR  successive over-relaxation
SPM  statistical parametric mapping
SSD  sum of squared differences
TPS  thin-plate spline
UNC  University of North Carolina
VOI  volume of interest
Notation

We employ standard mathematical notation. The following list of reserved symbols are employed in the text.

- $r(\vec{x}, \vec{y})$: distance function
- $\mathcal{D}(S,T)$: distance function; similarity measure
- $\mathcal{D}_d(S,T)$: sum of squared differences
- $\mathcal{D}_c(S,T)$: correlation coefficient
- $\mathcal{D}_l(S,T)$: mutual information between source and target
- $R(\vec{u})$: roughness measure or smoothness constraint
- $J(.)$: cost or energy functional
- $\lambda$: regularisation parameter
- $\lambda, \mu$: elasticity or viscosity parameters
- $S(\vec{x})$: source (deforming) image
- $T(\vec{x})$: target (matched) image
- $\Omega_S, \Omega_T, \Omega$: source, target, image domain
- $\Omega_i$: region of image
- $\mathcal{R}$: set of regions in image
- $I(S(\vec{x})), I(T(\vec{x}))$: pixel intensities in source / target
- $\phi(\vec{x})$: contrast or pixel intensity mapping
- $\vec{u}(\vec{x})$: transformation or domain mapping
- $\mathcal{U}$: set of transformations
- $\beta_k$: the $k$th basis function
- $w_i$: the $i$th weight function
- $C$: covariance matrix $= [c_{ij}]$
- $\bar{\Omega}$: inertia tensor $= [\Omega_{ij}]$
Chapter 1

Introduction

The hospital and medical research laboratory of today boasts a sophisticated range of imaging modalities capable of examining internal organs and tissue structures without the need for surgery.

The results of these interrogations are output as images, which are to be analysed by radiologists, specialists and researchers. In their analysis they are presented with a pair of images which are to be compared for significant differences. They are looking for changes in the size of a tumour, or for atrophy of the brain, causing enlargement of brain cavities called ventricles. Or perhaps they are monitoring neural activity in the brain, manifested as very slight increases in brightness or intensity in the image with respect to a base-line image of the brain at rest.

Classically the images are printed on film; the radiologist mounts each onto a viewing box and looks from one to the other making the assessment (figure 1.1).

Figure 1.1: How do the images differ?

But the differences may only be slight, or they may need exact quantification. There may be many such images to process, or the images may be fully 3-dimensional volumes. In all these cases, the radiologist’s task is made easier if the images, stored digitally by the scanners, are subtracted one from the other by computer so that only the differences remain (figure 1.2).

Generally however the anatomy to be compared has not been captured in exactly the same location or
orientation in each image. The patient has undoubtedly shifted even slightly between the first and second scan - especially if the second scan was performed on another day or in a different scanner. Subtraction of one image from the other will then result in plenty of structure in the difference image, representing misalignment rather than changes in anatomy (figure 1.3).

For the difference images to display relevant information only, the image pair must be correctly aligned. The correct alignment is termed registration. The image being repositioned is referred to as the source image, and the image it aims to match is the target image.

Registration consists of two stages:

1. First, correspondences must be found in the image pair. These are locations which are identified in both images as representing the same position anatomically. They may be selected point pairs known as landmarks, or they may be surfaces of anatomical structures; an holistic approach may be taken in which correspondences are to be found between each spatial location; alternatively, derived properties may be computed for each image, such as its geometric centre and orientation. These locations or spatially-dependent properties of the image are termed features.

2. The second stage is to move the source image relative to the target so that the corresponding features are superimposed. This movement is called a transformation.

Sliding or rotating one of the images may be sufficient to achieve alignment. In this case the transformation and registration are both said to be linear.

However, patient movement between the scans may have caused a change in shape of the organs or tissue structures scanned. In this case, to achieve alignment between corresponding structures and their boundaries or outlines, one of the images will have to be distorted or deformed. The distortion itself will not change the intensity of any part of the image, and the relative positions of the structures depicted will be unchanged. Most importantly, the image should not be torn or folded to achieve the registration. We can
think of the image as drawn on a sheet of rubber which is being stretched and pulled out of shape (figure 1.4).

Figure 1.4: Image painted on a hypothetical rubber sheet which is then stretched.

This manipulation of the image is termed a nonlinear deformation or a warp. To instruct the computer in how to deform the image (having supplied descriptions of the corresponding features to be matched), we may provide a model of some deformable physical material whose behaviour under the influence of external forces is described by a mathematical equation. The requirement that each pair of features be superimposed will replace the external forces in the equation, to drive the deformation. Examples of such physical models are elastic materials (which can be stretched) and pliable materials (which can be bent smoothly without introducing sharp creases).

Many registration algorithms, both linear and non-linear, have been developed over the past 20 years. The choice of an algorithm depends on its robustness and additionally on its degree of automation, referring to the amount of intervention required of or invited from the radiologist. A fully automatic algorithm has a minimum of reliance on human operators. Automation is preferable in that the radiologist is free to pursue other tasks requiring his/her expert knowledge - registration is laborious and time consuming when performed manually on large quantities of data. Automatic registration is also preferable in that it is infallible to subjective operator bias - for example, the computer does not get tired at the end of a long day! Nevertheless, some degree of human intervention is beneficial if only to monitor the process to ensure that for the given data a satisfactory result has been achieved - the role of the computer being to assist rather than replace the specialist.

Further criteria for selecting a registration algorithm are application dependent. For example, in section 1.2 we describe types of medical image data that may require registration. They vary in their picture quality and in the tissues represented. A common registration task is to match two images of the same subject but of different data type (acquired by different scanners), so that complementary information from both may be fused. In this case there may be few anatomical structures which appear similar in intensity and contrast in both data types. Identifying correspondences is then a more complex task.

Registration algorithms are tuned to perform best at a particular application. We list here terminology which has evolved to describe the application-specific type of registration.

**intra-subject** the source and target are of the same subject

**inter-subject** the source and target are of different subjects
inter-modality  the source and target are obtained from different scanning modalities

perioperative  the target image is acquired during a surgical procedure and the source is registered to it in real-time

pre-/post-operative  the images are acquired before and after a surgical procedure

cross-population  a set of source images is drawn from a population of data and registered to a common target

time studies  source image(s) of the same subject are obtained over a period of time such as several months

dynamic  source images of the same subject acquired at short time intervals (seconds or minutes)

Examples of perioperative applications are the real-time registration of a video image of the subject in the operating theatre to a pre-operative anatomical scan for tumour location, and/or to a labelled atlas identifying sensitive regions to avoid, and/or to a model with the planned surgical route marked. Pre- and post-operative anatomical scans are compared to evaluate the success of the operation such as the removal of a lesion: the cancerous tissue may be indistinguishable to the naked eye from normal tissue, and may only be visible in an MR scan where the variable tissue density is apparent. Registration is applied in cross-population studies for tasks such as the construction of an averaged ‘normal’ data set as a basis for an atlas, and identification of schizophrenics from normals by comparing sizes of ventricles in the brain. A subject scan may be matched to an atlas for passing labels or segmentations. Often the transformation required is highly non-linear and furthermore a homology* may not exist. Time studies are performed to monitor the course of an abnormality and its treatment - such as growth or shrinkage of tumours, or for the early detection of Alzheimer’s disease by detecting brain atrophy. There are many varied pathologies which introduce abnormal structures into the subject [Weir and Murray, 1998; Du Boulay, 1984; Partridge, 1985; Simpkins, 1988; Grainger and Allison, 1997]. The chest and abdomen contain highly-deformable tissue [Weir and Abrahams, 1997] so even intra-subject cases may be complex. Neither is registration of images of relatively rigid structures such as the brain [Talairach and Tournoux, 1988; Crossman and Neary, 1995; Thompson et al., 1996; Kapur et al., 1996] simple in the inter-subject or subject-atlas case where cross-population variability is high.

An automatic algorithm requires a method for determining whether registration has been successful. To this end, a distance function is computed which evaluates the closeness of the specified corresponding features in the registered image pair. In most algorithms, registration is a gradual process. The source image is repositioned or deformed slightly many times, reducing the value of the distance function at each step, until no further improvements can be made, at which point the algorithm terminates.

The distance function could hypothetically be plotted for each deformation of the source, giving a surface. The minimum of this surface is at the deformation which gives the lowest value of the distance function.

*two images exhibit an homology if for each structure, however small, in one image, the corresponding structure can be identified in the other image, and likewise with the roles of the two images reversed.
However, due to the complex nature of medical image data, there may be many deformations which match structures which appear similar but which are not the true correspondences. They give local minima in the distance function surface, but are not the true required deformation which gives the lowest or global minimum. Figure 1.5 illustrates global and local minima of a one-dimensional surface.

![Figure 1.5: Local minima (marked with thin arrows) and the global minimum (thick arrow) of a one-dimensional surface.](image)

Many registration methods therefore incorporate schemes for initially simplifying the distance function surface, removing many of the local minima. Some reduce details in the image pair leaving only the larger, more global structures to be matched first. Others restrict the initial deformations to be more global in nature, since it is more likely that strong, highly local deformations cause misregistrations. Once the global minimum has been approximated, complexity in the image data or in the range of allowable deformations is re-introduced, so that the solution can be finely tuned to find the optimum match. These methods of increasing complexity temporally are called hierarchical methods. We identify three different hierarchical strategies, as increasing in data complexity, in warp complexity and in model complexity. An additional advantage is that the initial matches are often performed rapidly due to a reduction in input data quantity or the calculation of a simplified transformation.

Medical images represent anatomies consisting of varying tissue types each of which deforms differently in nature - bones remain rigid, muscles stretch in preferential directions, etc. When registering images of the same subject, ideally these differences in true tissue behaviour should be imitated or approximated by registration schemes providing spatially varying deformation models. These are called inhomogeneous methods.

In this dissertation, we present three new inhomogeneous registration algorithms which are derived from a non-linear algorithm which models a viscous fluid. The fluid model allows severe local distortions and the creation of new regions from boundaries in the source. The three variants all have the ability to curtail the deformation in specified regions of the image, independently of the data in the image pair. In the first modification, referred to as modified fluid 1 (MF1), the data in the specified region are ignored in the application of forces to drive the registration, and hence the deformation of the region is due only to its context. An intended application is the registration of an image pair where the source contains a known additional structure for which no homologue exists in the target, but which may be confused with a region of similar intensity in the target.

The second variant, modified fluid 2 (MF2), enables specified regions to remain completely undeformed and unaltered. We intend it as the precursor to an algorithm which allows rigid body motion within the
1.1 Overview of General Image Registration

Registration of images is also required in non-medical imaging tasks, the two major non-medical fields being industrial vision (matching images of objects to templates for identification), and remote sensing of land images. In both cases, the images to be matched require only a slight deformation, if any, to achieve registration. An extensive survey of registration in these two fields, together with the medical case, is given in [Brown, 1992].

In industrial computer vision, images are of object shapes which can be approximated as combinations of geometric primitives, and so pattern matching concentrates on edge detection and template matching. Remotely sensed data [Heipke, 1997; Buiten and van Putten, 1997; Fonseca and Manjunath, 1996; Fogel, 1996; Goshtasby, 1988a] consist of images of a terrain acquired by a sensor mounted on aircraft or satellite. Depending on the geographic location and height of the sensor, the images may contain a few or many shapes approximating geometric primitives such as lines (roads, rivers) and circles (manhole covers, volcanic craters, lakes) as well as well-contrasted regions and edges. Image pairs presented for registration may be multi-sensor [Fonseca and Manjunath, 1996] (to fuse images which provide complimentary information) and multi-temporal (to detect changes over time) but always represent the same or fluid.

The third variant restricts the deformability of specified regions by altering the behaviour of the underlying physical model, by increasing its viscosity. The mathematical equation for the fluid model has been amended by Dr Kalvis Jansons of the Department of Mathematics, UCL, to effect the adaptation.

Having developed and/or selected a particular algorithm, and before introducing it into general use, it should first be evaluated using a representative set of image pairs to establish the accuracy of its results. Registration accuracy is measured locally in specified regions or at specified structures, or a global estimate is obtained for the image pair as a whole. The former case involves measuring the distance of features in the deformed source from their corresponding features in the target. For accurate registration, the features will be superimposed, and the distance will therefore measure zero (or a minimum). Appropriate features are landmark pairs or surface boundaries. A global estimate of registration is obtained by evaluating a function known as a (dis)similarity measure which compares the intensities (or a derivation thereof) of each spatially corresponding point in the image pair.

We have now introduced the main classes of medical registration algorithms according to their ability to reproduce the complexity of deformation required to register any reasonable image pair, and have briefly discussed the choice of algorithm. In the next part of the introduction we will set medical image registration in a broader context. First we will summarise image registration as applied in the non-medical tasks of industrial computer vision and remote sensing. We will then return to the medical field and consider registration as one element of a series of operations performed on medical image data. In particular, we devote a section to the acquisition of the data from four main modalities (CT, PET, ultrasound and MR).
1.2 Medical Image Data

This section aims to provide an understanding of the methods of acquisition of medical images and of the information contained within them. It is however only a brief summary. For further information on image acquisition the reader is referred to Brown and Semelka [1995]; Gadian [1995] and Webb [1988]. Westbrook [1994] and Weir and Abrahams [1997] give a less technical overview, with the latter illustrated with images from each modality.

Digital medical image data consist of 2D images, sets of 2D slices through a volume or full 3D volume images. The data are acquired by one of various scanning modalities: computed tomography (CT), magnetic resonance (MR) and ultrasound (US) providing anatomical information, and positron emission tomography (PET) and functional magnetic resonance (fMR) providing functional information, mainly used to monitor neurological activity. The term tomography originates from the Greek τόμος (tomos, section) and γράφω (graphos, writing), since these images are (often) acquired as (parallel) 2-dimensional slices through the 3-dimensional volume. The images are sampled into regular 2D arrays of pixels or 3D arrays of voxels. In the full 3D case we may expect isotropic voxel dimensions. However, when the volume is sampled as a series of discrete 2D slices, the inter-slice spacing may be relatively high compared to the inter-pixel distances within each slice; in this case we describe the latter as the in-plane resolution and the former as the between-plane, inter-slice or axial resolution.

X-ray CT

X-ray CT images (commonly referred to as CT) are produced by sending X-ray radiation through the subject and measuring the strength of the signal exiting on the opposite side. As in conventional X-rays, an image of the subject is formed by the differential in attenuation of the radiation by the various tissue densities. Bone reduces the signal the most compared to soft tissue, fat and water, and air attenuates the least; thus CT shows good contrast of hard tissue and soft tissues show up as shadows. For CT studies of soft tissue such as in the abdomen, a contrast medium (opaque to X-radiation) is administered in solution prior to scanning, timed so that it is passing through the volume of interest at scanning time.

Conventional X-ray imaging projects the radiation in parallel rays through the subject, capturing the exiting signal directly onto photographic film in a single plane giving a 2D image. CT creates a 3D volume of data by shifting and rotating the X-ray projector around the subject (figure 1.6, left); the subject may be passed through the scanner so that the data is collected in a helical path (figure 1.6, right). CT image...
1.2. Medical Image Data

resolution may be as high as 0.5 mm in all three dimensions. A computational method such as

![Figure 1.6: X-ray CT (left) projector and arc of detectors; (right) helical data collection.](image)

filtered back-projection (figure 1.7) is used to compute the estimated tissue density at grid points over the subject volume.

![Figure 1.7: X-ray CT: (left) strengths of attenuated X-rays are measured after passing through the subject. (right) back projection during estimation of tissue density at each voxel location.](image)

PET

PET (positron emission tomography) provides maps of function or neural activity. A radioactive isotope bound to an appropriate compound is injected into the subject’s bloodstream and perfuses the brain. Active areas of the brain require greater blood volume and hence have a higher uptake of the radioactive isotope. Radiation counts are detected by gamma cameras arranged around the surface of the head. As with X-ray CT, reconstruction of the image is an inverse problem, requiring some algorithm to compute the estimated isotope density at locations within the head. PET studies are often performed as time studies, with neural activity monitored while the subject performs simple physical, visual, aural or mental tasks. In this case, an isotope with a very short half-life, such as O^{15} (half-life 2 minutes) is used, with small doses and scanning repeated at approximately 10 minute time intervals to enable sufficient reduction in radioactivity concentration from the previous scan. PET studies are also useful for monitoring
1.2. Medical Image Data

brain function during strokes; in this case an isotope with a longer half-life is administered and the patient is scanned at a later time when he/she is fit to enter the scanner. The resolution of the PET images is fairly low; typical values are around $3 \times 3$ mm in-plane and 8 mm between-plane.

Ultrasound

Ultrasound (US) images are formed by interrogating the subject with high frequency sound waves, which reflect off surfaces at boundaries between tissues of different acoustic impedance. Ultrasound can also monitor movement or measure blood flow velocity due to the Doppler effect (the reflected sound waves from a receding surface appear to have a reduced frequency). Ultrasound is the least expensive of the scanning modalities, with images commonly acquired using a simple hand-held scanner, and is completely non-invasive. However the images obtained are of poor quality, with a low signal-to-noise ratio. Ultrasound is mainly used for monitoring heart motion and foetuses.

MRI

Magnetic resonance imaging (MRI) measures the slight magnetic field of nuclei such as hydrogen isotopes which are paramagnetic due to their containing an odd number of neutrons plus protons. These nuclei precess, that is, they exhibit a double rotation: they are spinning around an axis which itself is rotating at an angle about an axis aligned to the magnetic field (figure 1.8). The nucleus can absorb and emit energy in the form of electromagnetic (EM) radiation at its resonance frequency determined by the particular isotope and the strength of the magnetic field. For hydrogen isotopes in strong magnetic fields (between 0.5 and 7 tesla) the resonance frequencies are within the radio-frequency (RF) range. Hydrogen is abundant in soft tissues, in particular in body fluids and fat. A clinical MR scanner contains a strong magnet (field strength 1.5 tesla), and auxiliary magnets whose field strengths are gradated in each of the $x$, $y$ and $z$ directions. Hydrogen isotopes in the soft tissues of a subject inside the scanner align with the magnetic field (figure 1.9). The scanner also contains a transmitter coil and a receiver coil. The former sends out pulses of electromagnetic radiation at the resonance frequencies of the hydrogen nuclei; given a sufficiently long pulse (called a saturation pulse), the orientation of the magnetic fields of the nuclei flip to $90^\circ$ to that of the scanner (figure 1.10 A and B). The nuclei immediately start to lose their additional energy by emitting it at their resonance frequency at an exponential rate, called free induction decay or FID. These emitted signals are detected in the receiver coil. The particular frequency received specifies the strength of the magnetic field in which the emitting nucleus lies and hence identifies its location in
1.2. Medical Image Data

relation to the scanner magnet. The strength of the received signal is proportional to the density of nuclei aligned at an angle to the scanner’s magnetic field. At this stage, immediately after the saturation pulse, a spatial image can be formed of the subject by applying a Fourier transform to the data of magnitudes and frequencies of the detected RF signals. It is called a proton density image since it measures the density of excited nuclei at each spatial location.

MRI can provide further information about the subject’s tissues by varying the protocol of emitted pulse sequences and the length of delay between emitting and detecting RF signals. The two other most common MR images are a result of T1 and T2 relaxation. T1 relaxation is the loss in received signal due to FID returning the excited nuclei to their natural state, aligned to the scanner’s magnetic field (figure 1.10 C,D). Due to a differential in FID time for the hydrogen isotopes inside different tissue types, at an interval after the saturation pulse nuclei within some tissues will have relaxed back to their natural state while others will still be emitting RF signals detectable by the receiver coil, and the resulting reconstructed image will show the contrast in signal intensity.

Figure 1.10: T1 relaxation. A: Axis of nuclear spin aligned with magnetic field; B: after receiving RF pulse, axis at 90° to magnetic field; C,D: nucleus loses energy by emitting electromagnetic radiation and gradually re-aligns with magnetic field.

T2 relaxation is caused by some of the nuclei precessing at different rates, with the result that at some time after the saturation pulse the net magnetic field perpendicular to that of the scanner is zero (figure 1.11). As with T1, different tissues have different T2 relaxation times, and so a contrast image is formed.
1.2. Medical Image Data

Figure 1.11: T2 relaxation. A: After receiving RF pulse, the collection of nuclei in the region exhibit a net magnetic field perpendicular to that of the scanner magnet. B: Net perpendicular magnetic field becomes zero due to variable precession rates of the nuclei.

fMRI

Functional magnetic resonance imaging (fMRI) is a variant of MR used for monitoring functional activity, usually neural activity within the brain. Deoxyhaemoglobin (present in deoxygenated blood) is paramagnetic, unlike oxyhaemoglobin. The presence of paramagnetic 'impurities' causes a varying magnetic field, leading to a drop in the MR signal proportional to the concentration of deoxygenated blood. Neuroactivity requires increased bloodflow - more oxygenated blood goes to the areas of the cortex where there is activity, giving a positive change in the MR signal. However the change is only very slight and requires a very high magnetic field to be noticeable. Subtraction of an fMR image obtained during neuroactivity from a base-line image obtained at rest identifies regions of increased oxyhaemoglobin levels. Obtaining high temporal resolution to monitor the rapid fluctuations in oxyhaemoglobin levels necessitates a reduction in spatial resolution. Thus prior to a functional MR study at low spatial resolution, a high-resolution conventional anatomical MR image is obtained in the same scanner. The functional data are overlaid on the anatomical image to more accurately locate the areas of functional activity.

Comparison of MRI and fMRI to CT and PET

In the thesis we have used MR image data. The advantages of MR, especially in the context of neuroimaging, are its high contrast for soft, deformable tissues, and that it does not employ ionising radiation. In comparison to US and PET data, MR exhibits high signal-to-noise ratio. In addition, for functional studies, MR can produce both anatomical and functional data. MR is unsuitable for subjects with metal implants (such as metal prosthetics, pacemakers or metal fragments) due to the strong magnetic field. Apart from distorting the magnetic field they are likely to damage the subject by rapidly passing through him/her in attraction to the magnet, or by vibrating rapidly, causing friction burns. The magnet produces a loud banging noise as it moves when the field changes, and the bore of the magnet is fairly restrictive so most scanners are unsuitable for subjects who are claustrophobic. There are however new open-bore scanners coming into use; these also enable perioperative scanning.
1.3 Medical Image Processing

Medical images provide an abundance of anatomical and functional information. To extract and manipulate this information requires a series of tasks which may be described as an image processing pipeline. Figure 1.12 is a simple illustration of the medical image processing pipeline. The first step after image acquisition in extracting useful information is pre-processing. These methods enhance the quality of the image by emphasising and clarifying the useful information which it contains and suppressing or removing confusing or random information (noise) which may have been introduced by the scanner. Examples of pre-processing tasks are noise reduction [Jain, 1989, chap. 8], [Rudin et al., 1992; Black et al., 1998; Sijbers et al., 1998], enhancement of contrast between tissue types [Gauch, 1992] and corrections of geometric distortions introduced by the scanner [Maurer, Jr et al., 1996].

The pre-processing tasks improve image quality without examining what the information represents. The next two tasks in the pipeline, registration and segmentation, analyse and manipulate the data at a higher level, by considering the underlying anatomical (or functional) information represented by the image. Either one or both of these two tasks may be performed, and additionally one might be used as a tool to aid the other. Segmentation identifies different structures or tissue types within the image. Structures such as the cortical surface of the brain may be segmented prior to registration of two head volumes by a method which matches their cortical surfaces. Conversely, registration of a subject image to a labelled atlas is a means of obtaining a segmentation of the image.

The following is a list of typical structures which may be segmented from medical image data:

**surface of the brain** [Atkins and Mackiewich, 1996; Kapur et al., 1996]; for registration [Davatzikos, 1996] and in particular for registration of images acquired by different scanning modalities [Itti et al., 1997]; for localisation of activity monitored by electrical methods [Dale and Sereno, 1993]; to construct a *probabilistic* surface atlas, that is, to define the expected cortical shape using information from a population of subjects [Thompson et al., 1996]; or to describe the cortical shape of a particular individual [Schnabel, 1997]
1.3. Medical Image Processing

**deep structures of the brain** for shape analysis or radiation treatment planning [Szekely et al., 1996] or to build an anatomical atlas [Wells III et al., 1996; Sonka et al., 1996; Bucholz et al., 1997; Nowinski et al., 1997]

**the spinal chord** to describe its shape [Schnabel, 1997]

**tumours** as a diagnostic or surgical aid [Peck et al., 1996]

**contours of the face** prior to 3D visualisation [Cohen and Cohen, 1992]

**structures of the knee** for image-guided surgery [Smith, 1998]

**cavities of the heart** from noisy ultrasound or low spatial resolution - high temporal resolution MR data [Cohen and Cohen, 1990; Belohlavek et al., 1996; Kucera and Martin, 1997]

The above list also suggests some of the applications of the image manipulation. Main headings for these are listed in the final column of figure 1.12. **Data analysis** tasks extract further quantitative or qualitative information about the subjects(s). Typical quantitative information includes estimation of tumour volume [Peck et al., 1996] or brain volume [Saeed et al., 1997] since this is known to atrophy in subjects with Alzheimer's and schizophrenia. In functional imaging, the aim is to extract qualitative information regarding location of cortical areas responsible for specific functions. Examples in the literature are the somatosensory cortex [Gelnar et al., 1998], subdivisions of the auditory cortex [Wessinger et al., 1997], areas connected to cognitive tasks [Desmond et al., 1998] and to language [Thiel et al., 1998].

**3D visualisation** of the image volumes is particularly helpful for a wide range of applications [Toriwaki and Mori, 1998; Linney and Alusi, 1998]. These include diagnosis and surgery planning, where a 3D visualisation clarifies the location of tumours in relation to sensitive structures such as the optic nerve [Wells et al., 1996]. 3D visualisation techniques are divided between **volume rendering** and **surface rendering**; in the latter case the required surfaces must first be segmented from the image. In functional imaging studies, where the functional images are of poor spatial resolution, the locations of functional activity are extracted and superimposed as colour-coded regions on high resolution 2D anatomical images. Similarly, two scanning protocols may be applied to a subject, one giving high contrast for tumours and another giving better resolution overall for the anatomical structures; the two images obtained are then superimposed. In both cases, the images must be registered to ensure accurate superimposition.

The image data may be manipulated further in interactive systems for training, surgery simulation and image-guided surgery [Gibson et al., 1998; Troccaz et al., 1998; Colchester et al., 1996; Pareras and Martin-Rodriguez, 1996].

Having introduced the main concepts and context of medical registration, we will proceed to define the aims of this work and to specify its scope.
1.4 Definitions, Statement of the Problem, Aims and Scope

1.4.1 Definitions

We give here a list of definitions of terms to which we shall refer throughout the thesis. Further definitions relating specifically to inhomogeneous registration algorithms are given in chapter 6.

We will refer to both pixels and voxels as pixels when the context is applicable to both the two-dimensional and the three-dimensional case.

Definition 1.1 (Transformation) Let $\Omega_T$ be the domain $\{\vec{x}\}$ of pixel locations in the target image $T(\vec{x})$ and $\Omega_S$ be the domain $\{\vec{z}\}$ of pixel locations in the source image $S(\vec{z})$. Furthermore, let the source and target have the same domain such that $\Omega_T = \Omega_S = \Omega$. A transformation $\tilde{u}(\vec{x} \in \Omega_S) \mapsto \vec{y} \in \Omega_T$ of the source image is a function or a set of functions mapping $\vec{x} \in \Omega_S \to \vec{y} \in \Omega_T$.

Definition 1.2 (Registration) Let $U$ be a set of transformations $\{u_i(\vec{x}) = \vec{y}\}$. A transformation $u(\vec{x})$ is a registration of two images $S(\Omega), T(\Omega)$ if $\exists$ some distance function $D(S(\vec{x}), T(\vec{y}))$ of a feature space of $S$ and $T$ such that $D(S(u(\vec{x})), T(\vec{x})) \leq D(S(u_i(\vec{x})), T(\vec{x})) \forall u_i(\vec{x}) \in U$.

Typical feature spaces are

- pixel intensities $\{I(S(\vec{x})), I(T(\vec{z}))\}$;
- curvatures or spatial second derivatives of the intensities;
- segmented structure boundaries
- manually-selected landmark locations in the source and target image pair.

Definition 1.3 (Linear) A registration $u(\vec{x})$ of a source $S(\Omega)$ and target $T(\Omega)$ image is linear if $u$ is a linear function of $\vec{x}$ over $\Omega$.

We now consider dividing the image into a set of regions, or isolating regions within an image. This will have application in chapter 6. First we will defined adjacency for the 2-dimensional and 3-dimensional cases.

Definition 1.4 (4-adjacent) Two pixels $\vec{p} = (p_x, p_y)$ and $\vec{q} = (q_x, q_y) \in \mathbb{R}^2$ are 4-adjacent if one of the following two statements is true:

\[
|p_x - q_x| = 1 \quad \text{and} \quad |p_y - q_y| = 0
\]

\[
|p_x - q_x| = 0 \quad \text{and} \quad |p_y - q_y| = 1
\]

Definition 1.5 (8-adjacent) Two pixels $\vec{p} = (p_x, p_y)$ and $\vec{q} = (q_x, q_y) \in \mathbb{R}^2$ are 8-adjacent if they are 4-adjacent or if

\[
|p_x - q_x| = 1 \quad \text{and} \quad |p_y - q_y| = 1
\]
Definition 1.6 (6-adjacent) Two voxels \( p = (p_x, p_y, p_z) \) and \( q = (q_x, q_y, q_z) \) are 6-adjacent if one of the following three statements is true:

\[
\begin{align*}
|p_x - q_x| &= 1 \quad \text{and} \quad |p_y - q_y| = 0 \quad \text{and} \quad |p_z - q_z| = 0 \\
|p_x - q_x| &= 0 \quad \text{and} \quad |p_y - q_y| = 1 \quad \text{and} \quad |p_z - q_z| = 0 \\
|p_x - q_x| &= 0 \quad \text{and} \quad |p_y - q_y| = 0 \quad \text{and} \quad |p_z - q_z| = 1
\end{align*}
\]

Definition 1.7 (18-adjacent) Two voxels \( p = (p_x, p_y, p_z) \) and \( q = (q_x, q_y, q_z) \) are 18-adjacent if they are 6-adjacent or if one of the following three statements is true:

\[
\begin{align*}
|p_x - q_x| &= 1 \quad \text{and} \quad |p_y - q_y| = 1 \quad \text{and} \quad |p_z - q_z| = 0 \\
|p_x - q_x| &= 0 \quad \text{and} \quad |p_y - q_y| = 1 \quad \text{and} \quad |p_z - q_z| = 1 \\
|p_x - q_x| &= 0 \quad \text{and} \quad |p_y - q_y| = 0 \quad \text{and} \quad |p_z - q_z| = 1
\end{align*}
\]

Definition 1.8 (26-adjacent) Two voxels \( p = (p_x, p_y, p_z) \) and \( q = (q_x, q_y, q_z) \) are 26-adjacent if they are 18-adjacent or if the following statement is true:

\[
|p_x - q_x| = 1 \quad \text{and} \quad |p_y - q_y| = 1 \quad |p_z - q_z| = 1
\]

Definition 1.9 (n-Connected) Two pixels \( \bar{p}, \bar{q} \) are n-connected if there exists a set \( \{ \bar{r}_i \}, \ i \in \{1, \ldots, L\} \), such that

- \( \bar{p} \) and \( \bar{r}_1 \) are n-adjacent,
- \( \bar{r}_L \) and \( \bar{q} \) are n-adjacent, and
- \( \bar{r}_i \) and \( \bar{r}_{i+1} \) are n-adjacent

for \( n = 4, 8 \) if \( \bar{p}, \bar{q} \in \mathbb{R}^2 \) and \( n = 6, 18, 26 \) if \( \bar{p}, \bar{q} \in \mathbb{R}^3 \). We then write \( C_n(\bar{p}, \bar{q}) \).

Definition 1.10 (n-simply-connected) A set \( \mathcal{A} \) is n-simply-connected if \( C_n(\bar{p}, \bar{q}) \) for all \( \bar{p}, \bar{q} \in \mathcal{A} \).

Definition 1.11 (Region) A region \( \Omega_i \) in the domain \( \Omega \) is an n-simply-connected set of pixel locations \( \bar{z} \in \Omega \), where \( n \in \{4, 8\} \) if \( \Omega \subset \mathbb{R}^2 \) and \( n \in \{6, 18, 26\} \) if \( \Omega \subset \mathbb{R}^3 \).

Definition 1.12 (Partition) A partition of the domain \( \Omega \) is a set \( \mathcal{R} = \{ \Omega_i \}, i \in \Theta \) of regions \( \Omega_i \subset \Omega \) where \( \Theta \) is an indexing set,

\[
\bigcup_{i \in \Theta} \Omega_i = \Omega
\]

and

\[
\bigcap_{i \in \Theta} \Omega_i = \emptyset
\]
1.4. Definitions, Statement of the Problem, Aims and Scope

Definition 1.13 (Piecewise-linear) A registration \( \bar{u}(\bar{x}) \) of a source \( S(\Omega) \) and target \( T(\Omega) \) image pair is piecewise-linear if \( \exists \) a partition \( R \) of \( \Omega \) and \( \bar{u}(\bar{x}) \) is a set of functions \( \bar{u}_i(\bar{x}) \) where for each region \( \Omega_i \in R, \bar{u} \) is a linear function of \( \bar{x} \) constant over \( \Omega_i \).

Definition 1.14 (Non-linear) A registration \( \bar{u}(\bar{x}) \) of a source \( S(\Omega) \) and target \( T(\Omega) \) image pair is non-linear if \( \bar{u}(\bar{x}) \) is a higher-order function of \( \bar{x} \). If the registration \( \bar{u}(\bar{x}) \) consists of a set of functions \( \bar{u}_i(\bar{x}) \) defined over a partition \( R \) of the source domain \( \Omega \) then \( \bar{u} \) is non-linear if there exists at least one region \( \Omega_i \in \Omega \) such that its mapping function \( \bar{u}_i(\bar{x} \in \Omega_i) \) is a higher-order function of \( \bar{x} \).

Definition 1.15 (Deformation) A mapping \( \bar{u}(\bar{x}) \) of a source image \( S(\Omega) \) is a deformation if it is piecewise-linear or non-linear.

1.4.2 Statement of the Problem

The objective is to find a mapping \( \bar{u}(\bar{x}) \) from a source image \( S(\bar{x}) \) to a target image \( T(\bar{x}) \) so that homologous or anatomically corresponding features are identically located in the target and deformed source.

This goal is in most cases too demanding and is modified to require finding the mapping from a given set of functions which gives the closest match of corresponding features in respect to a selected distance measure.

An assumption is that \( T \) is a deformed version of \( S \), being images of the same anatomical structures. In practice this is not true, especially if \( T \) and \( S \) are scans of different individuals, obtained at different times during progression of, or treatment of, disease, or were produced by different scanning modalities.

However, for the simple case, we assume the problem to be:

**Determine the mapping \( \bar{u}(\bar{x}) \) such that**

\[
D(\phi(T(\bar{x})), S(\bar{u}(\bar{x})))
\]

**is minimised.**

where \( D \) is a given distance measure and where \( \phi(.) \) is an optional contrast mapping to modify the pixel intensities of the target.

The following are also required of the registration algorithm:

- the registration must be reproducible - deterministic - that is, given the same data twice, it should give the same registration result both times.

- a smooth mapping is encouraged so that the integrity of the image is maintained, without tearing or folding (although specific algorithms may be constructed to intentionally flaunt this rule).

- the registration should be robust with respect to noise, and without terminating in local minima

- a considerable degree of automation may be desirable
• occlusions or missing data should not affect the outcome of the registration.

• we do not require that any one algorithm be generically applicable to all registration applications, but we do require that the range of data for which it is valid be specified. We may require that the algorithm be suitable for inter-modality registration.

1.4.3 Aims

The aims of this project, within the scope stated below, were to:

1. survey and analyse the state of the art

2. evaluate the fluid algorithm with respect to other registration methods, and to

3. extend the fluid algorithm to the inhomogeneous case

1.4.4 Scope of this work

We limit the scope of our work to lie within the following boundaries:

• Other than in the chapter on linear registration, we will restrict our analysis to the intra-modality case, and in our evaluation we will use either real MR or synthetic data.

• Other than in the chapter on linear registration, we will not consider occlusions or missing data; we will assume the ideal case that the corresponding feature exists in the target for each feature in the source and vice versa.

• We will concentrate on the registration of static images - we will not discuss problems specific to gated heart studies or the real time perioperative case.

• The source and target images are both provided as intensities of a pixel lattice, stored in a sampled grid in two or three dimensions; the output deformed target is to be similarly presented.

• We will not research methods of optimising the computational speed or efficiency of algorithms.

• Other than the discussion on the use of hierarchical schemes to avoid misregistrations, we will not analyse the various search algorithms.

• Similarly we will not review or evaluate segmentation methods.

• The discussion on scale spaces will be restricted to its use as a tool to generate a rough hierarchy of detail. We will not study at depth the generation of the scale spaces nor optimise its temporal sampling.

• The evaluation of algorithms will be a demonstration rather than a thorough clinical trial.
1.4.5 Original contributions

A list of our contributions is given in the final chapter. We give here a brief summary.

Whereas there have been extensive surveys and evaluations of linear registration algorithms, there has not been a systematic analysis of non-linear algorithms. This we have achieved, while providing a structure within which any registration algorithm may be described according to its ability to determine in a robust manner the appropriate transformation of a given complexity.

We have constructed three variants of the viscous fluid algorithm, one for each of the three classes of inhomogeneity an algorithm may exhibit.

We have comparatively evaluated the fluid, landmark-based spline and SPM models, as well as the inhomogeneous versions of the fluid and spline.

To conclude, we are satisfied that we have met our objectives.

1.5 Structure of the Thesis

The thesis surveys a broad spectrum of medical image registration techniques, in particular for the application to MR images of the brain. We have arranged the review into four main chapters, dealing with linear, single-level non-linear, hierarchical and inhomogeneous methods.

Chapter 2 reviews linear registration methods, classifying them into four groups: principal axes, matching extracted features, matching by pair-wise pixel similarity measures and optical flow. We include this chapter for reasons explained in its introduction.

Chapter 3 commences a survey of basic non-linear algorithms by outlining representations of the deformations they compute, which are of three main types: a full discrete pixel displacement field; coefficients of a basis function expansion; or a combination, consisting of a sparse set of landmarks or nodes, with splines or finite elements as interpolants. Continuing the general analysis of non-linear methods, registration is described as an optimisation problem, using regularisation theory, and statistical and mechanical models to derive the cost function.

Rough analogies are drawn with linear methods in the presentation of specific algorithms in three groups: matching extracted features such as landmarks, with spline interpolation; registration with basis-function expansions driven by statistical similarity measures; and fluid registration, which, as in optical flow, follows image gradients over time. A fourth group is the elastic registration, whose historical development we summarise. The mathematical model of the spline is presented as a non-linear extension of the matrix methods for linear registrations, and we survey several variants. Both elastic and fluid registration can be computed using either a basis function decomposition or using finite differences in a relaxation method such as SOR. We chose to describe the computation of fluid registration using the latter.

Chapter 4 gives a new classification of hierarchical methods, as those of increasing data complexity, warp
1.5. Structure of the Thesis

We chose to concentrate on the analysis of scale space methods for increasing data complexity, and in chapter 5 give a quantitative comparison of the progress of the fluid algorithm in isotropic Gaussian scale space, closure scale space and without scale space. Inhomogeneous scale spaces may be of particular interest in the future development of anisotropic elasticity operators for inhomogeneous registration.

In the first part of chapter 5 we present results of a quantitative comparison of one algorithm from each of the three groups of deformation descriptors: the thin-plate spline (combination), SPM (basis function expansion) and fluid (discrete pixel displacement field).

In chapter 6 we classify inhomogeneous algorithms into those of variable data influence, of variable deformability, and of variable model type. We review three published inhomogeneous algorithms and present two modifications to the implementation of the viscous fluid algorithm. We then develop the fully inhomogeneous fluid model, where the viscosity is allowed to vary spatially over the source image domain. Chapter 7 presents an evaluation of inhomogeneous registration algorithms, giving results of application to a selection of 2D and 3D data.

Chapter 8 concludes the main body of the thesis with a summary and suggestions for future work.
Chapter 2

Linear Registration Methods
2.1 Introduction

We begin our analysis and review of medical registration algorithms by devoting a chapter to those in which the movement of the source image is restricted to linear transformations only. Although this dissertation aims to present non-linear registration methods, we have included this chapter for three reasons. Firstly, since historically the linear registration problem was the precursor to the non-linear. Secondly non-linear registration is almost always preceded by a linear registration, which is usually much faster. Lastly, many of the considerations and techniques used in the solution to the linear problem are applicable in the non-linear case.

For intra-subject matching of anatomies whose physical deformation is minimal, such as the brain which is constrained within the cranial cavity, rigid body motion may be assumed*. That is, the required mapping of the source will consist of a simple translation and/or rotation. Where the image dimensions or voxel dimensions in the source and target are different, the transformation may also require a scaling. These transformations are all linear, that is, they are of the form

\[ \bar{u}(x, y, z) = \begin{pmatrix} a_0 + a_1x + a_2y + a_3z \\ a_4 + a_5x + a_6y + a_7z \\ a_8 + a_9x + a_{10}y + a_{11}z \end{pmatrix} \]

where the \( a_i \) are all constants. Depending on the constants chosen, the transformation may involve shears or changes in perspective^1. Medical image registration rarely requires these transformations, other than to correct for distortions introduced by the imaging device^2.

Brown [1992] identifies four key components in the development of a registration algorithm: the feature space, where features to be matched are selected in the image pair; the search space, or range of transformations to be used; the search algorithm for finding the optimum transformation within the search space; and the similarity measure or distance function to measure the optimality of a transformation for the selected feature set. In practice a similarity measure may be replaced by a dissimilarity measure^3, which is to be minimised. For linear registration, the search space is simply the set of all linear transformations, which are described by a small number of parameters.

The search strategy is chosen from one of several standard optimisation algorithms, which are discussed in any good numerical analysis text; Press et al. [1995] is also a good source of these. Generally an initial estimate of the linear mapping can be refined by repeatedly updating the transformation parameters to improve on a similarity measure extracted from the image pair. We recall that in image registration the function to be minimised is not a continuous surface but is discrete due to the discrete nature of the images. In addition, it may contain an abundance of local minima and so the minimisation algorithm to be selected

---

*Although the brain itself can deform very slightly. Additionally there are conditions in which the normal or abnormal brain may grow or shrink (atrophy), and in inter-modality studies one of the modalities may introduce a slight deformation. Finally the assumption is invalid if the studies are pre-/post-operative or are time studies of developing lesion(s).

^1 using a modified form of equation 2.1 with 15 constants, see page 47.

^2 Although shears may be used as decompositions of rotations as in Danielsson and Hammerin [1992]

^3 This is nevertheless often referred to as a similarity measure.
should be robust to local minima traps. To avoid local minima, some hierarchical strategy is suggested, such as matching in scale space. Hierarchical strategies are described in chapter 4.

The variety in linear registration methods is due to the selection of similarity measures and feature spaces, in particular with respect to their applicability to inter-modality registration. The algorithms concentrate on extracting features in the image pair, and on determining correspondences between them.

In the summary that follows we group linear registration algorithms into four:

1. The principal axes transformation
2. Feature matching
3. Pixel matching methods
4. Directed search based on optical flow

The first three groups are according to the major choices of feature spaces. Optical flow is an interesting technique which may be thought of as the linear analogue to the non-linear fluid algorithm which we describe in chapter 3. All except for the optical flow have adaptations for application to inter-modality matching. We give a brief overview of the groups here and then study each separately at greater depth in sections 2.3–2.6.

Principal axes

Registration of principal axes is almost universally used only for a fast rough registration prior to other algorithms which require a good initial approximation. Essentially each image is implicitly approximated as an ellipsoid and the matching is by the location and orientation of their principal axes. This is the only group of the four which derives directly in one step an analytic solution to the matching problem; the remaining methods require iterative approximations to the optimal solution and thus have an increased computational burden.

Feature matching

Under this heading we include matching by geometric invariants extracted from the images in preprocessing steps which may or may not be automated. The implication is that the image information is summarised by a finite number of significant features. A general requirement is the guarantee (however not always given) that the features selected for matching form corresponding sets and are accurately isolated, although statistical methods may take outliers into account. For inter-modality matching the assumption is that the features chosen are readily identifiable in each modality.

Pixel matching

This group of linear registration algorithms minimise a statistical measure such as cross-correlation which compares the intensity values of the image pair pixel-wise, [Woods et al., 1992]. Most have been specif-
2.2. Linear transformations

We assume the reader is fully conversant with all the material in this section; we include it only for comparison with that presented later in section 3.5.1.

Linear registration methods apply linear transformations globally to the coordinates of the pixel set. The transformations may be written in matrix notation, with matrix operators for the elemental transformations of rotation, translation and scaling (and shearing and change in perspective if required). To allow translation (and perspective change) to be performed by matrix multiplication, the pixel coordinates must be written as homogeneous coordinates, that is, raised into a space of dimension one higher. 2D coordinates \((x, y)\) are represented as \(\alpha(x, y, 1)^T\) and 3D \((x, y, z)\) as \(\alpha(x, y, z, 1)^T\), where \(\alpha\) is any strictly positive scalar. The transformations applied to 3D images are represented by the following elemental matrices:
2.2. Linear transformations

scaling

\[
S_x = \begin{pmatrix}
  s_x & 0 & 0 \\
  0 & 1 & 0 \\
  0 & 0 & 1
\end{pmatrix}, \quad
S_y = \begin{pmatrix}
  1 & 0 & 0 \\
  0 & s_y & 0 \\
  0 & 0 & 1
\end{pmatrix}, \quad
S_z = \begin{pmatrix}
  1 & 0 & 0 \\
  0 & 1 & 0 \\
  0 & 0 & s_z
\end{pmatrix}
\]

where \( s_x, s_y \) and \( s_z \) are scaling parameters in the \( x, y \) and \( z \) directions respectively. These three elemental matrices can be combined by matrix multiplication into one rescaling operator \( S_{x,y,z} \)

\[
S_{x,y,z} = \begin{pmatrix}
  s_x & 0 & 0 \\
  0 & s_y & 0 \\
  0 & 0 & s_z \\
  0 & 0 & 1
\end{pmatrix}
\]

rotation, angle \( \theta \) about \( x \) axis

\[
R_{x;\theta} = \begin{pmatrix}
  1 & 0 & 0 & 0 \\
  0 & \cos \theta & -\sin \theta & 0 \\
  0 & \sin \theta & \cos \theta & 0 \\
  0 & 0 & 0 & 1
\end{pmatrix}
\]

rotation, angle \( \phi \) about \( y \) axis

\[
R_{y;\phi} = \begin{pmatrix}
  \cos \phi & 0 & \sin \phi & 0 \\
  0 & 1 & 0 & 0 \\
  -\sin \phi & 0 & \cos \phi & 0 \\
  0 & 0 & 0 & 1
\end{pmatrix}
\]

rotation, angle \( \psi \) about \( z \) axis

\[
R_{z;\psi} = \begin{pmatrix}
  \cos \psi & -\sin \psi & 0 & 0 \\
  \sin \psi & \cos \psi & 0 & 0 \\
  0 & 0 & 1 & 0 \\
  0 & 0 & 0 & 1
\end{pmatrix}
\]

translation

Properly, an elemental translation matrix produces translation in only one of the \( x, y \) or \( z \) directions. So for elemental matrices two of the three translation parameters \( t_1, t_2, t_3 \) in the following are set to zero:

\[
T_{x,y,z; t_1,t_2,t_3} = \begin{pmatrix}
  1 & 0 & 0 & t_1 \\
  0 & 1 & 0 & t_2 \\
  0 & 0 & 1 & t_3 \\
  0 & 0 & 0 & 1
\end{pmatrix}
\]
change in perspective

\[
P_{x,y,z,p1,p2,p3} = \begin{pmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
p1 & p2 & p3 & 1
\end{pmatrix}
\]

with only one of \(p1, p2, p3\) non-zero for the elemental case.

concatenation

Two or more elemental operators may be concatenated by matrix multiplication to produce a more general linear transformation; note that as matrix multiplication is not commutative, the order of the operations is important and the matrix decomposition of a general linear transformation is not unique.

A general 3D linear operator has the form

\[
M = \begin{pmatrix}
R & T \\
P & 1
\end{pmatrix}
\]

where \(R\) is a block matrix determining rotations, scaling and shears, \(T\) is a column vector of translations and \(P\) introduces perspective distortions.

The 3D linear transformation may thus be summarised by a number of parameters, depending on its flexibility. Six parameters are needed for a rigid body transformation (three translations, three rotations), and a further three for scaling (or one for uniform or isotropic scale). The more general affine transformation allowing shears requires twelve in total, and a full perspective transformation requires fifteen.

### 2.3 Principal Axes

The principal axes transformation for medical image registration originated in papers by Gamboa-Aldeco and Chen [1986], and by Faber and Stokely [1988]. In [Gamboa-Aldeco and Chen, 1986] the authors apply the transformation using only the surface pixels of the volume of interest, and thus errors in outlining the surface have greater effect.

The basic principal axes method gives a rough estimate for the six affine transformation parameters of translation and rotation; a more sophisticated version will additionally estimate scaling in three mutually perpendicular directions. The theory may be approached from either a physical or statistical viewpoint. Either way, for each image, a matrix is constructed describing the image shape approximated as an ellipsoid, of which a centre is determined. The centre is not necessarily the geometric centre of the shape but may be weighted by the variation of pixel intensities across the image. Perpendicular axes crossing at the
centre are calculated, together with weighting or distance measures along each axis from the centre. The images are matched by translating the source until its centre coincides with that of the target, and rotating it until the axes are identically orientated (figure 2.1). Finally, the source may be scaled using ratios of the calculated distance measures.

![Figure 2.1: Principal axes registration of a source (S) and target (T) image pair.](image)

The Physical Model

Using theory from classical mechanics [Synge and Griffith, 1959], the position and orientation of each image is specified by its centre of mass and principal axes determined from an inertia tensor.

The image is considered a rigid body consisting of point masses, these being the pixel intensity values. The centre of mass is given by

$$
\bar{x} = \frac{\sum I(\vec{x})}{\sum I(\vec{x})}
$$

where $I(\vec{x})$ is the image intensity at pixel $\vec{x} = (x_1, x_2, x_3) \equiv (x, y, z)$. Alternatively, a volume of interest (VOI) may be defined and $I(x)$ considered as a binary function:

$$
I(\vec{x}) = \begin{cases}
1 & \text{if } \vec{x} \text{ inside VOI} \\
0 & \text{if } \vec{x} \text{ outside VOI}
\end{cases}
$$

in which case $\bar{x}$ is the geometric centre of the VOI. This is an approach used in inter-modality image registration [Alpert et al., 1990; Dhawan et al., 1995], where the pixel intensities have different distributions in each image but where the corresponding volumes of interest approximate the same shape. Using the binary function to give the support of the images reduces the problem to an equivalent of the intra-modality case.

For a continuous rigid body with homogeneous density $\rho$, the inertia tensor $\Omega$ is given by

$$
\Omega_{ij} = \int \rho \left[ r^2 \delta_{ij} - x_i x_j \right]
$$

where $r^2 = x_1^2 + x_2^2 + x_3^2$.

For the discrete version as applied to image registration, we set the density as $\rho(\vec{x}) = \frac{I(\vec{x})}{\sum_{\vec{x} \in \Omega} I(\vec{x})}$ where $\Omega$ is the image domain. Then

$$
\hat{\Omega}(\vec{x}) = \begin{pmatrix}
\Omega_{xx} & -\Omega_{xy} & -\Omega_{xz} \\
-\Omega_{xy} & \Omega_{yy} & -\Omega_{yz} \\
-\Omega_{xz} & -\Omega_{yz} & \Omega_{zz}
\end{pmatrix}
$$
2.3. Principal Axes

with

\[
\begin{align*}
\Omega_{xx} &= \sum_{(x,y,z)} \left( (y - \bar{y})^2 + (z - \bar{z})^2 \right) \rho(\bar{z}) \\
\Omega_{yy} &= \sum_{(x,y,z)} \left( (x - \bar{x})^2 + (z - \bar{z})^2 \right) \rho(\bar{z}) \\
\Omega_{zz} &= \sum_{(x,y,z)} \left( (x - \bar{x})^2 + (y - \bar{y})^2 \right) \rho(\bar{z}) \\
\Omega_{xy} &= \sum_{(x,y,z)} (x - \bar{x})(y - \bar{y}) \rho(\bar{z}) \\
\Omega_{xz} &= \sum_{(x,y,z)} (x - \bar{x})(z - \bar{z}) \rho(\bar{z}) \\
\Omega_{yz} &= \sum_{(x,y,z)} (y - \bar{y})(z - \bar{z}) \rho(\bar{z})
\end{align*}
\]

The 2D image is analogous to a thin-plate, modelled as having infinite thinness in the vertical direction. In the continuous case, the density function \( \rho(x, y, z) \) is replaced by \( \sigma(x, y) \delta(z) \) where \( \sigma(x, y) \) gives the mass per unit area and \( \delta(z) \) is a dirac-\( \delta \) function. The inertia tensor then reduces to

\[
\begin{pmatrix}
\int \sigma y^2 & -\int \sigma xy & 0 \\
-\int \sigma xy & \int \sigma x^2 & 0 \\
0 & 0 & \int \sigma (x^2 + y^2)
\end{pmatrix}
\]

In the discrete 2D case we use the matrix \( \tilde{\Omega}(\bar{z}) = \begin{pmatrix} c_{xx} & c_{xy} \\ c_{xy} & c_{yy} \end{pmatrix} \) (symmetric)

where

\[
\begin{align*}
c_{xx} &= \frac{\sum_{(x,y)} (y - \bar{y})^2 I(\bar{z})}{\sum_{x,y} I(\bar{z})} \\
c_{xy} &= \frac{-\sum_{(x,y)} (x - \bar{x})(y - \bar{y}) I(\bar{z})}{\sum_{x,y} I(\bar{z})} \\
c_{yy} &= \frac{\sum_{(x,y)} (x - \bar{x})^2 I(\bar{z})}{\sum_{x,y} I(\bar{z})}
\end{align*}
\]

The principal axes of the body are given by the eigenvectors of \( \tilde{\Omega} \) and the eigenvalues of \( \tilde{\Omega} \) are its principal moments of inertia.

The centres of mass and principal axes are found for each image and then the required translation, rotation (and scale - optional) are applied to bring them into register. The exact parameters for each will vary depending on the order in which these operations are done, since the transformations do not commute.

The images supplied to a principal axes algorithm may first require a preliminary adjustment if there are pixels whose intensities are represented by negative values: all pixels values in the image are first raised by the absolute size of the minimum pixel intensity to avoid negative 'masses'.
We compare here three variants that apply the transform to slightly different registration tasks. Alpert et al. [1990] match whole binarised brain volumes without rescaling. First a translation is applied to match the centres of mass. The eigenvectors of the inertia tensor give the orientation of the image coordinate axes from the principal axes. Hence the relative rotations (around the three coordinate axes of the source image) required to bring the source into register with the target are found by multiplying the matrix of eigenvectors of the source by the inverse of the matrix of eigenvectors of the target.

Arata et al. [1995] likewise operate on binary images, this time of specific brain structures such as the ventricles, to generate an atlas of (normal) shape variation. In this respect an isotropic scaling is required, for which they use the cube root of the ratio of volumes of target and source. Again, the normalised eigenvector matrix of the inertia tensor is equated to the general form of rotation matrix to give the required angles about each axis. All structures are registered into the coordinate system of their principal axes.

An anisotropic scaling is applied by Lester [1995]. The eigenvalues are the moments of inertia about the principal axes. Since the moment of inertia of a particle about an axis equals the mass of the particle multiplied by the square of its distance from the axis [for example Synge and Griffith, 1959], an approximation to the length of a shape along its axes can be obtained from the square root of its eigenvalues. In addition, since \(\Omega\) is symmetrical about its leading diagonal, its eigenvalues are always real. Hence the required scaling along each principal axis can be estimated from the square root of eigenvalue ratios. The scalings and rotations of the image are performed about its centre after a centre-of-mass translation to the origin of the world coordinates (figure 2.2).

Discussion

The Principal Axes method is analytic; the transformation is determined directly without the need for an iterative approach and hence it is fast and there are no convergence problems. It only gives a rough solution since many mass distributions have the same principal axes. It is affected by incomplete data (such as when registering CT to MR: often only part of the head is imaged with CT to reduce radiation exposure) since this shifts the centres of mass and their locations no longer approximate anatomical correspondences. Dhawan et al. [1995] give a variant on their algorithm in [Arata et al., 1995], which compensates for this problem by iteratively selecting a sub-volume from the larger volume with which to match the smaller.

For shapes with many axes of symmetry (such as the circle which has infinitely many), multiple pairs of
2.4. Feature Matching Methods

In this section we subdivide feature matching into landmark-based methods, surface matching, matching of crest lines and of involutions in the convoluted surface of the cortex. For recent surveys, see Maintz et al. [1996]; Merickel [1988]; Maguire Jr et al. [1985].

2.4.1 Landmark-based

In landmark-based matching methods, homologous points known as landmarks or fiducials are identified either manually or automatically at corresponding anatomical locations in the image pair, and their relative positions are used to compute the required transformation. Some medical images are acquired with fiducial markers placed on the subject for this purpose [Hill, 1993]. Generally though it is difficult and time-consuming to identify corresponding landmark sets. If more than four landmarks are identified in the 3D data sets, uncertainty in the landmark locations can be considered and a solution found which minimises the distance function in a least-squares sense.

2.4.2 Surface matching

Members of this group are the most ‘ad hoc’, in that the only unifying theory is the minimisation of inter-surface distances from pre-segmented structures. The methods vary in their segmentation, definition of the distance function and choice of search algorithm. Typically, curves or surfaces are extracted by slice-by-slice edge detection or morphological processes of dilation and erosion, and the image pixels are assigned classification values of inside/outside/surface. Matching algorithms are then applied which minimise the least squares distances of the (corresponding) surfaces. Accuracy of the method is limited by the initial feature extraction, upon which the registration is highly reliant, as often no other information is taken into account once the surfaces have been segmented. An advantage is that if the features are identifiable in both images, the method can be applied to inter-modality matching; two related structures may also be matched, such as the inside of the skull (in CT) to the outer surface of the brain/meninges (in MR). Brain surface matching may also be applied to PET images if transmission scans are used (with each emission scan, a transmission scan is produced, which is used to correct attenuation in the emission scan).

The ‘Head-hat’ algorithm by Levin, Pelizzari, Chen, Chen, and Cooper [1988]; Pelizzari, Chen, Spelbring, Weichselbaum, and Chen [1989] is the classic surface matching algorithm. The skull or brain surfaces are first segmented slice by slice; the larger of the image surfaces is stored as the ‘head’, represented

\footnote{If the spatial sampling distances in the z-direction were known in relation to the in-plane sampling, this could be corrected by weights when calculating the centres of mass and the inertia tensor.}
as the set of structure contours; the surface of the other image is stored as a set of sampled points. Rays are projected from the centre of mass of the ‘head’ out to each ‘hat’ point, and the point of intersection with the ‘head’ surface is computed by trilinear interpolation of the slice contours. The mean square distance between the intersection points and the corresponding ‘hat’ points is used as the error metric to be minimised.

A similar method, for 2D images, is proposed by Cai et al. [1996] where the rays this time are parallel to a coordinate axis, and the mean square distance between corresponding curves is minimised. They used Powell minimisation as a search strategy [Press et al., 1995], but suggested simulated annealing, downhill simplex and quasi-Newton. Clearly convergence to the correct solution was problematic. They also mention the problem of a curve being multi-valued with respect to the axes (figure 2.3), due to involutions in its surface. They aid the matching by including user-selected landmark pairs whose distance measure must likewise be minimised.

![Figure 2.3: A 1-dimensional curve multi-valued at $x = x_1$.](image)

The initial relative positions of the surfaces has a particular effect on the guarantee of convergence - especially if more than one pair is to be matched; since the algorithm does not know which pairs correspond, it will match those which are closest. Nelson et al. [1997] address the problem of requiring an initial good matching estimate by first performing volume matching using the binarised volume inside the surface. Gilhuijs and van Herk [1993] use chamfer matching which is a directed search: The chamfer distance transform (appendix A.1) supplies a distance gradient of each pixel from the surface; the surfaces are moved relatively up the distance gradient until coincidence. Hill et al. [1993] describe two methods one of which uses prior anatomical knowledge of the relative locations of different anatomical structures. A modified chamfer transform which has increasing positive values for pixels distant from the outside surface and decreasing negative values for internal pixels, supplies internal/external information, described as ‘containment’. Convergence problems are addressed by a stochastic search over scale space (section 4.2). In this way, the initial matching is simplified - gross structure is registered first, and remaining local minima may be escaped by the random minimisation process.

### 2.4.3 Feature matching based on crest-lines

Thirion et al. [1992] match crest lines, defined as the loci of maxima of extremal surface curvature, the 3D equivalent of corners. The surface is assumed to be parametrised by $\mathbf{r}(u, v)$, where $(u, v) \in V$, where $V$ is some open subset of the real numbers $\mathbb{R}$. Crest lines are computed from the extremality criterion

$$\nabla k_M \cdot \hat{e}_M = 0$$
where \( k_M \) is a local maximal curvature of the parametrised surface \( \vec{r}(u, v) \) and \( c_{\vec{M}} \) is its corresponding principal direction (figure 2.4), being respectively maximum eigenvalue and associated eigenvector of the characteristic equation:

\[
\det \left\{ \left( \begin{array}{cc} L & M \\ M & N \end{array} \right) - k \left( \begin{array}{cc} E & F \\ F & G \end{array} \right) \right\} = 0
\]

where

\[
\begin{align*}
L &= \langle \vec{n}, r_{11} \rangle \\
M &= \langle \vec{n}, r_{12} \rangle \\
N &= \langle \vec{n}, r_{22} \rangle \\
E &= \langle r_{11}, r_{11} \rangle \\
F &= \langle r_{11}, r_{21} \rangle \\
G &= \langle r_{22}, r_{22} \rangle
\end{align*}
\]

and

\[
\vec{n} = \frac{r_{11} \wedge r_{22}}{|r_{11} \wedge r_{22}|}
\]

is the normal to the plane tangent to the surface, spanned by two independent tangents \( r_1 = \frac{\partial \vec{r}}{\partial u} \) and \( r_2 = \frac{\partial \vec{r}}{\partial v} \), \( r_{11}, r_{12} \) and \( r_{22} \) are the second derivatives of the surface with respect to its parametrisation, \( \langle \vec{a}, \vec{b} \rangle \) denotes the inner (scalar) product and \( \wedge \) the vector cross-product. The reader is referred to standard texts on differential geometry [Spivak, 1979; do Carmo, 1976].

As an aside, the mean and Gaussian curvatures, \( H \) and \( K \) are related to the maximal and minimal curvatures \( k_M \) and \( k_m \) by:

\[
\begin{align*}
H &= \frac{1}{2} (k_M + k_m) \\
K &= k_M k_m
\end{align*}
\]

Figure 2.4: Principal directions \( c_M \) and \( c_m \) lying in the tangent plane.
2.4. Feature Matching Methods

Thirion et al. [1992] finds the most likely correspondences between the identified crest lines as follows. Each point on each crest line has a spatial location in 3D space together with the normal and principal directions which give the orientation of the curve. Any crest point pair from the two images therefore has a unique 6-parameter affine relative transformation. All possible such transformations are binned, with the bins being 6-element vectors in the transformation space. The fullest bin indicates the most likely transformation to bring the images into register.

2.4.4 Feature matching based on involutions

The approach of Banerjee et al. [1995] is similar in that selected points on a segmented 2D edge are assigned parameters representing its geometry. Whereas the method of Thirion et al. [1992] is applied to skull surfaces, Banerjee et al. match 2D cortical boundaries and concentrate on their involutions (called sulci). These are extracted by subtracting the cortex surface from its convex hull (figure 2.5).

![Convex hull of a surface.](image)

The depth of each involution together with its two entry/exit angles at the surface of the convex hull defines a triangle which can be transformed into a standard triangle whose shape is described as being of canonical isosceles form using a 6-parameter affine transformation. The involution correspondences are decided by finding those with the same transformation to the canonical isosceles form. Finally they check the final registration by comparing the centres of mass of each image.

![Cortical involution, with convex hull as dotted line, showing depth and entry/exit angles to convex hull](image)

2.4.5 Discussion

Feature-matching methods lend themselves easily to the inter-modality case as long as corresponding features can be extracted from both images. They do however require the prior segmentation of the features,
which is a difficult procedure both when performed manually or automatically, and the identification of correspondences between the extracted features. Where only few features are extracted, accuracy of registration depends heavily on the reliability of their location, while other information present in the image pair is ignored. In the next section, a match is obtained using all (or most) of the available information, by considering the sets of pixel intensities of the image pair.

2.5 Pixel Matching Methods

The third group of algorithms discussed in this chapter compute a (dis)similarity measure from comparisons of pixels at identical locations in the image pair. The transformation is adjusted iteratively until the (dis)similarity measure has been minimised. Three popular pixel intensity (dis)similarity measures are the sum of squared differences, the correlation coefficient and the mutual information measure. We define each of these below and then elaborate on the use of the latter two in linear registration. A further measure is the variance of intensity ratios for which we reproduce the steps in its computation.

2.5.1 Pair-wise pixel intensity measures

Sum of Squared Differences

The sum of squared differences is computed as

$$D_d(S, T) = \sum_{\vec{x} \in \Omega} (S(\vec{x}) - T(\vec{x}))^2$$  \hspace{1cm} (2.2)

where the sums are taken over all pixel locations $\vec{x} \in \Omega$. It ideally measures zero for an intra-subject intra-modality image pair in complete registration but in practice this lower bound is not attained due to noise present in both images, and possible underlying fluctuations in intensity caused by inhomogeneities in the imaging device. The squaring strengthens the misleading effect of noise. Further, it is inapplicable to inter-modality matching since in this case we cannot assume that corresponding tissues are represented by similar intensities in both modalities - often the images are complimentary, in that they display disparate contrasts of tissue structures. The sum of squared differences is popular for driving nonlinear registration methods since it is fast to compute and its derivative is easy to obtain. Refer to sections 3.4.1, 3.6.1 and 3.8.

Correlation coefficient

The correlation coefficient is a normalised coefficient measure calculated as [Gonzalez and Woods, 1992]

$$D_c(S, T) = \frac{\sum_{\vec{x} \in \Omega} (S(\vec{x}) - \bar{S}) (T(\vec{x}) - \bar{T})}{\left(\sum_{\vec{x} \in \Omega} (S(\vec{x}) - \bar{S})^2 (T(\vec{x}) - \bar{T})^2\right)^{\frac{1}{2}}}$$  \hspace{1cm} (2.3)

where $\bar{S}$ and $\bar{T}$ are mean intensity values of source and target images respectively. The correlation coefficient rises to the value 1 for perfectly registered images. An alternative is to use the correlations of gradients of images. This avoids possible bias from intensity inhomogeneities due to a varying field in the scanning device, which, as slowly varying, are not so apparent in the intensity gradients.
2.5. Pixel Matching Methods

Mutual Information

Mutual information [Maes et al., 1997], derived from information theory, measures the statistical dependence of one image on another. It is associated with the probability of accurately predicting the pixel intensities in one image given the intensities at corresponding locations in the other, and is calculated as

$$D_i(S,T) = \sum_{s,t} p_{ST}(s,t) \log_2 \frac{p_{ST}(s,t)}{p_S(s)p_T(t)}$$  \hspace{1cm} (2.4)

from the joint distribution $p_{ST}$ and marginal distributions $p_S$, $p_T$ taken from histograms of pixel intensities in the source and target images. Mutual information gives a greater value for better registration, reflecting the higher probability of accurately predicting pixel intensities in the deformed source given those in the target.

Derivation of mutual information measure

The source and target images $S$ and $T$ are each a set of pixel intensities $\{s\}$ and $\{t\}$. The histogram $p_S(s)$ is the distribution of intensities $s$ in the source, and for a specific value of $s$ gives the probability of any pixel having that value. It is computed as

$$p_S(s) = \frac{\text{number of occurrences of intensity } s \text{ in } S}{\text{total number of pixels in } S}$$

$p_T(t)$ is similarly defined.

The joint distribution $p_{ST}(s,t)$ is a 2-dimensional histogram showing the distribution of pairs of intensity values in the source/target image pair. In practice, $p_{ST}(s,t)$, $p_S(s)$ and $p_T(t)$ are computed only for the region of overlap of the two images. $p_S(s)$ and $p_T(t)$ are marginal distributions and are found by summing columns or rows of the joint distribution $p_{ST}(s,t)$.

To understand the mutual information measure, some elementary probability theory is required.

Two random variables $A$, $B$ (grey level intensities in source and deformed target image) are said to be independent if

$$p_{AB}(a,b) = p_A(a)p_B(b)$$  \hspace{1cm} (2.5)

and maximally dependent if a one-to-one mapping $\phi$ exists, as included in equation (1.1), such that

$$p_A(a) = p_B(\phi(a)) = p_{AB}(a,\phi(a))$$  \hspace{1cm} (2.6)

so given a pixel value in $A$ we can predict the pixel value at the same location in $B$.

Additionally we must define the concept of entropy:

The information of a particular greylevel $k$ is given by [Jain, 1989]

$$I_k = -\log_2 p(k) \text{ bits}$$

The average information over the whole image is $\sum_{k=1}^{n} I_k p(k)$ which is [Jain, 1989; Sonka et al., 1993]

$$H = -\sum_{k=1}^{n} p(k)\log_2 p(k)$$  \hspace{1cm} (2.7)
This is called the *entropy* of the image. For independent pixels, $H$ is estimated from the histogram of pixel intensities. Maximum entropy is achieved when the distribution is uniform, that is, any pixel intensity in the range $k = 1, \ldots, n$ is equally likely, and so for all $k$, $p(k) = \frac{1}{n}$. Then, from equation (2.7),

$$\max_p H = -\sum_{k=1}^{n} \frac{1}{n} \log_2 \frac{1}{n} = -n \left( \frac{1}{n} \log_2 \frac{1}{n} \right) = \log_2 \left( \frac{1}{n} \right) = -\log_2 n$$

To measure the mutual information in the image pair, we sum the average information (entropy) in both source and target and subtract the total information content of the two images taken together.

$$D_i(S, T) = -H(S, T) + H(S) + H(T)$$

$$= \sum_{s,t} p_{ST}(s,t) \log p_{ST}(s,t) - \sum_s (\log p_S(s)) p_S(s)$$

$$- \sum_t (\log p_T(t)) p_T(t)$$

by equation (2.7)

$$= \sum_{s,t} p_{ST}(s,t) \log p_{ST}(s,t) - \sum_s \left( \log p_S(s) \left( \sum_t p_{ST}(s,t) \right) \right)$$

$$- \sum_t \left( \log p_T(t) \left( \sum_s p_{ST}(s,t) \right) \right)$$

(the reverse of summation to marginal distributions)

$$= \sum_{s,t} p_{ST}(s,t) \log p_{ST}(s,t) - \sum_{s,t} p_{ST}(s,t) \log P_S(s)$$

$$- \sum_{s,t} p_{ST}(s,t) \log P_T(t)$$

$$= \sum_{s,t} p_{ST}(s,t) \log \left( \frac{p_{ST}(s,t)}{P_S(s) P_T(t)} \right)$$

as given above in equation (2.4).

As an interesting aside, mutual information is the discrete version of the Kullback-Leibler measure of divergence [Ripley, 1996] between two continuous distributions on the same domain with densities $p, q$, given by:

$$d(p, q) = \int p(x) \log \frac{p(x)}{q(x)}$$

so $-D_i$ is the divergence of $p_{ST}(s,t)$ from the independent case $p_S(s) p_T(t)$ (equation 2.5) and thus measures the independence of image $S$ from $T$, and so mutual information measures the *dependence* of $S$ on $T$ and gives higher values when $S$ is more similar to $T$.

Maes et al. [1997] show that mutual information can be extended further to more general *f-information* measures.

**Variance in intensity ratios**

Woods et al. [1992] use an alternative similarity measure: the variance of the pixel-wise intensity ratios
between the two images. The assumption is that these ratios will have a constant value across the image domain if the image pair is in register, although in practice there will be some slight variation due to noise, shape differences in the images and partial volume effects\(^I\). When the image pair is mis-registered, the ratios will vary more markedly across the image domain and the variance of their distribution will give an indication of the quality of registration.

The measure is computed as follows:

- for every voxel \(i\), the ratio between source \(S\) and target \(T\) is found:

\[
 r_i = \frac{S(i)}{T(i)}
\]

This gives a ratio volume \(R_v = \{r_i\}\)

- all \(r_i\) where \(S(i)\) is less than a threshold (empirically set at 21.5 \% of the maximum value in \(S\)) are re-set to zero to give a masked ratio volume.

- a measure of the variance in the distribution of \(R_v\) is calculated as

\[
\frac{\sigma_r}{\bar{r}_{\text{mean}}}
\]

where \(\sigma_r\) is the standard deviation of \(R_v\) and \(\bar{r}_{\text{mean}}\) is the mean of remaining non-zero voxels in \(R\)

- the above calculations are repeated with the roles of \(S\) and \(T\) reversed, giving

\[
\frac{\sigma_{r'}}{\bar{r'}_{\text{mean}}}
\]

which is then averaged with

\[
\frac{\sigma_r}{\bar{r}_{\text{mean}}}
\]

to give a final variance measure

\[
\frac{\sigma_{\text{brain}}}{\bar{r}_{\text{brain}}}
\]

In this way there is no bias from assignment of one image as source and the other as target in the variance measure.

Having described above a selection of pixel intensity similarity measures, we will now elaborate on their use in linear registration algorithms.

### 2.5.2 Cross-correlation variants for adaptation to the inter-modality problem

We recall the problem of intra-subject inter-modality matching such as CT to MR: although the image sets represent the same underlying structure, there is a differential in how the various tissue types respond to the imaging

\(^I\) one pixel representing intensities from two or more neighbouring tissues.
2.5. Pixel Matching Methods

device (section 1.2). CT shows a high contrast between hard tissue (high intensity) and soft tissue/water/air (low intensity or background), whereas in MR this is roughly inverted, with bone, air and CSF** taking low or background intensities, and soft tissues responding with highest intensities.

We noted in section 2.4.5 that feature-matching techniques are readily adapted to the inter-modality problem: once corresponding features have been segmented, the registration proceeds as for the intra-modality case. A little more thought is necessary for the adaptation of cross-correlation-type matching for the inter-modality case.

Three generations of algorithm address the problem with solutions of increasing sophistication. The simplest and most straightforward is to apply a pixel intensity mapping to the CT image to make it appear more like an MR image, prior to progressing with a regular cross-correlation optimisation search. This is the pre-processing approach. The second generation applies cross-correlation to the unaltered image but includes some grouping strategy within the calculation of correlation measures - this is effectively considering each tissue type as an image in itself. We discuss these under the heading grouping during measure. Finally, using the mutual information measure instead of simple cross-correlation, the original technique may be applied directly without any pre-processing or segmentation.

2.5.3 Pre-processing

The pre-processing method aims to convert the inter-modality image pair to one imitating the intra-modality case before commencing registration. Recall the optional contrast mapping \(\phi(T(x))\) in equation 1.1. The aim is to map the intensities of each tissue type of one image to the intensities that would have been produced by the imaging modality that produced the other image - that is, to enforce the trend similarity. The (linear) maps use prior knowledge or assumptions of the intensities manifested by corresponding tissue types in each modality. Richardson and Bury [1996] preprocess the CT image by applying a threshold to force low-intensity pixels to take on background intensity (zero), and then inverting the entire greyscale range. Low intensity values in CT are from soft tissue, high intensities from bone. By removing the soft-tissue shadow and inverting the bone to black and remainder to high intensity, the cross-correlation will now register optimally bone to bone which are now black in both images. This is an over-simplified approach since the MR is also zero-intensity for non-bone regions such as air, and we can expect a high frequency of misregistrations.

van den Elsen et al. [1994] choose two intensity maps with higher selectivity: the first, a delayed ramp, enhances the differences between CT and MR, and the registration then aims to minimise the correlation measure. The alternative (the triangle map) enforces a segmentation on the CT image by varying the mapping depending on the original intensity; low intensity pixels representing background are thresholded to zero, the lower-intermediate intensities (soft tissue) are increasingly enhanced with a rising ramp, upper-intermediate intensities (soft bone) are dampened with a decreasing ramp, and the highest intensities (bone) are thresholded to zero. The cut-off points for each intensity grouping are pre-determined

**depending on image acquisition strategy
empirically and fixed. One drawback will be that if the image suffers from inhomogeneities induced by
a varying field in the scanner, they may be seriously enhanced if they are in regions close to the cut-off
points, with false structures resulting. The algorithm using the first mapping also requires a few manually-
placed landmark pairs (their positioning need not be exact) to guide the registration to avoid misregistrations.
It is suggested that landmarks be used likewise in the triangle map since this suffers misregistration
when mapping incomplete volumes (such as MR brain to CT whole head). The search progresses up a
multiresolution pyramid, with region size for parameter values shrinking as resolution increases, reflect-
ing a higher confidence level in the estimated parameters produced by the previous level. Several options
may be followed for the value of each parameter, each having its own route up the pyramid, until on ter-
mination the parameter value set giving lowest correlation value is determined to be the best fit.

2.5.4 Grouping during measure

In the second method of dealing with inter-modality cases, consideration of the varying tissue contrasts is
delayed until calculations of similarity measures during the registration itself. Woods et al. [1993] adapt
an earlier method [Woods et al., 1992] using variance in intensity ratios (section 2.5.1) for inter-modality
applications by individually considering the match of each tissue type manifested in the more varied of
the two images to its counterpart in the other image. The assumption is that for images in register, the
intensity ratios are constant within each tissue type. In effect, prior to calculating the variances of the
ratios, the intensity ratios of pixels at the same location in both images are binned according to the inten-
sity in the more complex modality. The variance of each bin is computed and then finally all variances
are averaged, giving a global estimate of the goodness of registration. Registration is applied to edited
brain images (scalp removed) to avoid misregistration. Convergence is not guaranteed: oscillations be-
tween two parameter values are blocked on detection by averaging the two values; longer loops cannot
be resolved, in which case the registration exits with failure.

Instead of averaging all correlation variances, an alternative is to predetermine which tissue types (in the
complex image) are most influential in directing the registration, and to compute the variances only for
these intensities. Hill et al. [1993] thus utilise only certain regions in the joint histogram of intensity ratios.

The ideal answer to the inter-modality problem is to use a dissimilarity measure which is optimised at the
correct registration regardless of the inter-modality difference in tissue contrasts. One such measure is
that of mutual information, whose use we describe next.

2.5.5 Mutual Information

Mutual information can be applied directly to the inter-modality image pair without preprocessing or in-
terference during measurements, since it is a measure of statistical dependence rather than actual inten-
sity differences. Furthermore it shows greater robustness with respect to occlusions and incomplete data.
Maes et al. [1997] show that the algorithms in both [Woods et al., 1993] and [Hill et al., 1993] mentioned
in the previous section are restrictions of a full mutual-information implementation. They note that the
2.6 Directed search based on optical flow

The optical flow [Horn and Schunk, 1981] method registers source to target by slowly updating its transformation according to intensity gradients in the target.

The optical flow technique is usually used for object tracking in video sequences. However, for two identical images slightly out of register, it can be applied to find the affine transformation to map one to the other. We follow here the method of Barber et al. [1995].

The source \( S \) and target \( T \) are assumed to be copies of the same image \( g \), the source having undergone a small displacement \((\Delta x, \Delta y, \Delta z)\). The target and original are identical:

\[
T(x, y, z) = g(x, y, z, t = t_0)
\]

(2.10)

while the source is the original after a small time interval \( \Delta t \):

\[
S(x, y, z) = g(x, y, z, t + \Delta t)
\]

(2.11)

The source is an affine transformation of the target,

\[
T(x, y, z) = g(x + \Delta x, y + \Delta y, z + \Delta z, t + \Delta t)
\]

(2.12)
where the displacement vector is given by

\[
\begin{pmatrix}
  a_{11} + 1 & a_{12} & a_{13} & a_{14} \\
  a_{21} & a_{22} + 1 & a_{23} & a_{24} \\
  a_{31} & a_{32} & a_{33} + 1 & a_{34} \\
  0 & 0 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
  x \\
  y \\
  z \\
  1
\end{pmatrix} =
\begin{pmatrix}
  z + \Delta x \\
  y + \Delta y \\
  z + \Delta z \\
  1
\end{pmatrix}
\]  

(2.13)
equating (2.10), (2.11) and (2.12) and Taylor expanding about time \( t = t_0 \) we have:

\[
T(x, y, z) \overset{(2.10)}{=} g(x, y, z, t = t_0)
\]

\[
T(x + \Delta x, y + \Delta y, z + \Delta z, t + \Delta t) \overset{(2.12)}{=} g(x, y, z, t = t_0) +
\Delta z \left( \left. \frac{\partial g}{\partial x} \right|_{t=t_0} \right) + \Delta y \left( \left. \frac{\partial g}{\partial y} \right|_{t=t_0} \right) + \Delta z \left( \left. \frac{\partial g}{\partial z} \right|_{t=t_0} \right) + \Delta t \left( \left. \frac{\partial g}{\partial t} \right|_{t=t_0} \right)
\]

(2.14)

Since \( S \) and \( T \) are copies of \( g \) displaced over time \( \Delta t \), we have the additional approximation

\[
\left. \frac{\partial g}{\partial t} \right|_{t=t_0} = \left. \frac{\partial T}{\partial t} \right|_{t=t_0}
\]

and hence (2.14) becomes

\[
0 = \Delta x \left. \frac{\partial T}{\partial x} \right|_{t=t_0} + \Delta y \left. \frac{\partial T}{\partial y} \right|_{t=t_0} + \Delta z \left. \frac{\partial T}{\partial z} \right|_{t=t_0} + \Delta t \left. \frac{\partial T}{\partial t} \right|_{t=t_0}
\]

(2.15)

Using the affine transformation in (2.13) gives for (2.15):

\[
T(x, y, z) = S(x, y, z) + (a_{11}x + a_{12}y + a_{13}z + a_{14}) \left. \frac{\partial T}{\partial x} \right|_{t=t_0} + (a_{21}x + a_{22}y + a_{23}z + a_{24}) \left. \frac{\partial T}{\partial y} \right|_{t=t_0} + (a_{31}x + a_{32}y + a_{33}z + a_{34}) \left. \frac{\partial T}{\partial z} \right|_{t=t_0}
\]

(2.16)

where the coefficients \( a_{ij} \) in the inverse transformation of equation (2.13) will map the source back to the target. Since approximations have been made in the Taylor expansions and the derivatives, (2.16) is not exact and so must be solved iteratively: in each iteration, twelve voxels \( (x, y, z) \) from the target and current source image are used in (2.16) to compute transformation coefficients \( a_{ij} \). These are used to transform the source to a new position and the process repeated until the true match has been found (when the \( a_{ij} \) are all zero). To avoid blurring of the source by successive resampling, the transformation matrix with the new coefficients is concatenated to the previous ones and the total transformation is applied to a copy of the original source image.

Application of the algorithm is limited to images where only a small displacement is required to bring them into register. Large displacements render the Taylor expansion inaccurate and then convergence is not guaranteed. Additionally, it is assumed the images are noise free, else the erroneous values of \( S(x, y, z) \) will knock the registration off course and the gradients will be incorrect. In both these respects, a prior smoothing of the images improves convergence as it reduces noise and lowers the higher derivatives (which have been ignored in the Taylor expansion).
The optical flow registration is the linear analogue of the fluid method which will be described in the following chapter on non-linear methods. Before proceeding to fully non-linear methods we first mention algorithms which use local linear transformations to approximate a non-linear deformation.

### 2.7 Piecewise-linear methods

Piecewise-linear registrations were defined in definition 1.15. They involve subdividing the images into small regions for which independent linear transformations are computed to register them individually. Constraints are applied to ensure that the transformations of neighbouring regions are sufficiently similar to ensure some degree of continuity, which is further enforced by applying some form of interpolant between them. An example is the algorithm of Collins et al. [1994], used for registration of a population of brains into a common coordinate system to create an ‘average’ brain. The coordinate system chosen is that proposed by Talairach and Tournoux [1988]. Each brain volume is subdivided into twelve sub-volumes, each of which is linearly registered into the corresponding sub-volume of the standard brain. The cross-correlation dissimilarity measure is minimised for each subvolume pair, computed from either corresponding pixel intensities or corresponding pixel gradients and then cubic interpolation is used to smooth the deformation field. Piecewise-linear registrations give the appearance of a non-linear deformation.

### 2.8 Summary

This chapter surveyed linear registration methods, classifying them according to the feature spaces used on which distance measures are to be minimised. The three main groups of feature spaces studied were: principle axes; landmarks and extracted structures such as surfaces; and pixel intensity pairs. In addition, the optical flow method was included as an interesting linear registration technique.

In the next chapter we discuss single-level non-linear methods, before proceeding to multi-level or hierarchical non-linear methods in chapter 4.

Further discussions and comparisons of linear registration algorithms can be found in West et al. [1997]; Banerjee et al. [1995]; Ayache [1995] and van den Elsen et al. [1993].
Chapter 3

Single Level Non-linear Registration Methods
3.1 Introduction

This chapter reviews registration algorithms which compute a non-linear transformation in order to register the source to the target. In the previous chapter, our discussion was centred on the adequate selection and extraction of features and the identification of their correspondences, and of the application of the algorithms to the inter-modality case. These considerations are equally relevant to the non-linear case. Non-linear algorithms however are distinguished by their greater complexity of search space, or range of allowable transformations. Of particular interest is the selection and adaptation of a model through which the deformations are conceptualised and implemented.

For this reason, we identify the realisation of a non-linear registration algorithm as consisting of three stages. First, a physical model is proposed which describes the ideal deformation of the target image. Typically the target is modelled as a deformable substance with properties which ensure the deformation is well behaved, most importantly that the material will not tear or fold. Alternatively, the model may be a statistical concept, describing a method for finding the most likely deformation to give a good match. Next, this is formulated as a mathematical model, in which the proposed transformations are described as equations and the values of parameters are identified. Finally, the model is translated into the practical implementation, which may only be an approximate realisation of the physical or mathematical models.

The practical implementation may be constrained by data structures (eg solution on a finite pixel lattice), memory or time constraints, and is often affected by inaccuracies introduced by sampling and truncations.

We work through these stages in reverse order, commencing with the practical implementation followed by the derivation of the mathematical model, and finally review one statistical and three physical models. Hence the chapter is divided into seven sections. The first two concern the description of a deformation and its practical implementation in respect to the regularly sampled pixel grids of the image pair. Section 3.4 describes various approaches to deriving the mathematical model. These are regularisation theory, Bayesian statistics, and the balancing of forces in a physical model.

The last four sections each analyse a physical or statistical model on which the registrations are based. These are:

- landmark matching with spline interpolation;
- statistical truncated basis function expansion methods;
- elastic registration based on a physical model;
- fluid registration.

These roughly correspond with the second, third and fourth groups identified in the linear registration methods on page 44: the thin-plate spline matches landmarks, 'elastic' registration in its basis functions and full deformation field versions can be implemented by minimising pairwise pixel intensity measures, and fluid registration is similar to optical flow techniques in that registration takes place over time by
3.2. Representations of deformations

Linear registration calculates an affine transformation that is applied globally to the image pixel locations. Non-linear registration finds deformations, where the mappings of the pixel locations vary over the image. A deformation may be described in the spatial domain as a full discrete deformation field on the pixel lattice, or as its dual in the spectral domain, as coefficients of some function basis. Alternatively, a combination of these two types of descriptors is used, as a sparse deformation set of nodes or landmarks, together with coefficients of interpolating functions to summarise the internodal deformations.

3.2.1 Deformation field vs. basis function expansions

We assume the target $T$ and source $S$ are related by a specific case of equation (1.1), with $D = S(u(\vec{x})) - \phi(T(\vec{x}))$:

$$S(u(\vec{x})) = \phi(T(\vec{x}))$$ (3.1)

and further we may assume that $\phi(.)$ is the identity function.

The deformation $u(\vec{x})$ may be implemented either as a displacement field or as an analytic function [Amit, 1994].

In the first case, $u(\vec{x})$ gives a displacement for every pixel in the lattice,

$$u(\vec{x}) = (u_1(\vec{x}), u_2(\vec{x}), u_3(\vec{x})) \quad \text{(in the 3D case)}$$

A full deformation field allows for the displacement of each pixel to be relatively independent. Examples of its use are found in versions of the elastic and fluid models when solved using finite differences (section A.4).

In the second case, the deformation $u(\vec{x})$ is expanded as

$$u(\vec{x}) = \left( \sum_k u_{1k} \beta_k(x_1), \sum_k u_{2k} \beta_k(x_2), \sum_k u_{3k} \beta_k(x_3) \right)$$ (3.2)

where $u_{1k}, u_{2k}$ and $u_{3k}$ are coefficients (to be found) and $\beta_k$ are a set of (orthonormal) basis functions.

Typically they are the Fourier or some other trigonometric basis (eg the cosine transform); alternatively they may be a wavelet basis [Amit, 1994] which has the advantage of producing more localised deformations, wavelets having local support. Ideally the sum is infinite ($0 < k < \infty$) allowing for all frequencies in the deformation; in practice the basis is truncated so that only the first $N$ of the lowest frequencies are used ($0 < k < N$).

The basis function expansion of the mapping can be considered a dual form of the displacement field, the latter being the a set of displacements $\{(u_1(x), u_2(x), u_3(x))\}$ in the spatial domain and the former a set...
of basis function coefficients \( \{u_1, u_2, u_3\} \) in the spectral domain.

### 3.2.2 Landmarks with interpolating functions

There are two types of the combination representation, namely landmark-based splines such as the thin-plate spline (TPS), and finite elements. Both consist of a sparse set of points (nodes or landmarks) for which the true deformation is specified, together with coefficients for a set of interpolating functions from which an approximation to the deformations at internodal points can be calculated.

**Thin-plate spline**

In the case of the thin-plate spline, the landmarks may be scattered in any formation over the image, although the spline approximation is only valid within the region bounded by landmarks (since it is an interpolant, not an extrapolant). Pairs of inter-landmark distances are given as arguments to a set of radial basis functions, from which a matrix is constructed to describe the initial landmark configuration. When inverted and applied to the final landmark positions (as defined by the sparse deformation field), the result is a set of coefficients to the interpolating functions which are applied to the entire image to deform it. Section 3.5.1 contains a fuller treatment of thin-plate spline deformations.

**Finite elements**

The target image is tesselated into elements (regions) bounded by nodal points, with neighbouring elements sharing nodes. Nodal displacements are solved using for example the elastic PDE (equation 3.66), while internal displacements (for pixels within the elements) are found by interpolation of nodal displacements using isoparametric shape functions. Practical finite element theory using isoparametric shape functions is explained well by Hinton and Owen [1977]. The elements are 4-sided with parabolic curved sides. This allows originally square-shaped regions to be mapped to deformed regions with the shape functions as the interpolator, reducing the effect of aliasing (whether the mapping is performed from source to target or in reverse). Aliasing is the next topic for discussion.

### 3.3 Pixel/voxel intensity interpolation and resampling

This section relates a computed deformation to its practical implementation within the regularly sampled pixel grids of the source and target image pair. The theory it summarises is also necessary for the implementation of linear algorithms, but we discuss it here since we additionally mention points of particular relevance to the non-linear case.

When applying a computed transformation to a source image, pixels are generally mapped into interstitial locations, whereas the required output transformed image must be stored and displayed on a regular pixel grid*. Resampling is necessary to deduce intensity values at the lattice points of the transformed image from the values at the interstitial locations. This is a particular case of interpolation. Interpolation is the

---

*We are not discussing finite element representations here.*
estimation of values at locations within the convex hull of a set of given data values, usually using a model of an underlying continuous function. When mapping the source image, one might think the underlying continuous function is the particular anatomy of the scanned subject. In reality however, the underlying function depends on the imaging modality and is an intermediary between the true physical anatomy and the pixel grid. The reader is referred to the descriptions of image acquisition in section 1.2. The image samples data collected by the scanner, which may be counts of isotopes (PET), estimated tissue densities in small regions (CT) or various density and contrast information reconstructed from the Fourier into the spatial domains (MR). In the latter case we speak of the point spread function which is the shape modelling the intensity surface of individual points in the reconstructed spatial domain.

Hou and Andrews [1978] identify interpolation as a theoretically two step process: first, building the underlying model, and second, re-sampling the model at the required locations. For interpolation at the pixel-level, these two steps are generally combined into one operation, with a pre-determined underlying model assumed before the process starts.

Interpolation comes in two classes for non-linear registration; we shall call these micro-level (or pixel-level) and macro-level interpolation.

Macro-level interpolation is an essential part of non-linear registration methods involving matching of landmarks or segmented structures. These features, for which a match is determined, populate the images sparsely; the intervening regions however also require a mapping from one image to the other. This is the macro-level interpolation, determining transformation vectors in unknown regions using the estimated maps at the landmarks. Here too since the true mapping is unknown, a model must be used, such as one of those introduced in section 3.2.2.

Micro-level interpolation is that applied at the pixel level to deduce intensity values at grid points from those computed at interstitial locations. It utilises one of a few models, of varying non-linearity, which we list in the following subsection.

So far we have considered applying a transformation to the undeformed source and resampling intensities from interstitial locations into pixel locations. In section 3.3.2 we will consider the reverse mapping, called back-sampling. This applies the inverse of the transformation to pixels in the deformed source, resamples pixel intensity values into interstitial locations in the undeformed source and paints these into the corresponding pixel locations in the deformed source domain. Since it is more common to apply back-sampling, we will discuss the interpolation models in this context.

3.3.1 Common interpolation models for pixel-level / micro interpolation

There are a few common models used for pixel-level interpolation. The choice of model depends on a compromise between speed and simplicity, and accuracy. For simplicity we will illustrate the shapes of the one-dimensional interpolants only. The two- and three-dimensional versions are obtained by repeated application of the 1-dimensional interpolation in each of the $x$, $y$ (and $z$) directions (the functions being
3.3. Pixel/voxel intensity interpolation and resampling

separable). The interpolation models are functions defined over a local coordinate system \( \{ \tilde{x} \} \) whose origin is placed over the required interstitial location. The heights of the interpolation model at the surrounding pixel locations determine the weights applied to their intensity. The weighted neighbouring pixel intensities are then summed to give the intensity for the interstitial location.

Nearest neighbour

The nearest neighbour model assigns to the interstitial pixel the value of the nearest grid pixel. This is the fastest algorithm - no computation is needed (other than finding the nearest pixel). No other pixel has any affect. If the interstitial pixel is on or close to a boundary between regions, the assignment may be inappropriate, as illustrated in figure 3.1.

![Figure 3.1: Nearest neighbour resampling at a boundary region.](image)

Linear interpolation

For 3D interpolation, there are eight closest grid pixels which contribute to the value of the interstitial pixel, and the weights are the inverse distances from grid to interstitial locations. Figure 3.2 illustrates the linear interpolant in the 1-dimensional case.

![Figure 3.2: Linear interpolation function](image)
3.3. Pixel/voxel intensity interpolation and resampling

Cubic interpolation

The one-dimensional cubic spline used for pixel interpolation weights the intensities of four neighbouring pixels and is given by Keys [1981]

\[ f(x) = \begin{cases} 
  f_2 = \frac{3x^3 - 5x^2}{2} + 1 & 0 \leq x < 1 \\
  f_1 = \frac{5x^2 - 2x^3}{2} + 2 & 1 \leq x < 2 \\
  0 & 2 \leq x 
\end{cases} \quad (3.3) \]

with symmetry about the origin, where \( \{x\} \) is the coordinate system local to the interpolation model. The function is shown in figure 3.3. Note that the function is zero outside the region of support. The bicubic interpolant gives an approximation to the sinc interpolant, which we describe next.

Sinc interpolation

\[ \text{sinc}(x) = \frac{\sin(\pi x)}{\pi x} \]

It has infinite support - the oscillations about the x-axis reduce in amplitude but continue indefinitely in both directions (to \( x = \pm\infty \)). As a practical interpolant, a truncated version must be used, using a neighbourhood of finite size. The full sinc function in the spatial domain corresponds to the pulse or rect function in the Fourier domain (figure 3.4, right). Hence if a rect function is used to sample a continuous function in the frequency domain (for example by the MR scanner) then the original function can be completely re-constructed in the spatial domain by convolution of the sample points with a sinc function.
3.3. Pixel/voxel intensity interpolation and resampling

Reconstruction by a truncated sinc will introduce error into the reconstruction since although its amplitude is low at any single point at a large distance $|x|$ from its centre, the sum of its energy over all $|x|$ is substantial, and part of this has been lost in the truncated version. The Fourier transform of a truncated sinc, figure 3.5, displays high frequency oscillations around the cutoff point. The result is that some frequencies beyond the cutoff point are included, and frequencies within the cutoff are either amplified or attenuated [Foley et al., 1992, p634], [and see also Jain, 1989].

![Cutoff](image)

Figure 3.5: Fourier transform of truncated sinc, an approximation to the pulse/rect.

A solution is to multiply the sinc by an *appodising*\(^1\) function, one that takes the value of 1 at zero and decreases smoothly to zero at the edge of the required neighbourhood (figure 3.6). Magnetic resonance images are acquired in the Fourier domain, and their frequency content is strictly band-limited, hence the sinc is the best in-plane interpolant, matching the point-spread function in the spatial domain. For 2.5D images, that is, data sets consisting of a series of 2D slices, the in-plane resolution exceeds that of the between-plane resolution and the point-spread function is anisotropic; the sinc function is still the best in-plane interpolant, but in the z-direction there is no advantage in using the sinc over cheaper interpolants\(^2\) [Hajnal et al., 1995b]. It is suggested [Hill et al., 1994] that the images be extrapolated before interpolation to avoid having a sharp edge at their boundaries which will cause artifacts in the Fourier domain\(^3\).

Lemieux et al [1998] found sinc interpolation gave a reduction of $< 1\%$ in the structure of difference images compared to that of linear interpolation.

![Appodised sinc](image)

Figure 3.6: Appodised sinc

---

\(^1\) 'cutting-off feet'

\(^2\) For true 3D images, the appodised 3D sinc is the best interpolant.

\(^3\) Since sharp edges are of high frequency
B-spline interpolation

The cubic spline is similar to the sinc in shape, without the infinitely oscillating tails. Like the sinc, it may introduce negative values into an interpolated image. The cubic B-spline [Unser et al., 1993a, 1991], however, is everywhere positive. The nearest neighbour, linear and cubic B-spline interpolation functions are a subset of a family of functions produced by repeated convolutions of the rectangle or 'sample-and-hold' function, figure 3.7. Jain [1989] refers to the nearest neighbour, or rectangle function, as the ‘zero-order hold’, the linear (triangle) as the ‘first-order hold’ and the splines as ‘nth-order hold'. As the order

\[
\begin{align*}
\text{rectangle / nearest neighbour} & : R_0(x) \\
\text{triangle / linear} & : R_0(x) \ast R_0(x) = R_1(x) \\
\text{quadratic} & : R_0(x) \ast R_1(x) = R_2(x) \\
\text{cubic} & : R_0(x) \ast R_2(x) = R_3(x)
\end{align*}
\]

Figure 3.7: First 3 convolutions of rectangle, giving the nearest-neighbour, linear, quadratic and cubic interpolation functions. Circles indicate pixel locations.

\(n\) of the interpolation function (ie the number \(n\) of convolutions of the rectangle in its construction) is increased, so are the number of pixels in its support. Note the similarity in shape of the higher order B-splines to a Gaussian; repeated convolution by the linear interpolant due to successive interpolation gives a blurring effect similar to a Gaussian smoothing. All the functions in the \(n\textsuperscript{th}\)-order hold family have only positive values within the region of support. In contrast, the cubic and sinc have zero crossings at regularly-spaced intervals on either side of the central peak (whose height = 1), giving both positive and negative weights, and so when centred on the given pixel return that pixel’s value exactly.

Comparisons of the various interpolants for micro-interpolation are given in [Parker et al., 1983; Tian and Huhns, 1986; Jain, 1989]

### 3.3.2 Interpolation for non-linear deformations

If the transformation of the source image is linear, the interpolation functions given above are sufficient and in the theoretical case of the full sinc function can give error-free resampling. This is of especial relevance to applications such as functional imaging, where the differences in intensity due to activity are only slight. Any artefacts introduced by interpolation error will be mistaken for functional activity within
3.3. **Pixel/voxel intensity interpolation and resampling**

the brain and so must be avoided. (See for example Bydder [1995]; Hajnal et al. [1995a]). Using the appodised sinc function, sub-voxel accuracy of registration is possible using a suitable linear registration algorithm, [Hill et al., 1994; Hajnal et al., 1995b].

Non-linear registration generates two problems for interpolation.

- Pixels are mapped to irregularly spaced locations; these are sparse within the pixel lattice if there have been large increases in region areas.
- The underlying continuous function, representing the data sampled onto the original pixel lattice, has been deformed.

The first is also a problem in linear registration if an area enlargement or shrinkage has occurred. Consider the forwards mapping of the source domain into the deformed source domain. During enlargement, neighbouring lattice pixels map to positions separated by more than the lattice spacing, so there will be regions in the transformed source to which no pixels have mapped. Interpolation in these regions will therefore return holes, areas of zero or low intensity (hypointense). Conversely, where area shrinkage has occurred, hyperintense regions will result. A back-sampling strategy is almost always employed to reduce this problem (figure 3.8); we mentioned this option on page 68. For back-sampling, the inverse transformation is computed which gives the mapping from the space of the deformed source into the original undeformed source. The original location of each lattice point of the deformed source is identified and the resampling is done in the un-deformed image. This will avoid larger holes in the deformed source. However if there has been severe compression or enlargement in some areas, the hyper/hypointensity problem will be reversed (and thus remain as a problem) since the inverse map will experience enlargement or compression respectively. Backwards re-sampling retains (or approximates) original intensity density within a region, whereas forwards re-sampling retains (or approximates) the sum total of original intensity within the region.

The second interpolation problem is that the underlying continuous function has been deformed. The interpolation function, ideally being a model of the underlying function, should be distorted likewise; the

![Figure 3.8: 'Backwards' interpolation. (right) The transformed source image is constructed by scanning the output lattice and mapping each pixel to the position from which it originated in the pre-transformed source. This is generally an interstitial location (centre). Its intensity value is obtained by a weighted sum of the values of neighbouring lattice pixels in the undeformed source (left). The weights and choice of neighbours is determined by the interpolation model.](image)
undistorted sinc may no longer be the best approximation to the (distorted) point spread function. Figure 3.9 illustrates the problem. The left-hand column shows some original underlying continuous function $f(x)$ (dotted) with the sinc interpolant overlaid (top row). For a linear transformation of the underlying function, the interpolant is shifted (second row) and used to resample the underlying function (third row), giving the final result (bottom row). The right-hand column gives case of a non-linear deformation. The top row shows the original underlying function (dotted) and the sinc interpolant as before, and also the deformed continuous function $f(u(x))$. The second row shows the sinc interpolant and the sampled $f(u(x))$; the third and fourth rows estimate the application of the sinc to the samples taken from $f(u(x))$ - the result bears hardly any similarity to the true $f(u(x))$. Instead of the sinc function, the required interpolant is the deformed sinc, $\text{sinc}(u(x))$ for forwards interpolation or its inverse for the backwards mapping. In practice this is unknown or impractical to apply.
SINC INTERPOLATION FOR DEFORMED SIGNAL \( f(\phi(x)) \)

Figure 3.9: Sinc interpolation after a non-linear deformation
3.4 Registration as an optimisation problem

Registration is an optimisation problem: we require a transformation to maximise the similarity between the image pair, while constraining the transformation to be smooth. Formally, we adapt the problem (1.1) to:

\[
\text{minimise } J(\bar{u}) = \mathcal{D}(S(\bar{u}), T) + \lambda R(\bar{u})
\]

where \( J(\bar{u}) \) is the cost function, \( \mathcal{D}(S(\bar{u}), T) \) is some dissimilarity measure between the target image \( T \) and the deformed source \( S(\bar{u}) \), and \( R(\bar{u}) \) is a measure of the roughness of the deformation field \( \bar{u} \).

The mathematical model (3.4) is derived using one of three standard theories (figure 3.10): regularisation, physical models of continuous media, or Bayesian statistics (closely related to decision theory). Regularisation theory gives a tradeoff between similarity measures and smoothness constraints to derive a cost function to be minimised. Physical models of continuous media balance external forces and internal stresses to derive a partial differential equation (PDE) to be solved. Bayesian statistics incorporates prior and likelihood probability distributions into a posterior distribution for which the highest probability is to be obtained. When Gaussian distributions are used, the statistical and physical models have links through statistical physics via Gibbs distributions and the Bayesian approach may be converted to the regularisation format by taking logs of the probability density functions. The connection between regularisation theory, Bayesian theory and physical models is given by Gee and Peralta [1995]. We give a summary of the three approaches in the following sections.

We note also that there may be an additional set of constraints implicit in the implementation of the algorithm such as approximations of the deformation using a truncated set of basis functions and premature termination due to time constraints. These prevent a true registration from being achieved.
3.4. Registration as an optimisation problem

3.4.1 Regularisation

Regularisation is the process of optimising an expression containing two or more terms whose relative contribution is governed by a hyperparameter(s) (also called the regularisation parameter) $\lambda$, as in

$$J = A + \lambda B$$  \hspace{1cm} (3.5)

where $A$ is an expression often depending on observed data while $B$ is a functional inserted into the problem to give stable solutions. $\lambda$ itself may be a variable quantity, in which case it is a ‘weak’ hyperparameter. The theory of regularisation is well developed; for a good introduction the reader is referred to Press et al. [1995]. Non-linear registration may be cast as such an optimisation problem by assigning $B$ to be a smoothness constraint and $A$ as a similarity measure, such as those introduced in section 2.5.1, giving equation (3.4). Regularisation then gives a tradeoff between maximising the similarity in the image pair and the smoothness of the deformation.

The transformation must be 1-1 and onto (that is, homomorphic) and continuous, so that the relative positions of structures in the deforming image are respected and no tears or folds are introduced. Such constraints are imposed by minimising some second-order derivative of the deformation, representing curvature or bending and stretching; examples are the Laplacian $\nabla^2$, the square of the Laplacian $\nabla^4$ or other differential operators derived from physical models. Where the deformation is represented as basis function expansions, smoothness is automatically guaranteed by limiting the expansion to the first few functions in the basis, which are of a lower frequency (see for example section 3.6).

3.4.2 Physical Models

A physical model of a continuous medium may be employed in deriving the smoothness constraints, or in modelling the entire registration process as the solution of a partial differential equation describing the behaviour of the physical material. In this case the regularisation achieves a balance between external and internal forces. The external forces are those which drive the registration to match corresponding features. Smoothing constraints are provided by internal forces or stresses maintaining the cohesiveness of the material.

driving forces

‘Forces’ driving the registration are derived from some distance function between features in the image pair. These may be pre-segmented landmarks, curves or surfaces, in which case the registration procedure aims to match these exactly or in a least squares sense, with the remaining image areas deformed using some smooth interpolant (macro-level interpolation, section 3.3). Alternatively, the driving forces are

\footnote{It may be of interest to refer to [Rudin et al., 1992; Alvarez et al., 1992; Evans and Spruck, 1991] who discuss $L_1$ and $L_2$ norms in the context of algorithms for noise removal. There the problem is to reduce noise while retaining edge strengths. The $L_1$ norm uses information from the gradient $\nabla$, rather than (gradient)$^2$. The constrained optimisation problem of

$$\text{minimize } \int_{\Omega} \sqrt{u_x^2 + u_y^2} \, dx \, dy$$  \hspace{1cm} (3.6)$$

in 2D is the $L_1$ norm of the gradient, being the size of $(\nabla x, \nabla y)$.}
3.4. Registration as an optimisation problem

derived from some statistical measure of difference between the full set of corresponding pixels in the
image pair (section 2.5.1), usually the sum of squared differences $D_d = \sum (S(\bar{u}(\bar{x})) - T(\bar{x}))^2$, although
as in linear registration, cross-correlation and mutual information are other possibilities. Using standard
gradient-based optimisation methods (see for example Press et al. [1995]), the forces are proportional
to the derivative of this function, which is a product of the difference image and gradients in the source
$(S(\bar{u}(\bar{x})) - T(\bar{x})).\n\ninternal cohesive forces

The most frequently used physical models providing smoothness constraints are the thin pliable sheet,
or thin-plate (section 3.5.1 and equation 3.13) and elastic media (section 3.7 and equation 3.66). The
thin-plate deformation minimises bending energy, given by Rohr et al. [1996]
\[
\sum_{\alpha_1+\ldots+\alpha_d=2}^2 \alpha_1! \cdots \alpha_d! \int_{\mathbb{R}^d} \left( \frac{\partial^2 \bar{u}}{\partial x_{\alpha_1} \cdots \partial x_{\alpha_d}} \right)^2 d\bar{x}
\]
(3.7)
where $d$ is the image dimension, ie $d = 2$ or $d = 3$.

Elastic media give an elastic smoothness energy cost term obtained from sums of extensional and shear
stresses which for 3D deformations is [Miller et al., 1993]
\[
\sum_{i=1}^3 \sum_{j=1}^3 \int_{\Omega} \left[ \lambda \left( \frac{\partial u_i(\bar{z})}{\partial x_i} \right) \left( \frac{\partial u_j(\bar{z})}{\partial x_j} \right) + \mu \left( \frac{\partial u_i(\bar{z})}{\partial x_j} + \frac{\partial u_j(\bar{z})}{\partial x_i} \right) \right]^2 d\bar{x}
\]
$\lambda$ and $\mu$ are Lamé elasticity constants and are derived in continuum mechanics theory; we include an
outline of the derivation in an appendix A.2. In the 2D version, there are no $x_3$-direction components of
extensional stress, nor shear stresses in the $(x_1, x_3)$ and $(x_2, x_3)$ planes, so
\[
\begin{align*}
\frac{\partial u_3}{\partial x_1} + \frac{\partial u_1}{\partial x_3} &= 0 \\
\frac{\partial u_3}{\partial x_2} + \frac{\partial u_2}{\partial x_3} &= 0
\end{align*}
\]
and also
\[
\frac{\partial u_3}{\partial x_3} = -\frac{\lambda}{\lambda + 2\mu} \left( \frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} \right)
\]
Christensen [1994] gives a linearised version where the cross-terms in the shear stresses are ignored:
\[
\frac{\lambda}{2} \left( \sum_{i=1}^3 \frac{\partial u_i}{\partial x_i} \right)^2 + \frac{\mu}{4} \sum_{i=1}^3 \sum_{j=1}^3 \left( \frac{\partial u_i(\bar{z})}{\partial x_j} + \frac{\partial u_j(\bar{z})}{\partial x_i} \right)^2
\]
These models will be analysed in greater detail in sections 3.5 and 3.7

3.4.3 Statistical Models and the Bayesian Approach

The Bayesian formulation [Ripley, 1996; Geman and Geman, 1984] is a version of the regularisation prob­
lem, where the outcome is not one single mapping, but rather a distribution of most probable deform-
3.4. Registration as an optimisation problem

The associated variance of a chosen probable mapping can be used as an estimate of its accuracy, given that the true form of the target is uncertain due to noise in image acquisition.

The application of Bayesian statistics to image registration [Gee and Peralta, 1995] generates an optimisation expression similar in structure to that derived from regularisation theory shown in equation (3.5). The terms \( A \) and \( B \) are derived from the likelihood and prior, probability distributions obtained respectively from information in the data (source and target images) and prior assumptions on the form of the mapping. These are combined to give a cost function derived from a posterior distribution whose peak indicates the most likely deformation of the target reasonably close to the form of the mapping suggested by the prior. The hyperparameter \( \lambda \) can be thought of as a global measure of confidence in either the prior or likelihood estimates.

The aim is to find the deformation \( \mathbf{u} \) which gives the most likely mapping of source into target. Three factors affect the decision: the data (the source image as is presented to the algorithm); a model of source degradation (usually what noise processes were involved when the source image was produced); and prior assumptions of what the mapping should look like. The first two factors combine in the likelihood, which is a function of the similarity between source and target.

Minimising the likelihood alone may encourage bad deformations. For example, neighbouring regions may exhibit disparate differences between source / target due to different tissue structures represented by similar intensities. Should two such structures overlap in the source and target, the algorithm will be confused into thinking this area is well-matched, while neighbouring areas exhibit high differences since they contain pixels from other tissue structures whose intensities are dissimilar. As a result, the first region will be given low displacements and the neighbouring regions high, with discontinuities in the mapping along the region boundary. To avoid discontinuous or non-smooth mappings such as these, the algorithm is also encouraged to prefer a deformation which is as smooth as possible; this is the third factor in the selection of an optimum mapping.

A Bayesian methodology is used to combine the likelihood and prior information.

Bayes' Theorem states:

\[
P(A|B) \propto P(B|A) \cdot P(A)
\]

that is, the probability of the occurrence of \( A \) given an event \( B \) is proportional to the products of the probability of event \( B \) given that \( A \) has occurred and the probability of event \( A \) happening in the first place.

Bayesian statistics combines likelihoods and prior distributions in the right hand side of (3.8), to give a posterior distribution on the left hand side. The true (unknown) mapping \( \mathbf{u} \) of source to target is one possibility drawn from a distribution \( \mathcal{U} \) of all possible mappings. Similarly the presented data \( \mathcal{S}(x, y) \), the source image, is just one arrangement of pixel intensities selected from a distribution \( \mathcal{S} \) of all possible colourings of the image pixel grid.

The likelihood \( f(S|\mathbf{u}, T) \) expresses the possibility of the given appearance \( S \) of the source image given
noise degradation and other intensity-altering processes which result in the target not being merely a spatial deformation of the source. In other words, were we to know the exact (inverse) mapping from target to source (i.e., given $u$ or $u^{-1}$), the likelihood estimates how well the mapped source would represent the version of the source that we see. If no noise were present, the two should be identical and in this case the likelihood would rise to a certainty (probability = 1). The likelihood allows us to incorporate explicitly our model of the noise degradation. For example, assuming a Gaussian noise distribution, the sum of squared differences similarity measure gives the likelihood distribution

$$f(S|\tilde{u}, T) = e^{-\frac{1}{2} \sum \varepsilon(S(\tilde{u}(x)) - T(x))^2}$$

The advantage of specifying the prior and likelihood distributions to be Gaussian is that the posterior distribution will be Gaussian also\footnote{although with Bayesian methods other distributions may be considered}. Taking logs of equation (3.8) in the form

$$e^{-J} = e^{-A} e^{-\lambda B}$$

converts it to the familiar form of equation (3.5).

The prior $\pi(\tilde{u})$ is an assumption of the form of the mapping, such as that it should be smooth. We can insert a model mapping to act as the prior. For example, if the mapping will depend on a weighted sum of certain basis functions, we can introduce suggested weights by means of the prior.

The form of Bayes' theorem for image registration is therefore:

$$\pi(\tilde{u}|S, T) \propto \pi(\tilde{u}) f(S|\tilde{u}, T)$$

(3.10)

with $\pi(\tilde{u}|S)$ being the posterior distribution giving the probability of $\tilde{u}$ being the true mapping, given the particular instantiation $S$ of the data.

Selection of the optimum deformation from the posterior distribution uses a statistical estimate such as the minimum mean squared error (MMSE) and the maximum a posteriori (MAP) [Gee et al., 1995b]. We give a brief description of the latter below:

MAP - Maximum a Posteriori estimate

Let us suppose, for simplicity, that both likelihood and prior have Gaussian-shaped distributions. (We can enforce such a distribution on the prior since it is determined by us. The likelihood may have a bimodal or multimodal distribution if there is more than one likely transformation; this is the case where there is more than one minimum in the optimisation problem).

A Gaussian (normal) distribution $X \sim N(\mu, \sigma^2)$ has, in the univariate case, density

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x - \mu)^2}{2\sigma^2}}$$

with mean $\mu$ and variance $\sigma^2$ (square of the standard deviation).
The posterior distribution is derived as shown above, using Bayes’ theorem to combine the prior and likelihood distributions. Since we have assumed both prior and likelihood to have a Gaussian distribution, the posterior will be a Gaussian also (fig 3.11). The optimal transformation parameters are those which have the highest posterior probability, hence we shall select those from the MAP (*maximum a posteriori*) estimate, which (for a symmetric Gaussian distribution) is at the mean.

![Figure 3.11: Prior, likelihood and MAP parameter estimates](image)

\[
\text{Mean of MAP} = \text{weighted average of: mean of prior distribution and mean of likelihood} \quad (3.11)
\]


**Gibbs distributions and statistical physics**

A brief discussion of statistical physics [Reichl, 1980] is helpful for the understanding of prior and posterior distributions on pixel lattice deformations. The image can be modelled as a closed *non-isolated* (or open) thermodynamic system consisting of a lattice of particles whose interactions are contained within a defined neighbourhood. Hence the probability of the state of a given particle is conditional on the states of its neighbours (but unaffected by those at a distance). The total energy of the system can fluctuate due to thermal and mechanical coupling to the outside world. The probability distribution for the total energy \( G \) of such a system is called a Gibbs distribution [Kotz et al., 1983], after J W Gibbs [Gibbs, 1960], and is given by

\[
P(G = g) = \frac{1}{Z} e^{-\frac{g}{kT}} \quad (3.12)
\]

where \( k \) is Boltzmann’s constant, \( T \) is the temperature of the system and \( Z \) is a normalising constant, called the partition function or the Boltzmann probability distribution. The energy function \( G \) is called the Gibbs potential, or Gibbs free energy (the stored energy of the system that can be converted to work).
The equilibrium state for the system (at a fixed temperature \( T \), for given internal and external forces) is found by minimising its associated energy function. This gives the maximum a posteriori (MAP) estimate in the case where the distribution modelled is Bayesian. Note that the Gibbs distribution is essentially of exponential form. Ignoring the constants and taking its logarithm obtains the energy function. If the prior and likelihood both have Gibbs distributions, so will the posterior (by Bayes' theorem; and the multiplication of two exponential functions gives an exponential):

\[
\begin{align*}
\text{prior} & \sim e^{\theta p} \\
\text{likelihood} & \sim e^{\theta l} \\
\text{posterior} & \propto e^{\theta p + \theta l} = e^{\theta p + g_l} \\
\log \text{posterior} & = \log e^{\theta p + g_l} = g_p + g_l
\end{align*}
\]

and so the energy function is a sum of those associated with the prior and likelihood.

Geman and Geman [1984] pioneered the use of this analogy for image restoration. The Gibbs distribution theory is useful because it enables the modelling of the pixel grid as a Markov random field (a lattice with conditional probabilities) and then simulated annealing [Press et al., 1995] can be applied for the optimisation.**

### 3.4.4 Solution of the optimisation problem

The method of solving the registration problem (equation 3.4) computationally depends on the final representation of the deformation \( \tilde{u} \). In the case of the thin-plate and smoothing splines (sections 3.5.1 and 3.5.2), an analytic solution is given by the matrix equations (3.29) and (3.40). However in general a gradient descent method is used to find the mapping \( \tilde{u}(x) \) which minimises the cost function \( J(u) \). Optimisation may take place in the spatial or spectral domain depending on the representation of \( u \). For solution in the spatial domain, differential terms in \( J(u) \) are approximated by finite differences and then equation (3.4) is solved by an iterative method such as multigrid [Press et al., 1995] or successive overrelaxation (as in section A.4). Solution in the spectral domain uses the method of Green’s functions [Dennery and Krzywicki, 1967], a classical analytic method for the solution of second order differential equations with given boundary conditions.

As with linear registration, the reader is referred to Press et al. [1995] or to any good numerical analysis text for standard algorithms for the solution of optimisation problems. Since the optimisation surface is usually complex, some strategy is required to avoid local minima traps. A reasonable approach is to successively solve a series of equations (3.4) starting with a high value of \( \lambda \) and so placing more emphasis on the smoothing term, in an attempt to prevent the algorithm from being distracted into local minima traps in the form of high frequency localised deformations. The resulting deformation is then used as a

**The amplitude of the distribution curve 3.12 can be controlled with the parameter \( T \) - a small value increases the amplitude, and hence reduces the variance and so reduces the probability of higher energy states. This is the principle behind simulated annealing, where the parameter \( T \) is gradually reduced. Refer to sections 3.6.1 and 3.7.1 for instances of gradual modifications to the regularisation parameter during minimisation of the energy function.**
starting estimate for the solution to the next version of (3.5), with a lower value for \( \lambda \). As registration progresses and the solution is closer to the global minimum, \( \lambda \) is gradually reduced - so that the smoothing constraint has less effect and the similarity measure is allowed to carry more weight. In the statistical parametric mapping (SPM) method [Friston et al., 1995] described in section 3.6 this is effected by multiplying \( \lambda \) by the sum of differences between source and target; as registration progresses the difference decreases, bringing a corresponding decrease in \( \lambda \). Further methods of avoiding local minima are dealt with in chapter 4 on hierarchical methods.

We now analyse four models (three physical and one statistical) used as the basis for the majority of non-linear registration algorithms developed over the past twenty years. These are: the landmark-based splines, statistical parametric mapping, elastic registration and the viscous fluid model. We commence with the landmark-based splines.

### 3.5 Model One: Landmark matching with spline interpolation

The most common non-linear landmark-matching method is that of the thin-plate spline (TPS). This requires corresponding landmark pairs to be matched in the source and target, and provides a smooth macro-level interpolant to specify the deformation of the remaining image. The derivation of the mathematical model is of particular interest when seen as a non-linear extension of the matrix methods described in section 2.2. Additionally, the landmark-based spline has several variants, both in the choice of interpolating functions and in the assumptions of error in landmark locations. The thin-plate spline is also a special case of kriging methods, but this is beyond the scope of this survey; we will list references for the interested reader. We begin by describing the original TPS and its mathematical model.

#### 3.5.1 The Thin-Plate Spline

**The Physical Model**

The term **spline** originates in the ship-building industry where it refers to a thin plank of wood that was bent to the required shape by the application of weights to points along its length. The thin-plate spline is modelled on an infinitely wide thin pliable sheet, originally lying in the \((z, y)\) plane that is distorted by displacing \(m\) given points \((z_i, y_i)\) in its surface to positions \(z_i\) above and below the plane. Due to the physical nature of the material, the sheet will adopt the surface shape \(f(z, y)\) which minimises an approximation to its total bending energy (equation 3.7), which in 2D is the functional

\[
\int_\Omega \left( \frac{\partial^2 f}{\partial x^2} + 2 \frac{\partial^2 f}{\partial x \partial y} + \frac{\partial^2 f}{\partial y^2} \right) \, dx \, dy \tag{3.13}
\]

while exactly fitting the displacements, ie

\[
f(x_i, y_i) - z_i = 0 \tag{3.14}
\]

at all the constraining points \(i\). The surface satisfying these conditions is [Bookstein, 1989]

\[
f(x, y) = a_1 + a_x x + a_y y + \sum_{i=1}^{m} w_i K ((x_i, y_i) - (x, y)) \tag{3.15}
\]
where

\[ K(z) = |z|^2 \log |z|^2 \]

For 2D image warping, the original image is modelled as a flat metal sheet which is subjected to a double within plane 'bending'. Selected points \( (x, y) \), referred to as landmarks are independently displaced within the plane. The displacements have both \( x \)- and \( y \)-directional components and these are considered separately, thus two thin-plate spline models are required to warp one 2D image. The \( x \)-directional displacements are used as the 'vertical' displacements in the first thin-plate spline model, and its resulting deformed surface \( f_x(x, y) \) prescribes the \( x \)-directional components of the warp. Similarly, the \( y \)-directional landmark displacements are inserted as the vertical displacements in the second thin-plate spline model, whose resulting surface \( f_y(x, y) \) provides the \( y \)-directional warp components. The minimisation of bending energy is a smoothness criterion which ensures that the warp consists of a smoothly varying deformation between the landmark points.

The thin-plate spline as a mathematical interpolator was established by Duchon [1976] and Meinguet [1979], used extensively in the field of geostatistics [Hutchinson and Gessler, 1994; Myers, 1994] and applied to image registration by Goshtasby [1988b] and to shape deformation by [Bookstein, 1989; Bookstein and Green, 1992, 1993, 1994]. Since then they have been a popular choice for general image warping, see for example Barrodale et al. [1993] and Arad et al. [1994], and in particular (in either 2D or 3D) to medical image registration, for example Kim et al. [1995], where the initial and final landmark positions are usually corresponding readily-identifiable physiologically-significant points in the source and target images, for example [Kim et al., 1996; Sprengel et al., 1996; Little et al., 1997]. We will not give the derivation of equation (3.15) from the physical model but in the following section present the thin-plate spline deformation as an extension to the general linear transformation, give the theoretical 3D thin-plate spline and conclude with particular aspects of the warp which have relevance to the image matching problem.

The Mathematical Model

We consider mappings of the \( n \)-dimensional space \( \mathbb{R}^n \) onto itself, where \( n = 2 \) or 3.

Recall from section 2.2 the affine map in 2D space of a set of points \( \{ P_1, \ldots, P_m \} = \{(x_1, y_1), (x_2, y_2), \ldots, (x_m, y_m)\} \) into new coordinates \( \{ Q_1, \ldots, Q_m \} = \{(u_1, v_1), (u_2, v_2), \ldots, (u_m, v_m)\} \) is given with homogeneous coordinates by the matrix equation

\[
\begin{pmatrix}
u_1 & u_2 & \cdots & u_m \\
v_1 & u_2 & \cdots & u_m \\
1 & 1 & \cdots & 1
\end{pmatrix} =
\begin{pmatrix}
a_{xx} & a_{yx} & a_{1x} \\
a_{xy} & a_{yy} & a_{1y} \\
0 & 0 & 1
\end{pmatrix}
\begin{pmatrix}
x_1 & x_2 & \cdots & x_m \\
y_1 & y_2 & \cdots & y_m \\
1 & 1 & \cdots & 1
\end{pmatrix}
\]

(3.16)
or equivalently as
3.5. Model One: Landmark matching with spline interpolation

\[
\begin{pmatrix}
1 & 1 & \ldots & 1 \\
u_1 & u_2 & \ldots & u_m \\
v_1 & v_2 & \ldots & v_m
\end{pmatrix}
= \begin{pmatrix}
1 & 0 & 0 \\
a_{x_x} & a_{y_x} & a_{y_x} \\
a_{x_y} & a_{y_x} & a_{y_y}
\end{pmatrix}
\begin{pmatrix}
1 & 1 & \ldots & 1 \\
x_1 & x_2 & \ldots & x_m \\
y_1 & y_2 & \ldots & y_m
\end{pmatrix}
\]  

(3.17)

where the coefficients \(a_{x_x}\) and \(a_{y_y}\) are the translation parameters and the submatrix \(\begin{pmatrix} a_{x_x} & a_{y_x} \\ a_{x_y} & a_{y_y} \end{pmatrix}\) determines rotations, scales and shears. The form (3.16) of the equation is that conventionally used when describing affine transformations; we have re-written it in the form (3.17) since this will assist in understanding the spline theory which follows.

For simplicity we will isolate just one coordinate pair \(P_1 \mapsto Q_1\) or \((x_1, y_1) \mapsto (u_1, v_1)\), which reduces (3.17) to

\[
\begin{pmatrix}
1 \\
u_1 \\
v_1
\end{pmatrix}
= \begin{pmatrix}
1 & 0 & 0 \\
a_{x_x} & a_{y_x} & a_{y_x} \\
a_{x_y} & a_{y_x} & a_{y_y}
\end{pmatrix}
\begin{pmatrix}
1 \\
x_1 \\
y_1
\end{pmatrix}
\]  

(3.18)

The thin-plate spline extends these mappings to include non-linear distortions, with mappings of the form

\[
\begin{pmatrix}
1 \\
u_1 \\
v_1
\end{pmatrix}
= \begin{pmatrix}
0 & 0 & 0 & 0 & 1 & 0 & 0 \\
w_{x_x} & w_{x_x} & \ldots & w_{x_x} & a_{x_x} & a_{y_x} & a_{y_x} \\
w_{y_y} & w_{y_y} & \ldots & w_{y_y} & a_{x_y} & a_{y_x} & a_{y_y}
\end{pmatrix}
\begin{pmatrix}
0 \\
K(r_{12}) \\
\vdots \\
K(r_{m1}) \\
1 \\
x_1 \\
y_1
\end{pmatrix}
\]  

(3.19)

\(K(.)\) is a function of the distances \(r_{12}\) between the point \(P_1\) and another point \(P_2\) of the original point set.

For example, suppose we are mapping the set of 4 points 
\(\{P_1, P_2, P_3, P_4\} = \{(x_1, y_1), (x_2, y_2), (x_3, y_3), (x_4, y_4)\}\) into their new positions 
\(\{Q_1, Q_2, Q_3, Q_4\} = \{(u_1, v_1), (u_2, v_2), (u_3, v_3), (u_4, v_4)\}\)

The mapping of \(P_1 \mapsto Q_1\) is given by the 2 equations

\[
u_1 = w_{x_x}.0 + w_{x_x}.K(r_{12}) + w_{x_y}.K(r_{13}) + w_{x_y}.K(r_{14})
+ a_{x_x}.1 + a_{x_x}.x_1 + a_{y_y}.y_1
\]

(3.20)

\[
v_1 = w_{y_y}.0 + w_{y_y}.K(r_{12}) + w_{y_y}.K(r_{13}) + w_{y_y}.K(r_{14})
+ a_{y_y}.1 + a_{y_y}.x_1 + a_{y_y}.y_1
\]

(3.21)
Looking at equations (3.20, 3.21) we see that

1. Translations depend on the constant 1 only, so the displacement of \( P_i \) to \( Q_1 \) due to the translation part of the map is uniform over the entire image.

2. The remaining parts of the affine map - rotation, shears and/or scales, depend on \( x_1 \) and \( y_1 \), so the displacement of a point \( P_i \) due to the remaining affine parts depends on the original position of the point \( P_i \).

3. The warp depends on the distances \( r_{ij} \), so the resulting displacement of \( P_i \) depends on its location relative to the other points in the set.

Extending equation (3.17) we can write the warp equation for all points in the set as

\[
\left( \begin{array}{cccc}
    u_1 & u_2 & \cdots & u_m \\
    v_1 & v_2 & \cdots & v_m \\
\end{array} \right) = \left( \begin{array}{cc}
    W & A \\
\end{array} \right) \cdot \left( \begin{array}{c}
    K \P \\
    0 \\
\end{array} \right)
\]

where

\[
\left( \begin{array}{cc}
    W & A \\
\end{array} \right) = \left( \begin{array}{cccc}
    w_{1x} & w_{2x} & \cdots & w_{mx} \\
    w_{1y} & w_{2y} & \cdots & w_{my} \\
    a_{x1} & a_{x2} & \cdots & a_{x_m} \\
    a_{y1} & a_{y2} & \cdots & a_{y_m} \\
\end{array} \right)
\]

and

\[
L = \left( \begin{array}{ccc}
    K & P \\
    p^T & 0 \\
\end{array} \right) = \left( \begin{array}{cccc}
    0 & K(r_{12}) & \cdots & K(r_{1m}) \\
    K(r_{21}) & 0 & \cdots & K(r_{2m}) \\
    \vdots & \vdots & \ddots & \vdots \\
    K(r_{m1}) & K(r_{m2}) & \cdots & 0 \\
\end{array} \right) \left( \begin{array}{c}
    1 & z_1 & y_1 \\
    1 & z_2 & y_2 \\
    \vdots & \vdots & \vdots \\
    1 & z_m & y_m \\
\end{array} \right)
\]

and we have dropped the extra (top) row in the homogeneous coordinates, although implicitly it is still there.

To summarise so far, the transformation is given as two functions, one mapping the point set to new \( x \)-coordinates, and one giving the \( y \)-coordinate mappings, each function being of the form given in (3.15) consisting of a linear part and a higher-order part, of a weighted sum of functions of distances between the original points.
3.5. Model One: Landmark matching with spline interpolation

Returning to equation (3.17) which gives a mapping for the whole plane \( \mathbb{R}^2 \) onto itself, the coefficients \( a_1, a_x, a_y \) are uniquely determined by solving (3.17) for three known points \( P_1, P_2 \) and \( P_3 \) with their corresponding new positions \( Q_1, Q_2 \) and \( Q_3 \). For the 3-dimensional most general affine map,

\[
\begin{pmatrix}
    u_1 & u_2 & \ldots & u_m \\
    v_1 & v_2 & \ldots & v_m \\
    t_1 & t_2 & \ldots & t_m \\
    1 & 1 & \ldots & 1
\end{pmatrix} = \begin{pmatrix}
    a_{xx} & a_{yx} & a_{zx} & a_{1x} \\
    a_{xy} & a_{yy} & a_{zy} & a_{1y} \\
    a_{xz} & a_{yz} & a_{zz} & a_{1z} \\
    0 & 0 & 0 & 1
\end{pmatrix} \begin{pmatrix}
    x_1 & x_2 & \ldots & x_m \\
    y_1 & y_2 & \ldots & y_m \\
    z_1 & z_2 & \ldots & z_m
\end{pmatrix}
\] (3.25)

we insert sets of \( m = 4 \) known points \( \{P_1, \ldots, P_4\}, \{Q_1, \ldots, Q_4\}, P_i = (x_i, y_i, z_i), Q_i = (u_i, v_i, t_i) \) into (3.25) and solve for the coefficients \( a_{1n}, \quad l, n \in \{1, x, y, z\} \). We then can apply (3.25) to any other points \( P_j \) in the plane to find their corresponding mapped positions \( Q_j \). The known points which determine the mapping are called fiduciaries or landmarks.

The same is true for the 2D warp mapping, except that here the number \( m \) of initial known landmarks needed to determine the coefficients is not fixed - or rather, the warp complexity depends on how many known fiduciary mappings there are. This is because the warp part of the map for any point depends on functions of the distances between the original coordinates of the landmark set. Given the matrix \( Q \) containing the mapped coordinates of the known points,

\[
Q = \begin{pmatrix}
    u_1 & u_2 & \ldots & u_m \\
    v_1 & v_2 & \ldots & v_m
\end{pmatrix}
\] (3.26)

from which an augmented matrix \( Q_a \) is constructed:

\[
Q_a = \begin{pmatrix}
    u_1 & u_2 & \ldots & u_m & 0 & 0 & 0 \\
    v_1 & v_2 & \ldots & v_m & 0 & 0 & 0
\end{pmatrix}
\] (3.27)

and given the matrix \( L \) containing information about the original absolute and relative locations of the landmarks, we solve twice for the coefficients \( a_x \) and \( w_z \):

\[
\begin{pmatrix}
    w_1 & w_2 & \ldots & w_m \\
    a_1 & a_x & a_y
\end{pmatrix} = L^{-1}Q_a
\] (3.28)

once using the first row of \( Q_a \) to solve for the coefficients of the mapping to \( x \)-coordinates in equation (3.20), and once using the second row of \( Q_a \) for the mapping (3.21) to \( y \)-coordinates.

The 3-D version of the warp (3.22) is easily constructed as

\[
\begin{pmatrix}
    u_1 & u_2 & \ldots & u_m \\
    v_1 & v_2 & \ldots & v_m \\
    t_1 & t_2 & \ldots & t_m
\end{pmatrix} = \begin{pmatrix}
    W & A
\end{pmatrix} \cdot \begin{pmatrix}
    K & P \\
    P^T & 0
\end{pmatrix}
\] (3.29)

where

\[
\begin{pmatrix}
    W & A
\end{pmatrix} = \begin{pmatrix}
    w_{1x} & w_{2x} & \ldots & w_{mx} \\
    w_{1y} & w_{2y} & \ldots & w_{my} \\
    w_{1z} & w_{2z} & \ldots & w_{mz}
\end{pmatrix} \begin{pmatrix}
    a_{1x} & a_{xX} & a_{yX} & a_{zX} \\
    a_{1y} & a_{yX} & a_{yY} & a_{zY} \\
    a_{1z} & a_{zX} & a_{yZ} & a_{zZ}
\end{pmatrix}
\] (3.30)
and

\[
L = \begin{pmatrix}
K & P \\
P^T & 0
\end{pmatrix} = \begin{pmatrix}
0 & K(r_{12}) & \cdots & K(r_{1m}) & 1 & z_1 & y_1 & z_1 \\
K(r_{21}) & 0 & \cdots & K(r_{2m}) & 1 & z_2 & y_2 & z_2 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & \vdots \\
K(r_{m1}) & K(r_{m2}) & \cdots & 0 & 1 & z_m & y_m & z_m \\
1 & 1 & \cdots & 1 & 0 & 0 & 0 & 0 \\
z_1 & z_2 & \cdots & z_m & 0 & 0 & 0 & 0 \\
y_1 & y_2 & \cdots & y_m & 0 & 0 & 0 & 0 \\
z_1 & z_2 & \cdots & z_m & 0 & 0 & 0 & 0
\end{pmatrix}
\]

(3.31)

For the thin-plate spline, the functions \( K(r_{ij}) \) are

\[
K(r) = \begin{cases} 
  r^2 \log r : & \text{2D version} \\
  |r^3| : & \text{3D version}
\end{cases}
\]

(3.32)

The 2D version is the fundamental solution \( z(x, y) = -K(r) \) of the biharmonic equation \( \nabla^4 K = 0 \), which determines the shape in 3D space adopted by a thin pliable sheet lying roughly parallel to the plane \( z = 0 \) and fixed to the origin, so as to minimise its bending energy, in the absence of gravitational forces.

The warps are determined by the in-plane mappings of landmarks in the image, which are derived from a model of forced out-of-plane (vertical) displacements of landmarks in the pliable sheet, which otherwise would lie parallel to the plane \( z = 0 \). Bookstein [1989] works through the example of a pliable sheet constrained by 4 landmarks displacements with their original in-plane coordinates at \((1,0), (0,1), (-1,0)\) and \((0,-1)\), shown in figure 3.13.

Figure 3.13: Vertical displacements upwards of landmarks originally at \((0, \pm1)\) in the plane \( z = 0 \) and downwards of those originally at \((\pm1,0)\)

The surface adopted by the sheet to minimise its bending energy is described in terms of the original landmark coordinates \( P_k \) as

\[
z(x, y) = K\left(\sqrt{x^2 + [y-1]^2}\right) - K\left(\sqrt{(x+1)^2 + y^2}\right)
+ K\left(\sqrt{x^2 + [y+1]^2}\right) - K\left(\sqrt{(x-1)^2 + y^2}\right)
= \sum_{k=1}^{4} (-1)^k K(||(x, y) - P_k||)
\]

(3.33)
3.5. Model One: Landmark matching with spline interpolation

which is an instance of the warp (3.20) with no affine component and weights \( w_i = \pm 1 \); it is a weighted sum of the fundamental solution \( K(\tau) \). Bookstein [1989] examines the behaviour of the plate far from the landmarks as follows:

Equation (3.33) when re-written with \( V(s) = K(\sqrt{s}) \) is

\[
\begin{aligned}
z(x, y) &= V(z^2 + [y + 1]^2) - 2V(x^2 + y^2 + 1) + V((x - 1)^2 + y^2) \\
&\quad - \{ V((x + 1)^2 + y^2) - 2V(x^2 + y^2 + 1) \\
&\quad + V((x - 1)^2 + y^2) \}
\end{aligned}
\]  

(3.34)

Using the second derivative finite difference approximation

\[
\frac{d^2f}{ds^2} \approx \frac{f(s + h) - 2f(s) + f(s - h)}{h^2}
\]

with \( h = 2y \) or \( h = 2x \), (3.34) reduces to

\[
z(x, y) \approx (2y)^2 \frac{d^2V(s)}{ds^2} - (2x)^2 \frac{d^2V(s)}{ds^2}
\]

with \( s = x^2 + y^2 + 1 \). But \( V(s) = s \log s \implies \frac{d^2V(s)}{ds^2} = \frac{1}{s} \)

So we have

\[
z(x, y) \approx \frac{4(y^2 - x^2)}{x^2 + y^2 + 1}
\]

or for large \( x, y \),

\[
z(x, y) \approx \frac{4(y^2 - x^2)}{x^2 + y^2}
\]

Finally, using the double-angle identity \( \cos 2\theta = \cos^2 \theta - \sin^2 \theta \) and \( \cos(\arctan(\frac{y}{x})) = \frac{x}{\sqrt{x^2 + y^2}} \),

\[
\sin(\arctan(\frac{y}{x})) = \frac{y}{\sqrt{x^2 + y^2}}
\]

\[
z(x, y) \approx -4 \cos(\theta) \quad \text{with} \quad \theta = 2 \arctan(\frac{y}{x})
\]

Hence far from the origin, the surface is sinusoidal (with small amplitude of 4). This displays an undesirable quality of the thin-plate spline for medical image warping, namely that it does not have finite support, i.e., perturbations of landmarks in one region of the image will necessarily involve small movements over the whole image. Note also that the spline was originally intended as an interpolant rather than an extrapolant and as such does not claim to accurately predict behaviour in regions far from the known data (landmarks).

The total bending energy for the deformation is given by \( W^T K W \), or equivalently \( Q^T L_{m}^{-1}Q \), where \( K \) and \( W \) are as given in equations (3.24) and (3.23) or (3.31) and (3.30) respectively, \( Q \) contains the mapped landmarks (equation 3.26) and \( L_{m}^{-1} \) is the upper \( m \times m \) submatrix of \( L^{-1} \). A description of the deformation is provided by the eigenvectors of \( L_{m}^{-1} K L_{m}^{-1} \), known as principal warps. These are directions within the initial landmark configuration, such as along axes describing the configuration shape, and the corresponding eigenvalues give ratios of contributions to the bending energy in the principal warp directions. For example, if a 2-dimensional warp summarised by \( L_{m}^{-1} K L_{m}^{-1} \) has eigenvectors \( \vec{e}_1, \vec{e}_2 \) and
with corresponding eigenvalues $\alpha_1$, $\alpha_2$ and $\alpha_3$, the $x$- and $y$- direction warps $f_x(x, y)$ and $f_y(x, y)$ given by equation 3.15 can each be expanded in terms of $\tilde{e}_1$, $\tilde{e}_2$ and $\tilde{e}_3$:

$$f_x(x, y) = c_1 \tilde{e}_1 + c_2 \tilde{e}_2 + c_3 \tilde{e}_3$$  \hspace{1cm} (3.35)

$$f_y(x, y) = d_1 \tilde{e}_1 + d_2 \tilde{e}_2 + d_3 \tilde{e}_3$$  \hspace{1cm} (3.36)

for some constants $c_i$, $d_i$. The total bending energy for the warp is given by

$$\alpha_1 (c_1^2 + d_1^2) + \alpha_2 (c_2^2 + d_2^2) + \alpha_3 (c_3^2 + d_3^2)$$  \hspace{1cm} (3.37)

The theory of principal warps is detailed in [Bookstein, 1989].

There are variations of the thin-plate spline which improve its usefulness and expand its range of applications. Firstly, the basis function $K(r)$ in equation (3.32) depends on the chosen physical model: this choice of $K(r)$ is the Green's function solving the biharmonic equation

$$\nabla^4 K = \delta$$

minimising the bending energy of a pliable sheet. Other radial basis functions are equally, if not better, suited, once strict adherence to the physical model is no longer required. For example, some basis functions introduce parameters governing the localisation of the warps. Also, if the exact interpolation conditions of the new landmark locations is not enforced, error or uncertainty in landmark locations can be built into the model, with the measures of uncertainty modifying the basis-function weights. All these variants will be discussed in the following sections.

### 3.5.2 Extensions to the Thin-Plate Spline

**Smoothing spline**

Often the exact locations of the landmarks is uncertain, when the landmarks have been selected manually or by an automatic landmark-finding algorithm. In this case the interpolating conditions $f(x_i, y_i) = (u_i, v_i)$ (for the $m$ landmark sets $\{(x_i, y_i), \{u_i, v_i\}, i = 1, \ldots, m\}$ can be relaxed in favour of a smoother warp [Arad et al., 1994; Myers, 1994]. The method will not correct for systematic errors, where all landmarks share an underlying smoothly varying error, but may correct for outliers - cases of only a few misplaced points from the landmark sets - which produce a higher degree of bending. The interpolating thin-plate spline finds the basis-functions coefficients to minimise the bending energy functional (3.13) while satisfying the interpolating landmark conditions (equation 3.14) $f(x_i, y_i) - (u_i, v_i) = 0$. This can be recast as a regularisation problem (equation 3.4) by introducing the similarity measure $D(S(f(x_i, y_i), T(u_i, v_i)) = ||f(x_i, y_i) - (u_i, v_i)||^2$ and requiring the basis-functions coefficients to minimise over the whole image a new combined functional

$$\tilde{J}(f) = \lambda R(f) + D(S(f), T)$$

$$= \lambda \int_\Omega \left( \frac{\partial^2 f}{\partial x^2}^2 + 2 \left( \frac{\partial^2 f}{\partial x \partial y} \right)^2 + \left( \frac{\partial^2 f}{\partial y^2} \right)^2 \right) dx \; dy +$$

$$\sum_{i=1}^m ||f(x_i, y_i) - (u_i, v_i)||^2$$  \hspace{1cm} (3.38)
where the roughness measure \( R(f(x_i, y_i)) \) is the thin-plate bending term
\[
\left( \frac{\partial^2 f}{\partial x^2} \right)^2 + 2 \left( \frac{\partial^2 f}{\partial x \partial y} \right)^2 + \left( \frac{\partial^2 f}{\partial y^2} \right)^2
\]  
(3.39)

The limiting case of \( \lambda \rightarrow 0 \) gives the interpolant, where the landmarks match exactly (satisfying 3.14). For an approximating or smoothing spline, the tradeoff parameter or regularisation constant \( \lambda \) can take any value between 0 and 1, with a higher \( \lambda \) placing a greater emphasis on the bending energy and relaxing the interpolation conditions.

The thin-plate spline solution, equation (3.29) is then replaced by
\[
\begin{bmatrix}
\mathbf{u}_1 \\
\mathbf{v}_1 \\
\mathbf{t}_1 \\
\mathbf{u}_2 \\
\mathbf{v}_2 \\
\mathbf{t}_2 \\
\vdots \\
\mathbf{u}_m \\
\mathbf{v}_m \\
\mathbf{t}_m
\end{bmatrix}
= \begin{bmatrix}
\mathbf{W} & \mathbf{A}
\end{bmatrix} \cdot \begin{bmatrix}
\mathbf{K} + \lambda \mathbf{I} & \mathbf{P} \\
\mathbf{P}^T & 0
\end{bmatrix}
\]  
(3.40)

where \( \mathbf{I} \) is the \( m \times m \) identity matrix.

Dyn and Wahba [1982] and Rohr et al. [1996] take this one stage further; we may have an estimate of the accuracy of each target landmark position, given by the variance \( \sigma_i \) of the probability distribution of its location. In this case, equation (3.38) becomes
\[
\tilde{J}(f) = \lambda \int_{\Omega} \left( \frac{\partial^2 f}{\partial x^2} \right)^2 + 2 \left( \frac{\partial^2 f}{\partial x \partial y} \right)^2 + \left( \frac{\partial^2 f}{\partial y^2} \right)^2 \, dx \, dy + \sum_{i=1}^{m} \frac{1}{\sigma_i^2} \left\| f(x_i, y_i) - (u_i, v_i) \right\|^2
\]  
(3.41)

which is solved by
\[
\begin{bmatrix}
\mathbf{u}_1 \\
\mathbf{v}_1 \\
\mathbf{t}_1 \\
\mathbf{u}_2 \\
\mathbf{v}_2 \\
\mathbf{t}_2 \\
\vdots \\
\mathbf{u}_m \\
\mathbf{v}_m \\
\mathbf{t}_m
\end{bmatrix}
= \begin{bmatrix}
\mathbf{W} & \mathbf{A}
\end{bmatrix} \cdot \begin{bmatrix}
\mathbf{K} + \lambda \mathbf{C}^2 & \mathbf{P} \\
\mathbf{P}^T & 0
\end{bmatrix}
\]  
(3.42)

where
\[
\mathbf{C} = \begin{bmatrix}
\sigma_1 & & 0 \\
& \ddots & \\
0 & & \sigma_m
\end{bmatrix}
\]

Automated Thin-Plate Spline

Kim et al. [1996] produce a (semi-) automated thin-plate spline, where an initial choice of source landmarks are iteratively repositioned so as to minimise the measure of mutual information in the image pair. They break down their algorithm into four separate tasks, which they describe as:

1. **registration**: determine spline coefficients
2. **reconstruction** (trilinear)
3. **MI** (mutual information) computation
4. multivariate **minimisation**
One could argue instead that the latter two tasks are the matching part of the registration, by the optimisation of a similarity measure. The first two tasks are merely interpolants of the estimated mapping which the matching process only specifies for the landmark points, with the first task and the second task being macro and micro interpolants respectively (section 3.3).

Matching edgels and other features

Landmark-based spline theory can be expanded to include spatial derivatives at landmark positions, called edgels. These are modelled by satellite landmarks infinitesimally close to existing landmark locations; altering the direction of the satellites from the landmarks gives a warp of the surface. With the addition of edgels to the spline model, directions of normals to isosurfaces in the images can be matched, which is an advantage where landmarks are less-easily identified. Bookstein and Green [1992] present the theory in detail. Joshi et al. [1995] present an extension to the theory, called the generalised Dirichlet problem, which includes curves and surfaces as features to be matched.

Alternative Basis Functions

The functions $K(r_{ij}) = K(x_i, x_j)$ form a basis for the space of non-linear transformations. They are called radial basis functions since for any point pair $(x_i, x_j)$ they depend only on the distance $r_{ij} = \|x_i - x_j\|$ (figure 3.14).

The first radial basis functions used for interpolation of discrete data to a smooth surface were introduced by Hardy [1971]. His aim was to estimate a continuous terrain of heights $z$ above an $x, y$ plane from a few known heights $z_i$. Previous attempts by others to use Fourier or polynomial series expansions had failed due to rapid oscillations in the reconstructed surface between the known data points. Hardy arrived at his basis functions empirically; they are families of quadrics or multiquadrics, which in 2D are of the form

$$K((x_i, y_i), (x_j, y_j)) = \left[ (x_i - x_j)^2 + (y_i - y_j)^2 + C \right]^{\frac{3}{2}}$$

(3.43)

The constants $\alpha$ and $C$ determine each quadric family, for example $\alpha = 1$ gives the family of circular hyperboloids $w \left[ (x_i - x_j)^2 + (y_i - y_j)^2 + C \right]^{\frac{1}{2}} = z$, and $\alpha = 2$ gives circular paraboloids $w \left[ (x_i - x_j)^2 + (y_i - y_j)^2 + C \right] = z$ (fig. 3.15).

The coefficient $w$ determines the steepness of the conical asymptote of the quadric surfaces, and decreasing $C$ pulls the apex of the surface towards the apex of the conical asymptote, thus setting $C = 0$ reduces
3.5. Model One: Landmark matching with spline interpolation

The quadric to a conic (fig. 3.16). These constants therefore regulate the jaggedness of the surface terrains that can be reproduced by the multi-quadric interpolants. In the case of image registration, they will affect the smoothness of the warps.

Arad et al. [1994] offers a further two types of radial function for image warping, the shifted log and Gaussian:

\[
\text{shifted log} : K(r) = \log \left( r^2 + c^2 \right) \frac{1}{2}, \quad c^2 \geq 1
\]

\[
\text{Gaussian} : K(r) = \exp \left( -\frac{r^2}{\sigma^2} \right), \quad \sigma > 0
\]

The shifted log is a variant of the circular hyperboloid. The Gaussian in profile is a bell-shaped curve whose steepness or width is controlled by the variance or localisation parameter \( \sigma \). Decreasing \( \sigma \) has the effect of making the warps more severe at the landmarks, that is, more localised to the cause of the deformations.
Kriging

We mention the classical use of the thin-plate spline interpolant in the discipline of geostatics, where it is used to generate intermediate values from irregularly and sparsely sampled data. The spline literature in this field is abundant and may be worth studying if further investigation into the behaviour of the thin-plate spline is desired. Specifically, the spline has been shown to be closely related to another more general interpolant process, kriging†† [Oliver and Webster, 1991; Yandell, 1993; Handcock and Stein, 1993; Le and Zidek, 1992; Kwaadsteniet, 1990; Hansen, 1993; Pohlmann, 1993], which makes (fuller) use of prior statistical knowledge of the data domain, estimates errors in the landmark distribution and can identify outliers. The connection between splines and kriging is explored in [Mardia et al., 1996; Myers, 1992; Wu and Schaback, 1993; Myers, 1994; Hutchinson and Gessler, 1994; Dubrule, 1984; Yandell, 1993]. Kriging has been used in medical imaging for between-slice interpolation and voxel-value interpolation during contour extraction, [Stytz and Parrott, 1993].

The elastic body spline

Finally, the elastic body spline (EBS), Davis et al. [1997], models the image as having elastic properties, replacing the thin-plate bending energy (3.39) with the elasticity energy functional

\[
\mu \nabla^2 \bar{u}(\bar{x}) + (\mu + \lambda) \nabla (\nabla \cdot \bar{u}(\bar{x}))
\]  

(3.44)

where \(\mu\) and \(\lambda\) are elasticity constants (not to be confused with the regularisation parameter \(\lambda\)) and \(\bar{u}(\bar{x}) = (u - x, v - y, t - z)\) are the displacements of the points in the source image. Deformations derived from the elasticity operator will be discussed further in section 3.7.

The matrix solution is more complex than (3.29), with the basis functions \(K(\bar{x})\) replaced in the 3D case by submatrices \(\tilde{K}(\bar{x})\) of dimension 3 x 3 given by

\[
\tilde{K}(\bar{x}) = \left[ \alpha |\bar{x}|^2 I - 3\bar{x}\bar{x}^T \right] |\bar{x}|
\]  

(3.45)

where \(I\) is the 3 x 3 identity matrix, \(\alpha\) is a constant given by

\[
\alpha = 12(1 - \nu) - 1
\]

††named after D G Krige
and $\nu$ is Poisson’s ratio:

$$\nu = \frac{\lambda}{2(\lambda + \mu)}$$

The sets of mappings $f(\tilde{x})$ in equation (3.15) are then

$$f(\tilde{x}) = a_1 + \tilde{A} \tilde{x} + \sum_{i=1}^{m} u_i \tilde{K}(|\tilde{x}_i - \tilde{x}|)$$  \hspace{1cm} (3.46)

where

$$\tilde{A} = \begin{bmatrix} a_{x1} & a_{x2} & a_{x3} \\ a_{y1} & a_{y2} & a_{y3} \\ a_{z1} & a_{z2} & a_{z3} \end{bmatrix}^T$$

$$a_{x1} = a_{x} \quad a_{y1} = a_{y} \quad a_{z1} = a_{z}$$

$\tilde{A}$ is collapsed into a vector $A$ and all the coefficients to be solved are stored in a vector of length $3N + 12$ ($N$ landmarks):

$$\begin{bmatrix} W | A \end{bmatrix} = \begin{bmatrix} w_1^T & \cdots & w_N^T & a_1^T & a_x^T & a_y^T & a_z^T \end{bmatrix}$$  \hspace{1cm} (3.47)

Likewise, the displacements in all three directions of all the landmark points are stored in one augmented vector of length $3N + 12$

$$Q = \begin{bmatrix} u_1^T & \cdots & u_N^T \end{bmatrix}$$

and then the modified system 3.28 is solved for the coefficients $(W | A)$ as before.

Davis et al. [1997] claims the EBS produces better results than the thin-plate spline for medical image applications.

## 3.6 Model Two: SPM and other truncated basis function expansion methods

Interpolating splines such as the thin-plate spline and radial basis function expansions have the disadvantage that corresponding landmarks must first be identified in the source and target images. This stage is time-consuming if performed manually and correspondence is rarely guaranteed when automated. In addition, most of the information content in the image pair is ignored. In linear registration there has been much success with pixel-based registration algorithms, where the transformation is automatically generated by the minimisation of a dissimilarity measure, such as those described in section 2.5.1, computed from all or a large number of source/target pixel intensity pairs. Non-linear equivalents have developed, in which the deformation is expanded typically in a Fourier or other trigonometric basis. The weights
of the basis functions are determined so as to optimise the match in terms of the dissimilarity measure, while smoothness is enforced by promoting deformations with heavier weights on the lower frequency functions. Since a truncated basis function set is used, the choice of basis in which the warp is expanded will also affect the smoothness and localisation of the deformations.

We analyse in this section an implementation of a purely statistical basis-function expansion registration called SPM (statistical parametric mapping) [Friston et al., 1995].

3.6.1 Solution using spectral decomposition

The target $T$ and source $S$ are assumed to be related by equation (3.1), and both defined over the domain $\Omega_{3D} = \{ (x \in [1, x_{\text{dim}}], y \in [1, y_{\text{dim}}], z \in [1, z_{\text{dim}}]) \}$. The function $\phi$ in (3.1) is an optional mapping of the intensities of the target, useful for example if the overall intensity of one image is higher than the other. $\phi$ may even vary spatially over the image [Friston et al., 1995], in which case $\phi$ itself may be expanded in terms of some basis

$$\phi(T(\mathbf{z})) = \sum_k \alpha_k \phi_k(T(\mathbf{z}))$$

(3.48)

We will assume here that the unknown intensity mapping $\phi$ is a constant.

The mapping $\mathbf{u}(\mathbf{z})$ in equation (1.1) is to be derived as a truncated expansion of the cosine basis $\{ \beta_m(x) = \cos(mx) \}$ with coefficients $\psi(\mathbf{u}_i) = \mathbf{u}_i^T C_0^{-1} \mathbf{u}_i$ for a given $M \times M$ diagonal covariance matrix $C_0^{-1}$, $M = n_x n_y n_z$, where $n_x$, $n_y$, $n_z$ are the number of basis functions in each of the $x$, $y$ and $z$ directions respectively. This gives a prior distribution of the deformation in the form

$$\pi(\mathbf{u}_i) = e^{-\frac{1}{2} \mathbf{u}_i^T C_0^{-1} \mathbf{u}_i}$$

(3.49)

The likelihood distribution $f(S|\mathbf{u}, T)$ is taken as a function of the sum of squares dissimilarity measure, $\mathcal{D} = \sum (S(\mathbf{u}(\mathbf{z})) - \phi T(\mathbf{z}))^2$. Assuming a Gaussian noise distribution, this gives

$$f(S|\mathbf{u}, T) = e^{-\frac{1}{2} \sum (S(\mathbf{u}(\mathbf{z})) - \phi T(\mathbf{z}))^2}$$

(3.50)

which models the noise in a perfectly deformed source $S$.

The prior (3.49) and likelihood (3.50) are inserted into Bayes' formula (3.10) to obtain the posterior:

$$\pi(\mathbf{u}|S, T) \propto e^{-\frac{1}{2} \mathbf{u}^T C_0^{-1} \mathbf{u} - \frac{1}{2} \sum (S(\mathbf{u}(\mathbf{z})) - \phi T(\mathbf{z}))^2}$$

$$\log(\text{posterior}) = \frac{1}{2} \mathbf{u}^T C_0^{-1} \mathbf{u} - \frac{1}{2} \sum (S(\mathbf{u}(\mathbf{z})) - \phi T(\mathbf{z}))^2$$

The negation of the log of the posterior gives the cost function (3.4)

$$J(\mathbf{u}) = -\log(\text{posterior}) = \frac{1}{2} \mathbf{u}^T C_0^{-1} \mathbf{u} + \frac{1}{2} \sum (S(\mathbf{u}(\mathbf{z})) - \phi T(\mathbf{z}))^2$$
3.6. Model Two: SPM and other truncated basis function expansion methods

\[ R^V \] (3.51)

where so far we have considered the prior and likelihood to be of equal relevance in determining the posterior, and so \( \lambda \) in equation (3.4) equals 1. The optimal solution maximises the posterior, or, equivalently, minimises the cost function. It is obtained iteratively within a matrix framework, as follows:

Let \( \vec{y} \) be a vector\(^{11}\) storing the pixel-wise differences between currently deformed source \( S(\vec{u}(\vec{x})) \) and target \( T(\vec{x}) \):

\[
\vec{y} = \begin{bmatrix}
\vdots \\
S(\vec{u}(\vec{x}j)) - \phi T(xj) \\
\vdots
\end{bmatrix} \uparrow \quad j = 1, \ldots, N
\]

(3.52)

where \( N \) is the total number of pixels in \( \Omega_{3D} \).

Using (3.52) the log likelihood is re-written as the scalar product

\[
D = \frac{1}{2} \sum_j (S(\vec{u}(\vec{x}j)) - \phi T(xj))^2 = \frac{1}{2} \vec{y}^T \vec{y}
\]

The mapping \( \vec{u}(\vec{x}i) \), where \( \vec{x}i = (x_i, y_i, z_i) \), is to be expanded with coefficients \( \{ u_{1k}, u_{2l}, u_{3m} \} \) in a basis \( \{ \beta_k, \beta_l, \beta_m \} \), \( k = 1, \ldots, nx \), \( l = 1, \ldots, ny \), \( m = 1, \ldots, nz \), that is:

\[
\vec{u}(\vec{x}i) = \sum_{k=1}^{nx} u_{1k} \beta_k (x_i, y_i) + \sum_{l=1}^{ny} u_{2l} \beta_l (y_i) + \sum_{m=1}^{nz} u_{3m} \beta_m (z_i)
\]

giving the deforming source

\[
S(\vec{u}(\vec{x}i)) = S \left( x_i + \sum_k u_{1k} \beta_k (x_i), y_i + \sum_l u_{2l} \beta_l (y_i), z_i + \sum_m u_{3m} \beta_m (z_i) \right)
\]

(3.53)

Taylor approximating (3.53) gives

\[
S(\vec{u}(\vec{x}i)) \approx S(\vec{x}i) + \sum_k u_{1k} \beta_k (x_i) \frac{\partial S(\vec{x}i)}{\partial x}
\]

\[
+ \sum_l u_{2l} \beta_l (y_i) \frac{\partial S(\vec{x}i)}{\partial y} + \sum_m u_{3m} \beta_m (z_i) \frac{\partial S(\vec{x}i)}{\partial z}
\]

(3.54)

The derivatives of the log likelihood \( D \) with respect to the coefficients \( u_{1k}, k \in \{1, \ldots, nx \} \) are

\[
\frac{\partial D}{\partial u_{1k}} = \sum_i (S(\vec{u}(\vec{x}i)) - \phi T(x_i)) \cdot \frac{\partial S(\vec{u}(\vec{x}i))}{\partial u_{1k}}
\]

(3.55)

and similarly for the derivatives with respect to the coefficients \( u_{1l}, l \in \{1, \ldots, ny \} \) and \( u_{3m}, m \in \{1, \ldots, nz \} \).

\(^{11}\)To construct \( \vec{y} \), the three-dimensionality of the images has been flattened by writing out the pixel intensities row by row, plane by plane. This is common when solving transformations by matrix methods, another example being the vector of coefficients in (3.47)
Combining (3.54) and (3.55) gives the differentials

\[
\begin{align*}
\frac{\partial \mathcal{D}}{\partial u_{1k}} &= \sum_i \left( S(\mathbf{u}(\mathbf{z}_i)) - \phi T(x_i) \right) \cdot \beta_k \left( x_i \right) \cdot \frac{\partial S(\mathbf{z}_i)}{\partial x} \\
\frac{\partial \mathcal{D}}{\partial u_{2i}} &= \sum_i \left( S(\mathbf{u}(\mathbf{z}_i)) - \phi T(x_i) \right) \cdot \beta_i \left( y_i \right) \cdot \frac{\partial S(\mathbf{z}_i)}{\partial y} \\
\frac{\partial \mathcal{D}}{\partial u_{3m}} &= \sum_i \left( S(\mathbf{u}(\mathbf{z}_i)) - \phi T(x_i) \right) \cdot \beta_m \left( z_i \right) \cdot \frac{\partial S(\mathbf{z}_i)}{\partial z} \\
\frac{\partial \mathcal{D}}{\partial \phi} &= \sum_i \left( S(\mathbf{u}(\mathbf{z}_i)) - \phi T(x_i) \right) \cdot T(\mathbf{z}_i)
\end{align*}
\]

This can be abbreviated as

\[
(\nabla \mathcal{D})^T = \mathbf{g}^T A,
\]

a row vector, where \( \mathbf{g} \) is given by (3.52) and \( A \) is the matrix

\[
A = \begin{bmatrix}
\vdots & \vdots & \vdots \\
\frac{\partial S(\mathbf{z}_i)}{\partial x} \cdot \beta_k & \frac{\partial S(\mathbf{z}_i)}{\partial y} \cdot \beta_i & \frac{\partial S(\mathbf{z}_i)}{\partial z} \cdot \beta_m & T(\mathbf{z}_i) \\
\vdots & \vdots & \vdots \\
\end{bmatrix}
\]

where \( i = 1, \ldots, N \).

We also have

\[
\begin{align*}
\frac{\partial^2 \mathcal{D}}{\partial u_{1k} \partial u_{2i}} &= \sum_i \beta_k \left( x_i \right) \frac{\partial S(\mathbf{z}_i)}{\partial x} \cdot \beta_i \left( y_i \right) \cdot \frac{\partial S(\mathbf{z}_i)}{\partial y} \\
\frac{\partial^2 \mathcal{D}}{\partial u_{1k} \partial u_{3m}} &= \sum_i \beta_k \left( x_i \right) \frac{\partial S(\mathbf{z}_i)}{\partial x} \cdot \beta_m \left( z_i \right) \cdot \frac{\partial S(\mathbf{z}_i)}{\partial z} \\
\frac{\partial^2 \mathcal{D}}{\partial u_{2i} \partial u_{3m}} &= \sum_i \beta_i \left( y_i \right) \frac{\partial S(\mathbf{z}_i)}{\partial y} \cdot \beta_m \left( z_i \right) \cdot \frac{\partial S(\mathbf{z}_i)}{\partial z} \\
\frac{\partial^2 \mathcal{D}}{\partial u_{1k} \partial \phi} &= \sum_i -T(\mathbf{z}_i) \cdot \beta_k \left( x_i \right) \cdot \frac{\partial S(\mathbf{z}_i)}{\partial x} \\
\frac{\partial^2 \mathcal{D}}{\partial u_{2i} \partial \phi} &= \sum_i -T(\mathbf{z}_i) \cdot \beta_i \left( y_i \right) \cdot \frac{\partial S(\mathbf{z}_i)}{\partial y} \\
\frac{\partial^2 \mathcal{D}}{\partial u_{3m} \partial \phi} &= \sum_i -T(\mathbf{z}_i) \cdot \beta_m \left( z_i \right) \cdot \frac{\partial S(\mathbf{z}_i)}{\partial z} \\
\frac{\partial^2 \mathcal{D}}{\partial \phi^2} &= \sum_i T(\mathbf{z}_i)^2 \nabla^2 \mathcal{D} = A^T A
\end{align*}
\]
3.6. Model Two: SPM and other truncated basis function expansion methods

$C_0^{-1}$ is a given diagonal matrix,

$$
C_0^{-1} = \begin{pmatrix}
C_{011} & 0 & \cdots & 0 \\
0 & C_{022} & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & C_{0NN}^r \\
\end{pmatrix}
$$

(3.58)

Setting $C_0^{-1} = I$, the identity matrix, the prior distribution would be $e^{-\frac{1}{2}d^T I d} = e^{-\frac{1}{2} \Sigma_k u_k^2} = \prod_k e^{-\frac{1}{2} \sigma_k^2}$. In this case, the prior probability is the product of the individual prior probabilities of the parameters which are all weighted equally. However, giving $C_0^{-1}$ diagonal elements increasing with row number $k$, produces the prior $e^{-\frac{1}{2} d^T C_0^{-1} d} = \prod_k e^{-\frac{1}{2} \sigma_k^2 C_{0k}^{-1}}$, where each parameter has a different weighting being the variance $= C_0^{-1} = \frac{1}{\sigma_k^2}$. The coefficients of the higher frequency basis functions are given lower weighting and those of lower frequencies are given higher variance so they are more likely to be non-zero.

From (3.51) we now have

$$
\frac{\partial R}{\partial u_{1k}} = u_{1k} C_{0kk}^{-1}
$$

and

$$
\begin{align*}
\frac{\partial^2 R}{\partial u_{1k}^2} &= C_{0kk}^{-1} \\
\frac{\partial^2 R}{\partial u_{1k} \partial u_{2l}} &= 0 \text{ for } l \neq k
\end{align*}
$$

or in matrix form:

$$
\nabla R = \begin{bmatrix}
\vdots \\
u_{1k} C_{0kk}^{-1} \uparrow \\
\vdots \\
\end{bmatrix}
\begin{bmatrix}
k = 1, \ldots, M \\
\text{row vector: } \quad (\nabla R)^T = \bar{u}^T C_0^{-1}
\end{bmatrix}
$$

(3.59)

and

$$
\nabla^2 R = C_0^{-1}
$$

(3.60)

The log posterior $J(\bar{u}) = R(\bar{u}) + D(\bar{u})$ is to be maximised. In practice a regularising multiplier $\lambda$ is applied to the log prior and a succession of optimising problems are solved whose coefficient solutions are input as initial estimates to the next iteration, with $\lambda$ decreasing as the match is gradually optimised. The decrease in $\lambda$ is effected by multiplying it by the normalised similarity measure

$$
d(S(\bar{u}), T) = \frac{\sum_{x} (S(\bar{u}(\bar{x}) - T(\bar{x}))^2}{\text{total number of pixels}}
$$

Thus the aim is to maximise

$$
J(\bar{u}) = D(\bar{u}) + \lambda d R(\bar{u})
$$

(3.61)

or equivalently to find the zeros of $\nabla J(\bar{u})$, with starting values $\bar{u} = \bar{u}_0 = \{0\}$, i.e starting with the identity as the initial mapping.
A gradient descent method [Press et al., 1995] is used to find the zeros of $\nabla J(\bar{u})$ iteratively. At iteration $n$, with a given estimate of the coefficients $\bar{u}$ describing the deformation, the gradient of the updated cost function $J(\bar{u} + \Delta \bar{u})$ is Taylor expanded:

$$\nabla J(\bar{u}^n + \Delta \bar{u}^n) \approx \nabla J(\bar{u}^n) + \Delta \bar{u}^n \nabla^2 J(\bar{u}^n)$$  \hspace{1cm} (3.62)

with the left hand side assumed to approximate zero for the update $\Delta \bar{u}^n$.

The computation is to be done by pre-multiplication of matrices by row vectors, hence (3.62) is constructed as

$$0 \approx (\nabla J(\bar{u}^n))^T + (\bar{u}^{n+1} - \bar{u}^n)^T \nabla^2 J(\bar{u}^n)$$  \hspace{1cm} (3.63)

Substituting from (3.61) gives

$$0 \approx (\nabla D)^T + \lambda d (\nabla R)^T + (\bar{u}^{n+1} - \bar{u}^n)^T \nabla^2 D(\bar{u}^n) + \lambda d \nabla^2 R(\bar{u}^n)$$

and now inserting the matrix expressions (3.59) and (3.60) for the gradients of the prior and (3.56) and (3.57) for the gradients of the likelihood gives

$$0 \approx \bar{g}^T A + \lambda d (\bar{u}^n)^T C_0^{-1} + (\bar{u}^{n+1} - \bar{u}^n)^T [A^T A + \lambda d C_0^{-1}]$$

$$- (\bar{u}^{n+1})^T (A^T A + \lambda d C_0^{-1}) \approx \bar{g}^T A + \lambda d (\bar{u}^n)^T C_0^{-1} - (\bar{u}^n)^T A^T A - \lambda d (\bar{u}^n)^T C_0^{-1}$$

$$\Rightarrow (\bar{u}^{n+1})^T \approx (A^T A + \lambda d C_0^{-1})^{-1} (\bar{u}^n)^T A^T A - \bar{g}^T A$$

which is the scheme for updating the basis function coefficients.

$\bar{g}$ is only an approximation to the true differences $S(\bar{u}(\bar{x})) - T(x)$ since the true deformation $\bar{u}(x)$ is unknown. Instead $\bar{g}$ contains the pixelwise differences of the current deformed source, $S$, and the target $T$. An alternative [Ashburner, 1997] is to fill $\bar{g}$ with the pixel intensities of $S(\bar{u})$, since the target $T$ is unchanging.

### 3.6.2 Derivation of basis functions for bending energy smoothness constraints

In SPM smoothness was imposed by an a priori assumption on the distribution of the basis function weights, setting lower values to those weighting the higher frequency functions. However, no particular rationale was given for the choice of the basis functions themselves. Since the space of possible mappings is limited by the truncated basis function set, the choice of functions and size of the basis will influence the variety of possible mappings. Amit et al. [1991] employ such considerations and their method is now described for a 2D transformation.

We wish to find a mapping $\bar{u}(\bar{x}) = (u_1(\bar{x}), u_2(\bar{x})) : S(\bar{x} + \bar{u}(\bar{x})) = T(\bar{x})$ with the conditions that $\bar{u}(\bar{x})$ is smooth and the boundary does not move. An assumption is that $u_1$ and $u_2$ are independent.

The mappings are to be expanded out in a set of basis functions, to be determined so that the conditions are met. The mappings are to be drawn from a 2D Gaussian distribution with the identity map as the mean,
that is, we want the mappings to be closely similar to the identity transformation, with the probability of
a transformation being allowed decreasing exponentially as it varies from this mean.

The Karhunen-Loève (K-L) series expansion of a random field \( f(x, y) \) is given by Jain [1989]:

\[
f(x, y) = \sum_{m,n=0}^{\infty} a_{mn} \phi_{mn}(x, y)
\]

where \( \{\phi_{mn}(x, y)\} \) are the eigenfunctions of the autocorrelation function (or alternatively the covariance
matrix) of the random field \( f(x, y) \) and the \( a_{mn} \) are orthogonal random variables.

In practice, the expansion must be used in a truncated form. The K-L expansion gives the lowest mean
squared error from the true \( f(x, y) \) of all truncated expansions for a given truncation length \( N \).

The K-L expansion is also known as the Hotelling transform or principal components analysis. The
eigenfunctions are ordered by decreasing size of eigenvalues so that those producing the strongest de­
formations are nearer the beginning of the list and the truncated tail of the expansion contains the more
subtle deformations only.

We now choose the Gaussian distribution to take into account the required smoothness and boundary
conditions. Smoothness is imposed by minimising the bending in the mapping \( (u_1(\bar{x}), u_2(\bar{x})) \), a measure
of the bending of \( u_i(\bar{x}) \) being the square of the Laplacian, \( \nabla^4 u_i(\bar{x}) = \left( \frac{\partial^2 u_i}{\partial x_1^2} + \frac{\partial^2 u_i}{\partial x_2^2} \right)^2 \). Let us call this
measure \( R \). The bending \( R \) is ideally zero, with transformations \( \bar{u} \) producing higher bending \( R \) being less
likely. \( R(\bar{u}) \) can be modelled this time as a Gaussian distribution, with mean zero and small variance, say
\( R \sim N(0,1) \). The transformation with zero bending is the identity map or some linear transformation of
it.

The model for \( \bar{u}(\bar{x}) \) is now

\[
\alpha (\nabla^2 u_1 + \nabla^2 u_2) = R
\]

(\( \alpha \) some scale factor) with Dirichlet\(^* \) boundary conditions for the edges \( x = 0 \) and \( x = 1 \) of the horizontal
map \( u_1 \) and the edges \( y = 0 \) and \( y = 1 \) of the vertical map \( u_2 \), and Neumann\(^1 \) boundary conditions for
the remaining sides of both maps.

We wish to solve (3.64) for \( \bar{u} \) and also to expand \( \bar{u} \) in an eigenfunction basis (the K-L expansion, \( R(\bar{u}) \)
being a Gaussian random field).

We take as our operator \( \tilde{L}^2 \) where \( \tilde{L}(\bar{u}) = \nabla^2 u_1 + \nabla^2 u_2 \), with the boundary conditions given above.
(Amit [1994] uses \( (\nabla^2 + \sigma)^2 (u_1, u_2) \)). The covariance matrix \( \hat{G} \) is constructed from the eigenvectors of
\( \tilde{L}^2 \):

\[
e_{nm}(x_1, x_2) = (2 \sin \pi n x_1 \cos \pi m x_2, 0)
\]
\[
e_{nm}(x_1, x_2) = (0, 2 \cos \pi m x_1 \sin \pi n x_2)
\]

and (repeated) eigenvalues

\[
\lambda_{nm} = \alpha \pi^2 (n^2 + m^2) \quad n, m \text{ positive integers}
\]

\(^* \bar{u} \text{ is specified at the boundaries.}\)

\(^1 \nabla \bar{u} \text{ is specified at the boundaries.}\)
The coefficients $C_{nm}$ are chosen, as in SPM, to minimise the variance of the posterior density
\[
e^{-\frac{1}{2}|\xi_{nm}|^2 - \frac{1}{2\sigma^2} (T(\theta) - S(\theta + \bar{\theta}(\theta)))}
\]

The smoothness criterion can be enhanced by using some even-degree polynomial of the eigenvalues, $p(\lambda_{nm}^2)$ instead of $\lambda_{nm}$ in the expansion (3.65). $p$ must be chosen to have no positive root, so that no real $\lambda_{nm}$ will give $p(\lambda_{nm}^2) = 0$, to avoid having an undefined expansion (zero in the denominator of (3.65)).

### 3.6.3 Other basis functions

Further examples of basis function expansions of transformations are given in Amit et al. [1991]; Gee et al. [1993]; Miller et al. [1993]; Grenander et al. [1992]; Christensen et al. [1993]. In particular, wavelet bases [Amit, 1994; Mallat, 1989b,a; Chui, 1992; Aldroubi et al., 1996] can also be used. These consist of dilations and translations of a generating function $h(x)$, giving a set

\[
h_{ab}(x) = |a|^{-\frac{1}{2}} h \left( \frac{x}{a} - b \right) \quad b \in \mathbb{Z},
\]

\[
a = 2^i, \quad i \in \mathbb{Z}
\]

We will meet Wavelets again in another context in chapter 4.

We now return to physical models, describing first the elastic and then the viscous fluid.

### 3.7 Model Three: Elastic Registration

The term elastic registration is occasionally employed to describe non-linear registration in general\footnote{For example, [Rohr et al., 1996] writes "A special class of general nonrigid transformations are elastic transformations which allow for local adaptivity and are constrained to some kind of continuity or smoothness. This contribution is concerned with elastic registration..." referring to his use of approximating thin-plate splines.}, but in its stricter sense the term refers to the use of the physical model of a continuous medium with elastic properties to impose smoothness constraints on the mapping [Bajcsy and Broit, 1982; Broit, 1981; Bajcsy et al., 1983; Bajcsy and Kovačič, 1989; Christensen, 1994]. An elastic medium is resistant to applied force; intermolecular forces or stresses ensure that the medium does not tear or fold and impose the same constraints which are required of the image mapping, namely that the transformations should be continuous, one-to-one and onto (so that pixels are mapped to unique locations, there are no discontinuities in
3.7. Model Three: Elastic Registration

3.7. Model Three: Elastic Registration

the mapping, and an inverse mapping exists). Driving forces, which force the source image to deform
until it matches the shape of the target, may be derived from similarity measures or difference functions
between corresponding features in the image pair, such as the surfaces at structure boundaries, which re­
quire a pre-deformation segmentation step. For an automated elastic registration, the driving forces are
derived from a statistical model of the likelihood of match between target and deforming source. We as­
sume the medium is isotropic, and so there is no preferential direction of forces, that is, forces of equal
magnitude will have the same effect regardless of their line of action (direction); and homogeneous (hav­
ing the same physical properties regardless of position within the medium), so forces of equal magnitude
will have the same effect regardless of their point of application.

The behaviour of the medium in terms of the deformations \( \tilde{u}(\tilde{x}) \) under external forces \( \tilde{f}(\tilde{x}) \) are described
by the Navier linear elastic partial differential equation (PDE)

\[
\mu \nabla^2 \tilde{u}(\tilde{x}) + (\lambda + \mu) \nabla \left( \nabla \cdot \tilde{u}(\tilde{x}) \right) + \tilde{f}(\tilde{x}) = 0
\]  

(3.66)

A derivation of the PDE is given in an appendix (section A.2). \( \mu \) and \( \lambda \) are Lamé's elasticity constants
and relate applied forces per unit area (stresses) to the resulting deformations (strains). Poisson's ratio \( \sigma \):

\[
\sigma = \frac{\lambda}{2(\mu + \lambda)}
\]

gives the ratio of lateral shrink of the medium to extensional strain when a tensile force is applied. Gener­
ally \( \lambda \) and hence \( \sigma \) are set to zero to ensure that deformations are only effected in the directions of applied
forces; \( \mu \) takes a value between 0 and 1. The forces are obtained from a similarity measure. A variety
of options exist: referring back to sections 2.4 and 2.5, these may be statistical pixel-intensity similar­
ity measures such as the sum of squared differences, or extracted features such as surfaces of anatomical
structures. (We have also mentioned the elasticity operator used as the smoothness constraint in an inter­
polating spline, section 3.5.2).

The PDE can be derived from two physical approaches, either by balancing forces or by minimisation
of energy (figure 3.10). The forces approach is concerned with balancing applied forces with internal
stresses to obtain the equilibrium situation where the material has settled into the shape which defers to
the applied forces, i.e. registers well, while also satisfying the elasticity constraints, i.e. obeys continuity
requirements. This approach is concerned with modelling the elastic behaviour of the medium so that
it imposes satisfactory continuity constraints on the transformations. The second approach depends on
the fact that the equilibrium position is at the lowest energy state of the material (to move away from
equilibrium requires doing work on the system and hence supplying it with more energy). This approach
is linked to the statistical model. The energy of the system is related to a probability distribution, with
the lowest energy state having the highest probability.

With \( \lambda \) set to zero, equation (3.66) can be seen as the regularisation problem (3.4), with \( \mu \) as the hyperpa­
rameter of a second order partial differential regularisation term which enforces smoothness. As well as
the Laplacian, the elastic smoothness functional includes the mixed derivatives \( \nabla \left( \nabla \cdot \tilde{u}(\tilde{x}) \right) \). Alterna­
tively, from a Bayesian viewpoint, the elasticity operator is a prior assumption of the form of transformation, and the body force is derived from the likelihood of the match. The various approaches to deriving the PDE are summarised in figure 3.18, which is an adaptation of figure 3.10 to the elasticity registration models.

**Figure 3.18: Variants in the derivation and solution of the mathematical model for elastic registration**

In elastic registration theory, the physical and statistical methods may be combined: the physical model to derive the prior constraints, and the statistical to model the likelihood of match given statistical factors such as noise processes in the target image. 'Elastic' registration can also be approached from a purely Bayesian statistical model [Gee et al., 1995b; Miller et al., 1993; Gee and Peralta, 1995]; the elasticity operator can then be substituted for a simpler measure of distortion, such as the square of the Laplacian alone [Amit, 1994], offering alternative choices of basis functions in an expansion of the deformation field. In this case, the 'elastic' registration is equivalent to the SPM model. See [Gee et al., 1997] for a
comparative evaluation of the two. Variants on the similarity measure are given in Gee et al. [1995b] and Gee et al. [1994a].

For a discrete spatial description of the deformation, the PDE may be solved using finite differences and a relaxation method such as successive overrelaxation (SOR) [Strikwerda, 1989; Press et al., 1995], possibly within a multigrid framework. A basis functions solution may be obtained as shown in section 3.6. Alternatively, the PDE may be solved for a subset of pixel locations distributed over the image, which are used as nodes in a finite element model of the deformation [Gee et al., 1994b,a]. Mappings of inter-nodal pixels are obtained from those of the nodes by finite element interpolation.

3.7.1 Historical development of elastic registration

The choice of the physical elastic model in registration originates with Broit [1981] and Bajcsy et al. [1983]. Broit specified the problem as arriving at the equilibrium state balancing internal and external forces by minimising the total cost function (3.4). Bajcsy et al. [1983] divided the images into equisized square regions; edge-shape features were identified in each region, and the similarity of features in corresponding regions in the images drove the registration, with weights to adjust the relative importance of each feature.

In a later paper Bajcsy and Kovacic [1989] solved the PDE (3.66) for a discrete deformation field \{\vec{u}\} where \((x_1, x_2, x_3) \mapsto (y_1, y_2, y_3) = (x_1 + u_1, x_2 + u_2, x_3 + u_3)\). They set \(A\) to zero and chose \(\mu\) to give the required elasticity: a large value gave a more rigid model, with reduced effect of external forces; a small value of \(\mu\) produced a more pliant model, that is, one which was more susceptible to noise in the images and hence highlighted error in the driving forces. The driving forces were proportional to the gradient vector of a local similarity measure \(D(\vec{x})\), chosen to be a normalised correlation of the projections of the images \(S\) and \(T\) onto a basis of orthonormal functions created from Hermite polynomials. A finite different approximation to the PDE was used for its solution. Additionally, a simple 3-step multiresolution approach was adopted, with the lower resolution images created by low-pass filtering (blurring) and resampling of the originals.

Amit et al. [1991] was the next advance in the history of elastic matching. The problem with using a purely discrete elastic registration is that it is very localised and so may suffer from local minima traps, which may be partially avoided by using blurred images as in the multiresolution approach of Bajcsy and Kovacic [1989]. Amit et al. [1991] introduced the concept of expanding the deformation field using some orthonormal basis, and matching in the spectral domain, allowing higher frequencies in the deformation to be postponed until the lower frequency global match had been found. This paper does not strictly describe an elastic match since the smoothing constraints were imposed by a simple Laplacian rather than the full elastic partial differential operator.

Gee et al. [1993] combined their previous approach [Bajcsy and Kovacic, 1989] with that of Amit et al. [1991], solving the elastic PDE for a basis-function expansion of the deformation field. Their similarity measure used only edge information from the surfaces of outer cortex and ventricles in initial trials, and

---

\[H_k(x)\] generated by the rule \(H_{k+1}(x) = 2xH_k(x) - 2kH_{k-1}(x)\) with \(H_0(x) = 1\) and \(H_1(x) = 2x\). Thus \(H_2(x) = 4x^2 - 2\) etc. They form an orthogonal basis. [Conte and de Boor, 1981]
they reported an improved registration if grey/white matter boundary information was also included. Of
interest is their experimentation with altering the value of the elasticity constant $\mu$, testing $\mu = 0.25$,
$\mu = 0.5$ and $\mu = 0.75$. They reported no significant difference in the matches obtained other than the
model with a higher value of $\mu$ being more susceptible to noise in the image data.

Miller et al. [1993] allowed nonlinearity in the driving forces, and introduced a more sophisticated hi-
erarchical approach: using the projection of the deformation field onto a set of basis functions as in Amit
et al. [1991], they initiated the registration using a highly truncated basis set and increased the number
of basis functions after registration had been achieved maximally for the current set. They chose as their
distance measure

$$ \frac{1}{\sigma^2} \int_\Omega (S(\tilde{u}(\tilde{x}) - T(\tilde{x}))^2 \, d\tilde{x} $$

($\sigma$ being the variance in the likelihood model) which in discrete form is the sum of squared differences.
They also introduced the Bayesian approach into the registration problem, with the optimal deformation $\tilde{u}$
being the mean of the posterior distribution. Statistical concepts were further utilised by the introduction
of a random noise term into the gradient of the posterior distribution, allowing a stochastic gradient search
as a further defence against local minima traps.

The chosen basis for their implementation was obtained from the eigenfunctions of the elasticity operator
(in a similar way to section 3.6.2. The eigenfunctions are listed in appendix A.3). These are functions
chosen to diagonalize the covariances associated with the elasticity operator (so that the covariances are of
the form given in equation (3.58)). They used a combined hierarchical approach, first obtaining a global
match by matching in the spectral domain, as in [Amit et al., 1991], and then proceeding to optimise in the
spatial domain for the most localised match. The method was applied to 2D images; later [Christensen
et al., 1994, 1993] they extended it to the full 3D application. In these papers they used $\lambda = 1$.

Gee et al. [1994b] used a finite element approach with isoparametric shape functions (section 3.2.2).

Christensen in his doctoral thesis [Christensen, 1994] proposed an Eulerian reference frame, where pixel
movements are monitored according to their final positions, as opposed to the Lagrangian reference frame
where pixel movements are tracked from their initial locations. Using an Eulerian reference frame auto-
matically solves for the backwards mapping (section 3.3.2), scanning along each pixel in the final de-
formed source image and locating its origin in the original source image. The similarity measure was
optionally fortified by dictating approximate landmark locations to guide the match. The regularisation
parameter was modified by a variance term derived from the similarity measure, in a similar way to SPM
(section 3.6.1). The finite element implementation also allows a hierarchical approach (chapter 4), with
initial registration using large elements, and hence converging rapidly, and the elemental size gradually
decreasing to improve on localisation of the deformations.

Gee et al. [1994a] applied a variant of the finite element match, using features such as edge and curvature

\footnote{Obviously likelihood information should not be obtained from nodal positions alone, since then the deformations are more susceptible to noise at the nodes and local minima; perhaps this is why their next paper [Gee et al., 1994a] tried edge and curvature information instead.}
information in the similarity measure, instead of pixel information at the nodes and landmarks.

A set of three papers in 1995 examined the Bayesian statistical approach to elastic matching. [Gee and Peralta, 1995] analysed the prior as the physical elasticity analogue. [Gee et al., 1995b] and [Gee et al., 1995a] referred to decision theory [Berger, 1985] and discussed the information provided by the posterior distribution: The optimal solution minimises the expected loss with respect to the posterior distribution, with options for the loss function. The maximum a posteriori (MAP) solution minimises the zero-one loss function, with minimum for this loss being at the maximum of the posterior distribution. Alternatively, the minimum mean squared error (MMSE) solution minimises the squared loss, and the optimum corresponds to the mean of the posterior distribution. They reported in [Gee et al., 1995b] no conclusive evidence to suggest that either is generally superior.

Summary

The elastic matching registration thus has developed over time to a more general form: a statistical likelihood function (such as correlations or sum of squares) drives the match, and a prior smoothing constraint (usually based on a physical model, such as the elastic differential operator - minimising extensional and shear strains, or the Laplacian - minimising bending) are selected to build the mathematical model, expressed as a partial differential equation in the deformations. The minimisation may be carried out in the spatial domain - as in the earlier attempts at elastic registration, or in the spectral domain, where the deformation field is projected on to an orthonormal basis, derived from the choice of smoothing operator (so that the functions with lowest eigenvalues correspond to the lowest bending/stretching deformations). To address speed and local minima considerations, a hierarchical approach is advised. We explore this suggestion in chapter 4. Early hierarchical methods matched within scale space; more successful seems to be the initial use of the spectral domain matching, with the number of basis functions being increased, and with the final matching at the highest resolution in the spatial domain. Alternatively, a finite element approach is adopted, with the element size being gradually reduced.

The statistical approach has several benefits. First, the assignment of probabilities to different energy states means that higher energy states, although having low probabilities, are not impossible, and serve as an escape route from local minima. Also the problem can be easily modelled as a Markov random field, and so associated statistical methods can be applied in the solution of the problem. Finally, the outcome has a probability distribution whose variance can be calculated, giving an estimate of the accuracy of the match.

3.8 Model Four: Fluid Registration

3.8.1 The fluid model

The physical elastic model has the disadvantage of disallowing extensive localised deformations due to the modelling of stresses that increase proportionally to the deformation strength. An alternative contin-
uous medium model is that of a viscous fluid [Christensen et al., 1995; Christensen, 1994; Freeborough and Fox, 1998], in which these internal forces relax as the source image deforms over time into the shape of the target. The fluid model therefore permits severe localised distortions including the formation of sharp corners and the growth of new regions from the edge of a structure. A resulting danger is that the opportunity for misregistrations is increased, generally involving the growth of one region instead of a shifting and/or distortion of another (figure 4.3). Additionally the freedom of the fluid model may not be ideal in some clinical situations, for example in pre-postoperative data where the fluid registration may attempt to reconstruct the removed tissue. The fluid model allows the creation of an object of any shape as long as there already exists within the image a minimal seed region of the required intensity from which the growth can proceed. These deformations cannot be mimicked by the thin-plate spline or any deformation formed by ultimately applying a single linear combination of basis functions, which can reproduce neither large and sudden localised region growths, twists and tears, nor create deformations with sharp corners - all of these deformations being excluded from the space of transformations of any of the truncated sets of basis functions.

The fluid model is based on the Navier-Stokes fluid PDE (equation 3.67) which is similar to equation (3.66), except that the elastic differentials operate on a velocity field \( \mathbf{v} \) describing pixel movements over time:

\[
\mu \nabla^2 \mathbf{v} + (\lambda + \mu) \nabla (\nabla \cdot \mathbf{v}) + f(\mathbf{u}(t)) = 0
\]  

(3.67)

\( \mathbf{u}(t) \) are pixel displacements describing the deformation at time \( t \); \( \lambda \) is generally set to zero and \( \mu \in (0, 1) \).

The driving forces \( f(\mathbf{u}(t)) \) are obtained from gradients of the sum of squared differences dissimilarity measure (equation 2.2):

\[
f(\mathbf{u}(x, t)) = \nabla (T(x) - S(x - \mathbf{u}(x, t))^2)
\]

\[
= -2S(x - \mathbf{u}(x, t)) (\nabla T)(x - \mathbf{u}(x, t))
\]  

(3.68)

Computationally, the algorithm contains nested iterations. Overall, registration proceeds through a set of timesteps. At the start of each timestep, the PDE (3.67) is updated using the current state of the deformed source to compute the force field and is then solved iteratively using the following numerical technique: equation (3.67) is expanded by finite differences using a scheme similar to that written out in appendix A.5 and solved pixelwise over the discrete image lattice by the iterative successive over-relaxation (SOR) method [Press et al., 1995]. This provides a continually modified velocity field \( \mathbf{v} \) which specifies the direction and magnitude of displacement-change \( \Delta \mathbf{u} \) for each pixel within the source image for the current timestep. Should the transformation in the neighbourhood of some pixel approach a singularity, regidding is applied globally as follows: the current deformation-change field is concatenated to the total deformation field and then reset to zero, and a new current ‘source’ image is obtained from the original source image by applying to it the concatenated deformation field \( \{ \mathbf{u} \} \). The continually-modified source is as a result seen to ‘flow’ over time into the shape of the target.

Recall the linear optical flow registration in section 2.6, of which the search strategy was to follow pixel intensity gradients over time. In this respect the fluid model is a highly non-linear equivalent. Although
the derivation of the forces in equation (3.68) is from the dissimilarity measure, the search strategy is to recompute the force field at each iteration and move slowly along gradients in the deforming source.

### 3.8.2 Derivation of the fluid PDE

Equation 3.67 is the linearised elasticity operator of equation 3.66 applied to the velocity field of the moving fluid, rather than directly to the field of displacements. Derivation of the elasticity operator is written out in appendix A.2. For the fluid model, we summarise here the main points in the derivation of its PDE, to which we will return in section 6.5.2.

The fluid PDE is obtained by balancing external forces $\vec{f}$ with internal stresses which are described by a stress tensor, which in form is a matrix of first derivatives of the deformation field: the smoother the deformation, the less rapidly it fluctuates and hence its derivatives are of smaller magnitude.

The viscous fluid PDE (3.67) is summarised by

$$\nabla \cdot \vec{\sigma} + \vec{f} = 0$$

(3.69)

where $\vec{\sigma}$ is the Cauchy stress tensor. For a regular viscous fluid, this is given by

$$\vec{\sigma} = -p\vec{I} + 2\mu\vec{D}$$

(3.70)

where $p$ denotes pressure acting on the fluid, $\vec{I}$ is the identity matrix and $\vec{D}$ is the rate of deformation tensor,

$$\vec{D} = \frac{1}{2} \left( \nabla \vec{u} + (\nabla \vec{u})^T \right)$$

(3.71)

The stress tensor can be given by its components:

$$\sigma_{ij} = -p\delta_{ij} + \mu \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right)$$

(3.72)

Christensen [1994] presents a special version of the model to include mass creation: the PDE is reformulated, by inserting into the mass conservation equation a mass creation term $\eta$ (which is later neglected as one of two inertia terms), and by adding a volume dilation term to the stress tensor which we now denote $\vec{\sigma}_C$:

$$\vec{\sigma}_C = (\lambda(\text{tr}\vec{D}) - p\vec{I}) + 2\mu\vec{D}$$

(3.73)

where $\text{tr}\vec{D}$ is the trace of $\vec{D}$ (sum of terms on its leading diagonal).

Inserting (3.73) into equation (3.69) gives

$$\mu\nabla^2 \vec{u} + (\lambda + \mu)\nabla(\nabla \cdot \vec{u}) + \vec{f}(\vec{u}(t)) = \nabla p$$

(3.74)

The pressure gradient $\nabla p$ prevents the fluid from changing its density, and is neglected in [Christensen, 1994]. Setting $\lambda$ to zero (as is done in the successive overrelaxation algorithm in [Christensen, 1994]), and so removing the volume dilation term, still allows regions of the deforming image to change size, due to the omission of $\nabla p$. The density of regions is artificially maintained by the regridding technique, where
the 'fluid' consists of pixels at constant density determined by the grid sampling, and intensities of mapped point locations are preserved rather than preserving sums of intensities of mapped regions. Hence mass creation in the fluid model is actually accomplished by regridding the deforming source several times during the registration.

The original fluid registration algorithm in [Christensen, 1994] using solution by finite differences is slow\(^1\). There are numerical improvements to SOR such as multigrid methods [Yavneh, 1996; Brandt and Mikulinsky, 1995; Zhang, 1996] but they are beyond the stated scope of this work. Bro-Nielsen and Gramkow [1996] give a faster version where (3.67) is solved by deriving a convolution filter from the eigenfunctions of the linear elasticity operator \(\mu \nabla^2 + (\lambda + \mu) \nabla (\nabla \cdot \cdot)\) with sliding boundary conditions (which are reproduced in the appendix A.3). As with any convolution method, it can only be applied if the operator is constant over the image, which will not be the case if the viscosity \(\mu\) is allowed to vary, as presented in section 6.5.2. Furthermore, if additional boundary conditions are introduced, derivation of the convolution filter is more complex, and cannot be accomplished automatically within the fluid algorithm.

### 3.9 Summary

In this chapter we have investigated one statistical and three physical models which form the conceptual basis for non-linear registration algorithms. We have analysed the derivation of the practical implementation of non-linear algorithms in general via mathematical models, using Bayesian statistics, regularisation theory and the balancing of forces in physical models.

This completes our survey of single-level registration algorithms. The choice of title for this chapter is explained by that of the following chapter, in which the algorithms described change in method or application up levels of a hierarchy during the course of the registration.

---

\(^1\) and was implemented by its authors using parallel processing.
Chapter 4

Hierarchical Methods
4.1 Introduction

Equation 3.4 describes registration as an optimisation problem, where the deformation giving the global minimum of the surface of a cost function is to be found (figure 1.5). Due to time constraints, the registration algorithm rarely performs an exhaustive search of all possible deformations. Instead it employs a search strategy such as one of the standard gradient descent methods. A common problem is that a local minimum in the cost function surface is mistaken for the global minimum, causing premature termination and hence misregistration, since corresponding structures have not been aligned. This is the local minimum problem.

The likelihood of falling into local minima is reduced if either some of the local minima are removed or if the search is commenced close to the global minimum. Local minima are removed, giving a smoother cost function surface, by one of two methods. In the first method, details are temporarily removed from both source and target image, leaving only larger, more global, corresponding structures. Since there are now fewer structures to be matched, the possibility of mismatch is reduced. Many local minima are caused by the extreme distortion of a structure in the source to match a nearby but non-corresponding structure in the target which is of similar intensity but of greatly differing shape. Hence the second method for avoiding local minima is to initially restrict the range of allowable deformations to those of a simpler, more global nature. In either case, registration within the simplified cost function surface is more likely to find the global minimum. Complexity in the image data or in the allowable deformations can then be re-introduced. The global minimum of the simplified cost function surface is close to that of the original, more detailed, surface, and so the transformation giving the global minimum of the former is used as a starting estimate for the search within the latter.

Thus the search for the optimum deformation is split into stages using a hierarchy either of data complexity or of deformation complexity; at each level in the hierarchy the search is initiated at a location closer to the global minimum. An alternative choice is to precede the chosen non-linear algorithm with methods which are more global in nature. This similarly increases the likelihood that the final model employed commences its search close to the global minimum.

Figure 4.1: Data set of (left) head1A, MRI head scan and (right) head1B, a warped version of head1A

Whereas non-avoidance of local minima only results in premature termination of a linear registration, in the non-linear case unnecessary distortions result. The effect is most severe when using a highly local
algorithm, such as the fluid deformation on a discrete pixel description of the mapping, where the constraints on strong localised deformations are weaker.

We demonstrate such a deformation using two data sets of 2D images sized 256 × 256. The first, labelled head1A and head1B, is an image pair consisting of an MRI head scan, and an image produced artificially from it by applying a thin-plate spline with 20 landmarks (figure 4.1 and see section 5.3). The second set, head2A and head2B, are MRI head scans of two different subjects at roughly the same slice location (figure 4.2). Head2A and head2B were supplied by John Ashburner of the Functional Imaging Laboratory*.

Figure 4.3 shows the fluid registration of head1B to head1A, where an initial misregistration of cortex to scalp boundary, with a subsequent eradication of the original scalp in the deforming image, is followed by the growth of a false scalp from the cortex and from two newly created edges in the scalp radial to the cortex. Figure 4.4 shows the first three corresponding difference images between head1A and the deforming head1B. The cause of initial misregistration is the confusing proximity of the cortex-skull boundary in head1B to the scalp-background boundary in head1A, both of which exhibit similar contrasts. Figures 4.5 and 4.6 show a similar misregistration of head2B to head2A.

* Institute of Neurology, Queen Square, London.
Figure 4.4: Subtraction images of the first, second and third deformed images in figure 4.3, with (left) source image, showing the cause of misregistration: the overlap of different structures with similar contextual contrast.

Figure 4.5: The first 8 timesteps in the fluid registration of test image head2B to head2A showing misregistration of scalp to cortex.
Figure 4.6: Difference image of head2A and head2B showing overlapping cortex-skull and scalp-background boundaries, the source of the confusion causing misregistration.
An additional advantage of hierarchical methods is that the earlier computations can often be performed rapidly, due to a reduction in the quantity of input data, or in the calculation of a simpler transformation.

Stochastic methods are another option to help escape these local minima traps - Kim et al. [1997] restart the search from a nearby pixel location after a minimum is found; Miller et al. [1993], Wells et al. [1996] introduce noise to the gradient of the optimisation surface; Amit et al. [1991] and Aykroyd and Mardia [1996] avoid a directed search altogether and opt for an MCMC (Markov Chain Monte Carlo, [Ripley, 1996]) method; Hill et al. [1993] have a choice of two methods both with a high level of randomness: a genetic algorithm or a stochastic simplex.

4.2 Increasing Data complexity

The first hierarchical method entails the initial reduction of information content within the images to be matched so that only the coarsest, most global structures remain; the absence of finer detail ensures avoidance of local minima traps by virtue of their eradication - ideally the optimisation surface initially contains only one maximum, in the region of the optimal global match. The optimisation can then be improved by gradual re-introduction of structural details at the same rate in both the images and the concurrent updating of the registering deformation to match corresponding re-introduced structures or features. Note the assumption: corresponding structures are re-introduced in the image pair at each level.

Thus this method requires the parallel generation of a family of images from each of the original source and target, prior to registration, where each successive member of the family contains an appropriate corresponding reduction in structural detail. It may be considered as a type of image pre-processing, in which copies of the images themselves are altered to aid an optimal match, rather than using a more intelligent matching algorithm which registers using its own perceived hierarchical description of the registration optimisation surface.

Such hierarchies of decreasing data complexity are provided by scale spaces, where the image size is constant in all levels, or by pyramids, where image size is also reduced in each successive level. In the latter case, an additional advantage for pixel-based matching schemes is that computation of optimal parameters is speeded up in the higher level (lower resolution) images due to the reduction in the amount of data to be processed.

Gaussian scale space

A discrete scale space consists of a set or stack of images which are increasingly simplified versions of the input image (fig. 4.7). In the Gaussian scale space, the images are increasingly more blurred versions of the original. Figure 4.8 shows selected slices of the isotropic Gaussian scale space of test image headA.

Gaussian scale spaces [Witkin, 1983; Babaud et al., 1986; Lindeberg, 1994; Nielsen et al., 1994; Florack, 1993] are generated by convolving an input image with a Gaussian smoothing function with successively higher values of standard deviation equal to $\sqrt{2i}$, where $i$ denotes the level in the scale space.
4.2. Increasing Data complexity

Figure 4.7: Scale space stack generated from an input image.

Figure 4.8: Slices of the isotropic Gaussian scale space of head1A, at scales of (left–right) \( \sigma = 2, 6, 15 \) and 20 pixel units.

Since the Gaussian is the Green’s function of the diffusion operator, the isotropic scale space stack consists of solutions for pixel intensity \( I \) of the partial differential equation

\[
\frac{\partial I}{\partial t} = \nabla^2 I
\]

at given values of \( t \). This is a special case of the heat diffusion equation

\[
\frac{\partial I}{\partial t} = \nabla . c \nabla I
\]

where the constant \( c \) is the conductance, which controls the rate of diffusion of heat \( I \).

Thus Gaussian smoothing is due to the dissipation of high pixel intensities into surrounding regions of lower intensity in the same way as a hot body cools gradually by spreading its heat.

An example of the Gaussian scale space used as the framework for a piecewise-linear method is in the algorithm of Collins et al. [1994], which consists of five stages:

1. an initial rough registration by principal axes;
2. cross-correlation matching using pixel intensities in the images smoothed with a Gaussian of \( \sigma = 8 \) mm
3. cross-correlation matching using pixel intensities in the images smoothed with a Gaussian of \( \sigma = 4 \), using the whole head images
4. cross-correlation matching using pixel intensities in the images smoothed with a Gaussian of \( \sigma = 4 \), using masked brain images
4.2. Increasing Data complexity

5. cross-correlation matching using intensity gradients, on the masked images.

The last two steps use masked images to avoid possible bias by a differential scalp thickness between the two subjects. Additionally it should avoid the common trap of scalp being registered to brain. However the initial registrations may have been on the whole head images to aid localisation of the scalp for removing it. The final step, of registering by gradient strengths, is to correct for bias by possible intensity inhomogeneities due to a varying field in the scanning device, which, as slowly varying, are not so apparent in the intensity gradients.

At lower resolutions, the optimisation surface is smoothed, so local minima are obliterated and the search can find the (approximate) location of the global minimum. Note that features such as ridges and minima are not stationary within the Gaussian scale space [Griffin and Colchester, 1995], so that the maxima/minima locations at lower resolution are initially approximate. At higher resolutions, features with smaller scale are reintroduced, enabling finer tuning of the matches. At each jump from one level to the next, the cost can be expected to increase due to the appearance of these additional features which generally will be slightly out of register. This will force the finer tuning. The next-resolution jump may be accompanied by a reduction in the search region for each parameter, to force convergence to the minimum. Alternatively, occasional large random displacements may be tried to prevent the search from settling into a local minimum.

Location of the global minimum or maximum is easier at low scale not only since “tempting” local minima have been removed, but also due to the widening of the peak of the maximum or trough of the minimum. This is especially true in the case of ridge-matching when registering sparse features such as sharp crest lines. The search space may then be almost uniformly flat, with the crest lines producing occasional sharp peaks. If the initial estimate is not given close to the (true) peak, the search may not make any considerable progress as it usually relies on gradients in the search space to drive it to the solution. In a lower resolution, the peaks will be smoothed to more gradual gradients covering a larger catchment area, making more likely the earlier commencement of descent to the global minimum.

Gaussian blurring not only weakens all edges but also moves them, as described by Griffin and Colchester [1995] (and see also [Bruce et al., 1996; Lifshitz and Pizer, 1990]). This problem was noted by Lifshitz and Pizer [Lifshitz and Pizer, 1990] in the case of segmentation: a pixel can change classification over scale space if located near a boundary region. In addition, the non-linear image registration methods such as SPM [Friston et al., 1995] and the elastic [Bajcsy and Kovačič, 1989; Gee et al., 1993; Miller et al., 1993] and viscous fluid [Christensen et al., 1995] models are driven by intensity gradients. It may therefore be preferable for a scale space to selectively retain edge strengths of detail being matched at that scale. These are examples of inhomogeneous scale spaces.

Inhomogeneous scale space

Inhomogeneous scale spaces can be generated by modifying the conductance term [Perona and Malik, 1990; Whitaker and Pizer, 1993], for example by forcing it to be lower at region boundaries and greater
in homogeneous regions.

The conductance term \( c(r, t) \) now varies in location both within the image and generally also through scale space. Perona and Malik [1990] introduced this concept by reasoning that, to retain edge sharpness, \( c \) should be a function of 'edgeness', or magnitude of gradient, monotonically decreasing with increasing gradient size [Whitaker and Pizer, 1993]:

\[
\frac{\partial I}{\partial t} = \nabla \cdot g(|\nabla I|) \nabla I
\]

where the conductance function

\[
g(|\nabla I|) = e^{-|\nabla I|^2/k^2}
\]

contains a constant \( k \) chosen empirically to give an appropriate response to edge strengths.

Note that generally an inhomogeneous scale space can no longer be generated by convolutions with Gaussian kernels (or equivalently by multiplication by Gaussians in the Fourier domain) since the blurring is no longer of the same strength everywhere in the image; instead a finite differences scheme is used to propagate the solution through time.

Whitaker and Pizer [1993] note that the inhomogeneous scale space as presented above requires for its generation knowledge of gradient strengths in the input image, for insertion in the conductance function. However, the calculation of gradients is an ill-posed problem - it is strongly affected by the presence of noise, and so should ideally be calculated on smoothed images. Hence they present an adaptation of the Perona-Malik diffusion where the gradients are calculated on an (isotropically-)Gaussian-blurred version of the current image:

\[
\frac{\partial I}{\partial t} = \nabla \cdot g(G(s) * |\nabla I|) \nabla I
\]

Additionally, they introduce a further scale parameter, \( s(t) \) into the Gaussian blurring, which is thus reduced as the Perona-Malik diffusion progresses, reflecting the consequential reduction of noise by the diffusion within the more homogeneous regions. The full inhomogeneous edge-affected Gaussian scale space is then:

\[
\frac{\partial I}{\partial t} = \nabla \cdot g(G(s(t)) * |\nabla I|) \nabla I
\]

Whitaker and Pizer [1993] suggest \( s(t) \) to be a decreasing linear function of time \( t \).

---

Figure 4.9: Application of anisotropic blur to head1A with conductance given by equation (4.2); (left-right): timesteps \( t = 1, 4, 7 \) and 10.
4.2. Increasing Data complexity

An alternative conductance term is

$$g(|I|) = -\frac{1}{1 + \frac{|\nabla I|^2}{k^2}}$$ (4.3)

This produces the blurs shown in figure 4.10.

Figure 4.10: Application of anisotropic blur to head A with conductance given by equation (4.3); (left-right): timesteps $t = 1, 4, 7$ and 10.

There are many other choices for the conductance term $c$, depending for example on higher derivatives of the image $I$. For more complete descriptions refer to [Vincken, 1995; Florack, 1993; ter Haar Romeny, 1994].

Inhomogeneous scale spaces allow edge strengths to be selectively reduced in areas of high detail only. Whitaker and Pizer [1993] highlight two potential problems with the Perona-Malik diffusion. Firstly, as already mentioned, the choice of $k$ in the conductance function (equation 4.1) must reflect the expected gradient strengths in the image. This choice is not simple if the image contains a wide range of edge strengths at structure boundaries. Secondly, a ringing phenomenon termed ‘staircasing’ sometimes manifests where edges are wide and smooth. Furthermore, the edge-affected diffusion process is not always well behaved. Diffusion at a boundary region consists of three stages: initially, the edge gradients increase while surrounding homogeneous regions are smoothed; after reaching some upper bound the edges do not stabilise but start to decay, slowly decreasing in edge strength; finally, after reaching some lower threshold of edge strength (below which the region is no longer classified as an edge but falls under the category of ‘noise’) the decay or leakage is rapid, the process in this region reverting to an ordinary Gaussian smoothing. This process is more complicated for an image exhibiting several structures of different contrasts, which at some moment in the scale space will be at different stages in this process. Structures with weaker contrast (i.e. those bounded by regions of a similar intensity to that within the region) will reach the decay stage of diffusion while structures with higher contrast are still within the initial stage of increasing edge strength. This may be a particular problem in registration if the contrast of larger or more global structures relative to their context is weak, while that of a few smaller possibly irrelevant structures is strong. Additionally this scale space still fails to address the problem of misregistrations such as those illustrated previously.

Figures 4.11 and 4.12 show leakage across an edge on further application of inhomogeneous blur following that in figure 4.9.
4.2. Increasing Data complexity

Figure 4.11: Application of inhomogeneous blur with conductance given by equation (4.3); timestep 20 with an edge region marked by a box

Figure 4.12: Continued application of inhomogeneous blur (4.3) in the box region marked in figure 4.11, timesteps 20-24, showing leakage across edges.

Morphological scale space

Gaussian scale space has classically been used in image analysis because it satisfies the causality criterion, i.e., that each structure at a particular level can be traced to a cause at some lower resolution, or roughly that no new structure is generated as the diffusion proceeds (this is only strictly true for one-dimensional images; for higher dimensional images causality is violated in some special cases [Lindeberg, 1994]). If this criterion is relaxed in favor of greater emphasis on edge-preservation and edge-location-preservation, then an alternative choice of scale space is one generated by mathematical morphology, such as the closure operation. Closure [Haralick et al., 1987; Preteux, 1992; Maragos, 1996; Sivakumar and Goutsias, 1996] has has two distinct advantages: firstly, an object boundary is neither smoothed nor shifted; secondly, it fills gaps to group objects into larger structures, thus the scale space contains a hierarchy of objects which is more intuitive to the registration criterion. The advantage of closure over median filtering is that peaks in the intensity surface are not levelled, but remain to drive the
Greyscale closure is the process of dilation followed by erosion by means of a mask known as a 'structuring element'. The outcome of the closure operation depends on the shape of the structuring element. For image registration, it is desirable that the scale space will not be biased in any particular direction, and thus an isotropic structuring element should be selected. To retain the sharpness of edges, the structuring element itself must have a sharp edge. The appropriate structuring element therefore for 2-D images is a disc, and for 3-D images a sphere. The scale space is created by application of the closure operation with successively larger structuring elements; at each level all detail finer than the size of the structuring element will be eradicated. Trough-like detail, consisting of concavities in the intensity surface, will simply be filled in. However, peak-like detail, or convexities, is retained, although it may be spread across a level region shaped by the structuring element, creating a 'false structure'. Figure 4.13 illustrates the application of closure to test image head1A, using discs of radius 20 and 5 pixels. For further discussions on morphological scale space, see [Alvarez and Morel, 1994; Jackway, 1996; Jackway and Deriche, 1996].

![Figure 4.13: Closure on test image head1A. (left) using a structuring element of radius 20 identifies two objects: scalp and 'inside-scalp'. (right) a structuring element of radius 5 is less successful at identifying 3 objects: scalp, skull and brain. False structures shaped by the structuring element are evident.](image_url)

The Gaussian Pyramid

A third type of data hierarchy is provided by the Gaussian pyramid [Lindeberg, 1994; Burt, 1981; Crowley, 1981], where each successive level is formed by applying Gaussian smoothing with increasing scale and downsampling the previous level. Also of interest is the Laplacian pyramid [Lindeberg, 1994; Burt, 1981; Crowley, 1981], which is approximated by the differences between successive levels in the Gaussian pyramid. The Gaussian pyramid was used for example by [Christensen et al., 1994] in the application of his fluid deformation [Christensen et al., 1995; Christensen, 1994] as a final stage after registration by the elastic (statistical basis-function) model. Figure 4.14 shows fluid registration of head1B to head1A in the first level of the pyramid (images downsampling to size 64 \times 64), at the termination of which the local minima at the scalp-skull-cortex boundaries have been bypassed.

The advantage of a pyramid is that the amount of processing in the initial stages is minimised due to the smaller image sizes; this is particularly appreciated in a slow algorithm such as the pixel-based fluid. The disadvantage in the case of fluid registration is that upsampling, that is, sampling from a coarser to a finer...
grid, is required between levels: the transformation calculated for the previous level (of lower resolution) must be interpolated into a higher-resolution grid to form the initial transformation for registration of the larger images in the next level. This may introduce block-shaped artefacts into the deformation and velocity fields. A similar problem is met when computing basis function coefficients using multigrid methods, where the computed solution to the numerical problem is resampled many times between grids of different levels of resolution. Amit [1994] suggests that in this case bilinear interpolation used in inter-level propagation may introduce unnatural deformations. This demonstrates that the three basic hierarchical types can not always be mixed - in this case hierarchies of warp and of data complexity may be incompatible. A Gaussian pyramid is used by Gee et al. [1994b] in their hierarchical finite element elastic registration. Here interpolation between levels is provided by the isoparametric shape functions [Hinton and Owen, 1977] which interpolate the deformations at nodes to inter-nodal locations. In this case the method neatly unifies two hierarchical types, namely increasing data complexity (using scale space) and increasing warp complexity (using an increasing number of smaller elements). In respect of the latter we will return to their method in section 4.3.2. Woods et al. [1992, 1993]; Alpert et al. [1997] use a simplification of a pyramid representation of the image data to speed up their (linear registration) algorithms for a 3D image: an approximate transformation is first computed using a sampling of every 81^{st} pixel pair in the source-target image pair; the transformation parameters are then refined using every 27^{th}, every 9^{th}, then every 3^{rd}, followed by finally the full image pair. Use of an image pyramid where subsampling is preceded by smoothing would reduce susceptibility to noise of the earlier registration levels.

Spline Pyramids

A refinement to the Gaussian pyramid is provided using spline pyramids by Unser et al. [1991, 1993b,a] and we summarise their discussion here. Spline pyramids are generated [Unser et al., 1993b] by successive application of steps consisting of prefiltering and subsampling. The prefilter is a spline interpolating function which is convolved with the image to obtain a representation of the image by coefficients of the spline. Unser et al. [1993b] gives a choice of four splines: the basic (B-spline), dual, cardinal or orthogonal cubic spline. Subsampling to the lower resolution is then performed by reading off values at pixel locations from the spline representation of the image. To move from a lower to a higher resolution, upsampling is followed by application of an appropriate postfilter, which is the spline interpolation function associated with the pyramid. The basic pyramid (for which the interpolating function is the cubic
B-spline) is the most appropriate to use in tasks requiring easy interpolation. The dual, named since the roles of pre- and post filter are the reverse of those of the basic, is the most simple to construct. The cardinal pyramid shows edges sharper than the dual; hence it is best used for computing derivatives and integrals. The B-spline approximates the Gaussian for increasing order \( n \) and hence levels in the basic pyramid exhibit slight blurring, whereas the cardinal spline in the limit tends to the sinc (the ideal interpolant) and hence has greater fidelity to the original image. The orthogonal cubic spline has the advantage that its prefilter and interpolation function are identical - practically this obviates the need for pre-filtering; however, it is not an exact interpolant and so reconstructed images are not true to the original.

Hierarchical registration schemes based on these polynomial splines are restricted in the literature to affine transformations only [Unser and Aldroubi, 1993; Thévenaz et al., 1995, 1998]. The optimum affine transformation parameters in a given level are propagated to the higher-resolution level by scaling (to account for the larger image size). Thévenaz et al. [1998] highlights an advantage of inserting the registration method within a spline pyramid hierarchy, which provides a consistent filtration and interpolation model for generating the data hierarchy, computing derivatives, applying transformations during registrations, and interpolating the final transformed image. Centered pyramids [Brigger and Unser, 1997] enable more accurate inter-level upsampling of the transformation parameters, since the location of the origin is preserved throughout levels.

Wavelet Pyramids

Wavelets have been introduced in section 3.6.3 as basis functions from which a deformation is constructed. Wavelet functions are used in another context to describe either the redundancy of information in an image or its detail. When this process is repeatedly applied, a pyramid is formed.

Wavelet pyramids are related to spline pyramids, in that they are likewise generated by successive filtering and subsampling. Pyramids can describe hierarchies of data in two ways: in the first paradigm, the pyramid levels consist of successively simplified versions or approximations of the input image (as in the Gaussian or spline pyramids), in which case all information in a lower resolution is contained within a higher resolution level. In the second paradigm, successively lower levels store only the detail, i.e. the additional information needed to generate a higher-resolution image; Laplacian pyramids [Sonka et al., 1993] are an example. In the construction of wavelet pyramids [Mallat, 1989b], both types are generated concurrently - hence the wavelet transform is a decomposition into approximation and detail images. Generally only the detail images are retained; an exception is the scheme used by Deubler and Olivo [1997]. For each dimension in the image two wavelet functions are used as pre-filters before subsampling. The first, a band-pass filter, generates the detail images; the band-pass filters are chosen so that at each successive level, detail of lower frequency is isolated. The second filter is derived from the band-pass filter and is used to generate the approximation image. The approximation image is recursively split into the detail and approximation signals of successive levels, until the lowest resolution level contains solely an approximation image.
4.3 Increasing Warp Complexity

Deubler and Olivo [1997] uses such a wavelet pyramid in a hierarchical non-linear registration scheme. The scheme itself concentrates on interaction with the wavelet pyramids of source and target. Registration starts at the highest pyramid (lowest resolution) levels by block-matching the approximation image pair. Within each subsequent level, matching is driven by differences in a feature space derived from approximation and detail images of corresponding levels in the source and target pyramids; horizontal and vertical displacements are estimated independently using the horizontal and vertical detail images respectively. Propagation down the pyramid hierarchy of the displacement fields is achieved by upsampling without interpolation. Such a scheme is not continued into the very highest resolution level since this consists of one image pair only (the original source and target). Instead, the displacement field obtained in the penultimate level is applied to the wavelet decompositions of the original source and the inverse wavelet-transform applied, giving a deformed source image.

4.3 Increasing Warp Complexity

The second hierarchical method is to increase the complexity of the transformation within a particular algorithm. This method can be applied (in different ways) to two subsets of algorithms: warps summarised by function coefficients and warps summarised by displacements of nodes or landmarks. In both the transformation is analytic, that is, a mathematical function describes the warp, as opposed to a discrete deformation field describing the mapping of each pixel individually.

4.3.1 Warps summarised by function coefficients

This subset includes the use of wavelet expansions and warps expanded in terms of a truncated set of basis functions (see for example Amit [1994] who uses both methods), which were described in their single-level implementation in section 3.6. In this case, the warps are summarised by a set of coefficients to the basis or wavelet basis functions, as in equation (3.2). The set of basis functions is ordered by increasing frequency, and thus coefficients of functions early in the list summarise more slowly varying, underlying trends in the warp, which are the more global descriptors of the transformation. Coefficients of higher frequency functions, appearing later in the expansion, describe more localised, finer detail in the warp. The hierarchical method is to start with only the first few (or one) basis functions and search for their coefficients which optimise the registration, then gradually increase the number of basis functions, adding finer variation in the warp. Once a coefficient has been found at one level, it remains unaltered for all succeeding levels [Amit, 1994; Aykroyd and Mardia, 1996] or is used as a good starting estimate for the next level [Downie et al., 1996].

Any warp can be completely reproduced by an expansion in an infinite set of basis functions. The choice of basis is only significant when a truncated set is used, since this set determines the space of possible transformations. A wavelet basis will have an advantage over a sinusoidal function basis since wavelets have local support and therefore can produce stronger deformations in regions identified as requiring further refinement.

Note the terminology in [Aykroyd and Mardia, 1996] and [Downie et al., 1996]: "resolution" refers to the number of wavelet coefficients used. Mallat [1989a] provides the theory.
4.3. Increasing Warp Complexity

ther warping for a good match without affecting the remaining image. For the use of wavelets in medical image registration see Amit [1994], and also Aykroyd and Mardia [1996], and Downie et al. [1996] who use a statistical approach to find the wavelet coefficients.

4.3.2 Warps summarised by displacements of nodes or landmarks

The second subset of the warp complexity hierarchy methods includes landmark-based splines, where the complexity of the warp increases with the number of relatively displaced landmarks. We include also the finite element method, although strictly speaking this gives a discrete displacement field for a set of nodes, the finite elements being interpolants to give displacements at inter-nodal locations. However the hierarchical concept in the splines and finite element methods is the same: increasing the number of landmarks or nodes increases the complexity in which the warp can be described.

Finite elements

The finite element elastic matching method by Gee et al. [1994b] uses an increasing number of elements of decreasing size. Thus the initial levels of registration are fast since the elastic PDE is solved for fewer nodal points. By itself the method will not solve the local minima problem, since the distribution of nodes is not based on prior information of image content and only image information at nodes is used. However, the hierarchical method is coupled to the use of a Gaussian pyramid, hence the information at a nodal point contains information averaged over a region in the level of next-highest resolution. In this respect, the method is inferior to basis function truncations, where coefficients calculated in one level are retained, or at least used as good starting estimates for subsequent stages in the registration. Goshtasby [1986], Goshtasby [1987] give single-level schemes using a triangulation of the images, which could be inserted into a similar hierarchical framework. Finally, as an aside, Flusser [1992] uses an adaptive tesse-lation of the image domain, not during registration itself, but for computing a deformation to approximate a thin-plate spline.

Landmark matching

The automated landmark-based spline method can be further exploited within a hierarchical framework, with increasing warp complexity provided by increasing the number of landmarks in each level [Feldmar et al., 1997]. The starting locations of the new landmarks are in areas identified as being the most poorly registered. For general non-linear registration where the whole image is expected to deform, such a method using radial basis functions is inferior to a statistical basis-function expansion in that the entire deformation must be recalculated at each level. However, currently the only published non-linear registration algorithm to incorporate rigid bodies is that of Little et al. [1997] (section 6.3.2) which is a modified radial basis function warp with user-supplied landmarks. The hierarchical automated spline or radial basis function warp with the rigid-body modification is a suggestion for future research to which we return in section 8.1.1.
4.4 Increasing Model Complexity

The third hierarchical strategy is to use increasingly sophisticated matching methods as the registration progresses.

The strategy is to perceive the image as different models and apply registration methods based on each model type. For example, using solely physical models, the target image can first be modelled as a rigid body, and an affine transformation found to map it to the source. Proceeding to the next level, the image is then modelled as consisting of an elastic medium, and equation (3.66) solved to find optimal coefficients in a truncated basis-function expansion of a non-linear warp. Finally, the fluid algorithm can be applied, in which the image is modelled as a viscous fluid, with a deformation field giving individual mappings for each pixel.

Examples are: hierarchical sinusoidal basis function expansion followed by high dimensional elastic (stochastic gradient search), with initial images pre-registered linearly [Miller et al., 1993]; matching of (manually-identified) points, lines and surfaces using an interpolating spline, followed by refinement using hierarchical basis function expansions [Joshi et al., 1995]; hierarchical sinusoidal basis function expansion followed by high dimensional fluid applied in pyramid scale space, (initial images pre-registered linearly) [Christensen et al., 1994]; matching of edges using rigid, then affine transformations, followed by a deformation using cubic B-spline basis functions [Declerck et al., 1997].

Some algorithms carry with them an assumed participation or otherwise in such a hierarchical scheme. The highly local fluid algorithm [Christensen et al., 1995] is recommended only after initial application of more global registrations have accounted for the the main differences between the images, since the algorithm is programmed to aim for the nearest local minimum. Conversely, the thin-plate spline (with user-supplied landmarks) has no need for an initial linear registration since the global linear transformation is calculated concurrently with the warp.

4.4.1 Restricting repeated resampling

The number of levels in this hierarchy is significantly few. At the end of each step the transformed image undergoes pixel-level interpolation since generally pixels have been mapped to interstitial locations. Although highly accurate interpolation models exist for application to affine transformations (such as sinc or bicubic interpolation, see for example [Parker et al., 1983; Tian and Huhns, 1986; Hajnal et al., 1995b; Keys, 1981; Hou and Andrews, 1978; Unser et al., 1991]), such accuracy is not obtained for interpolation after a non-linear transformation and a very slight blurring of the image results. The blurring of pixel intensities is particularly significant in functional MR imaging (fMRI) studies [Bydder, 1995; Hajnal et al., 1995a], where the pixel intensities indicate functional activity in the brain (and where non-linear registration may be used for the averaging of a multi-subject study). Restricting the hierarchy to two or three levels minimises the accumulative affect of such blurring.
4.5 Summary

This chapter has focused on strategies to reduce misregistrations in algorithms employing directed search methods by initially reducing the number of local minima in the optimisation surface. Hierarchies of complexity have been classified into three main groups: those of temporally increasing data complexity, those of increasing warp complexity and those of increasing model complexity. A similar classification will be applied in chapter 6 to algorithms modelling spatially-varying complexity. First we devote a chapter to evaluating a small selection of the methods described in this and its preceding chapter.
Chapter 5

Evaluations of Nonlinear and Hierarchical Methods
5.1 Introduction

Having chosen a registration algorithm for a particular application, an evaluation stage is required to establish that for the expected range of image data the registrations will be of sufficient accuracy. To this end, registration metrics are selected and applied to a representative set of image pairs. The metrics are usually equal or equivalent to the similarity measures used by the algorithms since these measure the closeness of match between the features identified as corresponding in the image pair. However, if additional correspondences are known for a particular test image pair, these may be employed to verify the algorithm as a whole, including the choice of feature space and similarity measure. Evaluations may be of the overall or global registration quality, or of that local to particular regions. Examples of global metrics are the pixel-based similarity measures described in section 2.5.1: the sum of squared differences, the correlation coefficient and mutual information, the latter being applicable to the inter- as well as intra-modality case. Local registration metrics compute the similarity in small corresponding regions in the image pair. In particular, distances between known corresponding features (such as landmarks or known surface boundaries) are measured.

In this chapter we demonstrate the evaluation of three main non-linear algorithms described in chapter 3 using the three global metrics mentioned above, and apply both a global and a local metric to evaluate two of the scale spaces reducing data complexity as described in chapter 4. We emphasise that these are demonstrations of the evaluation methods - although they suggest relative qualities of registration, we did not test the algorithms on a range of data. Furthermore, we restricted our evaluations to the intra-modality case.

5.2 Global registration metrics: comparison of the 2-D TPS, SPM and fluid registration

In this section we compare the quality of registration provided by a linear registration (an implementation of the iterative algorithm given in Woods et al. [1992]), two versions of the thin-plate spline, and the SPM and fluid methods. We used as test images the inter-subject data set of head3A and head3B (figure 5.1) which are masked* versions of the head2A and head2B (figure 4.2), with the scalp removed.

The first variant of the TPS tested was an automated hybrid. It used the fluid model to estimate correspondences in the target for a large number of landmarks in the source which had been selected by a modified corner-finding algorithm. The landmark pairs were then inserted into a thin-plate spline (TPS) warp to summarise the fluid registration of source to target. The advantage of the TPS summary of the fluid deformation is that only the initial and final landmark locations need be stored if a record of the deformation field is required, in contrast to the full pixel displacement field of the fluid. The method is described in more detail in section 5.2.1 below. The full fluid registration was retained for comparison.

*The xdispunc tool written by Dave Plummer (Medical Physics, UCL); it operates on the UNC (University of North Carolina) image file format. xdispunc contains automatic contouring and masking options; we used these to outline the outer boundary of the brain and to set to zero the intensities of all pixels outside the boundary.
5.2. Global registration metrics: comparison of the 2-D TPS, SPM and fluid registration

Figure 5.1: Data set of (left) head3A and (right) head3B

For the second TPS variant, we selected manually 18 landmarks in locations estimated subjectively as being homologous in the source/target image pair and supplied these to a TPS spline; the choice of landmarks is described in section 5.2.2. The aim of comparing both TPS variants was to determine whether a more complex warp (based on a large number of approximately corresponding landmarks) gives a better registration globally than a warp based on only a few landmarks but whose locations are subjectively defined by an expert as being accurate. In this latter case of manual landmark identification, only a few landmarks are chosen since this is the case in a clinical context due to the time-consuming and laborious nature of manual landmark extraction.

The source image was in addition registered to the target by John Ashburner at the Functional Imaging Laboratory, using 2D versions of SPM with 4, 8 and 12 basis functions in the truncation. Quality of registration of all six methods was assessed visually and using three global metrics; the results are presented in section 5.2.3.

5.2.1 TPS with automatic landmark extraction

The automated thin-plate spline summary method consisted of three stages: extraction of landmarks in the source, tracking of the landmarks using fluid registration, and determining final landmark coordinates in the deformed source. The latter were assumed to approximate the corresponding locations in the target, within the accuracy of the fluid registration.

The first stage, the extraction of landmarks from the source, used a modified corner-finding algorithm. Corners are defined as regions of high isophote curvature [ter H Romeny et al., 1993]:

\[ C(x, y) = \frac{2L_x L_y L_{xy} - L_y^2 L_{yx} - L_x^2 L_{xx}}{(L_x^2 + L_y^2)^{\frac{3}{2}}} \]  (5.1)

where \( L_x, L_y, L_{xx}, L_{yy} \) and \( L_{xy} \) are the first and second partial derivatives of the image intensity at the pixel \( (x, y) \). We wished to limit the number of landmarks chosen, and required the landmarks selected to be scattered evenly over the image, rather than concentrated in clumps in regions of the image exhibiting high curvature. We therefore modified the corner-finding algorithm as follows. The image domain, excluding a margin of 5 pixels around the border\(^1\), was divided into a grid; for the given data of image size \( 256 \times 256 \) the inter-gridline spacing was set empirically at 17 pixels. Within each grid square, the pixel

\(^1\)The margin was omitted since in the fluid deformation the window boundary is fixed with constant zero displacements.
with the highest curvature value above a threshold was selected as a landmark. The algorithm fixed the threshold automatically at a curvature value of around 150\(^4\). It could however be raised so that no corners would be selected in roughly homogeneous regions (where corner values would be lower overall).

The fluid registration was then started on the source/target. On termination, the resulting displacement field was applied separately for each landmark to an image consisting of a bright dot (intensity 256) painted at the landmark location and black (zero) elsewhere. The deformed image displayed the final landmark location as determined by the fluid registration. However, generally the deformed landmark image consisted of a small region of a varying non-zero intensity, the diameter of the region being proportional to the scale of area growth in the neighbourhood of the landmark. The intensity varied from approximately 256 at its central pixels to near zero at its boundaries due to partial volume effects. To describe the final landmark location by unique coordinates, the centre of mass of the region was computed (figure 5.2).

The sets of initial and final landmark coordinates were then supplied to a TPS to summarise the fluid deformation.

Essentially this technique for obtaining landmark pairs uses the fluid registration as a search strategy for finding corresponding locations in the image pair, where two locations are assumed to correspond if the pixel intensities in their neighbourhood minimise the (sum of squared differences) similarity measure, but where additionally correspondences at neighbouring areas of the image pair influence the choice, together with smoothness constraints imposed by the viscous fluid model. True correspondence as defined by homologous anatomical locations is not guaranteed. In contrast, the manual landmark selection provides landmark pairs which are homologous insofar as the expertise and patience of the operator allows. Manual determination of landmarks in three-dimensional data sets is not an easy task and is slow.

\[ \text{Specifically, the threshold } t \text{ was set as } 32767 \frac{C_{\text{max}}}{C_m} 220 \]

where \(C_{\text{max}}\) = maximum curvature in the image, \(C_{\text{min}}\) = minimum curvature in the image, \(C_m = \max (|C_{\text{max}}|, |C_{\text{min}}|)\)

Figure 5.2: Propagation of initial landmarks by the fluid algorithm. The top row illustrates the source image deforming over time; the bottom row shows the concurrent deformation of an image of a bright dot at a landmark location.
5.2.2 Manual landmark extraction

18 homologous landmarks were manually selected with the aid of an atlas [Ellis et al., 1994; Crossman and Neary, 1995]; two experts (Dr Andreas Kleinschmidt and Dr Ingrid Johnsrude of the Functional Imaging Laboratory, Queen Square, London) advised on the nomenclature of the labels. These are illustrated in figure 5.3 and listed in table 5.1, with the coordinates following image processing convention, ie y axis increasing downwards, and the labelling of ‘left’ and ‘right’ follows radiological convention, ie ‘left’ refers to the subject’s left, which is the right of the image.

<table>
<thead>
<tr>
<th>no.</th>
<th>description</th>
<th>head3A x</th>
<th>head3A y</th>
<th>head3B x</th>
<th>head3B y</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>R lateral ventricle, A tip of frontal horn</td>
<td>115</td>
<td>82</td>
<td>112</td>
<td>84</td>
</tr>
<tr>
<td>2</td>
<td>L lateral ventricle, A tip of frontal horn</td>
<td>146</td>
<td>86</td>
<td>146</td>
<td>82</td>
</tr>
<tr>
<td>3</td>
<td>R lateral ventricle, P border of frontal horn</td>
<td>125</td>
<td>110</td>
<td>122</td>
<td>111</td>
</tr>
<tr>
<td>4</td>
<td>L lateral ventricle, P border of frontal horn</td>
<td>136</td>
<td>111</td>
<td>133</td>
<td>113</td>
</tr>
<tr>
<td>5</td>
<td>R, most posterior border between mediosagittal cortex and convexity</td>
<td>118</td>
<td>218</td>
<td>116</td>
<td>211</td>
</tr>
<tr>
<td>6</td>
<td>L, most posterior border between mediosagittal cortex and convexity</td>
<td>129</td>
<td>217</td>
<td>126</td>
<td>210</td>
</tr>
<tr>
<td>7</td>
<td>P sectional tip of RP horn of lateral ventricle</td>
<td>97</td>
<td>171</td>
<td>96</td>
<td>160</td>
</tr>
<tr>
<td>8</td>
<td>P sectional tip of LP horn of lateral ventricle</td>
<td>158</td>
<td>172</td>
<td>149</td>
<td>162</td>
</tr>
<tr>
<td>9</td>
<td>A tip of inter-thalamic adhesion</td>
<td>130</td>
<td>134</td>
<td>128</td>
<td>130</td>
</tr>
<tr>
<td>10</td>
<td>frontal pole</td>
<td>125</td>
<td>39</td>
<td>122</td>
<td>41</td>
</tr>
<tr>
<td>11</td>
<td>R lateral termination of Heschel’s gyrus posteriorly</td>
<td>67</td>
<td>126</td>
<td>69</td>
<td>112</td>
</tr>
<tr>
<td>12</td>
<td>L, middle of Heschel’s sulcus</td>
<td>182</td>
<td>135</td>
<td>178</td>
<td>128</td>
</tr>
<tr>
<td>13</td>
<td>putamen, RP tip</td>
<td>104</td>
<td>130</td>
<td>100</td>
<td>132</td>
</tr>
<tr>
<td>14</td>
<td>putamen, LP tip</td>
<td>155</td>
<td>138</td>
<td>154</td>
<td>132</td>
</tr>
<tr>
<td>15</td>
<td>putamen, RA tip</td>
<td>109</td>
<td>98</td>
<td>107</td>
<td>96</td>
</tr>
<tr>
<td>16</td>
<td>putamen, LA tip</td>
<td>152</td>
<td>99</td>
<td>151</td>
<td>96</td>
</tr>
<tr>
<td>17</td>
<td>putamen, L medial tip</td>
<td>114</td>
<td>110</td>
<td>111</td>
<td>107</td>
</tr>
<tr>
<td>18</td>
<td>putamen, L medial tip</td>
<td>148</td>
<td>110</td>
<td>147</td>
<td>105</td>
</tr>
</tbody>
</table>

Table 5.1: Description of location of the manually-selected landmarks for insertion onto a thin-plate spline applied to registering head3A to head3B. A: anterior, P:posterior, L: left, R: right.

5.2.3 Quantitative comparisons of the registration methods

Figure 5.4 shows the 170 landmarks selected in the original source by the modified comer-finding algorithm (centre; with the source reproduced for reference on the left) and the final landmarks determined by the fluid (right, displayed as a binary image for visual clarity).
5.2. Global registration metrics: comparison of the 2-D TPS, SPM and fluid registration

Figure 5.3: Diagram showing locations of manually selected landmarks in head3 images. Landmark number-labels are those given in the first column of table 5.1.

Figure 5.4: (left) Source image (head3A); (centre) Landmarks automatically selected in the source image; (right) new positions of landmarks after fluid registration of source to target; points have spread to small regions (the support of the regions is shown by conversion to a binary image for greater visibility). The corner-finding technique arbitrarily selected landmarks in grid squares outside the brain region as well as finding the point of highest curvature within each grid square inside the brain region.

Registration quality was assessed visually in three ways: by inspecting the deformed source and comparing it to the target, by subtracting the source from the target and noting the amount of remaining structure in the difference image, and by applying the computed deformation to a regular grid to compare the complexity and localisation of the deformations. Images of the deformed source compared to the target (figure 5.5) appear to show good registration to the target shape. Images of the differences between deformed source and target show more clearly where registration was incomplete (figure 5.6). Those of the fluid and its TPS summary (figure 5.6 far right, top and bottom) indicate that the latter gives a good summary of the fluid deformation. However, the grid images in figure 5.7 (left, centre) show that the more intricate detail in the deformations has been lost in the spline. As a comparison, the grid in figure 5.7 (right) shows greater deformations in the centre of the image where the feature landmarks were displaced, and also shows the distortions caused by the mappings extrapolated to outside the envelope of the landmarks. In background regions this effect is irrelevant; however, regions of the lower (posterior) cortex boundary were also displaced, as shown clearly in figure 5.6 (top centre right). From figure 5.6 (bottom), the accuracy of SPM matching appears similar to that of the fluid registration in this test, regardless of the
5.2. Global registration metrics: comparison of the 2-D TPS, SPM and fluid registration

We then quantified the accuracy of registration using three global pixel-comparison metrics: the correlation coefficient (equation 2.3); the sum of squares of differences (equation 2.2); and mutual information (equation 2.4). The sum of squares of differences gives lower values for good registration whereas correlation and mutual information give higher values for good registration. The results are presented in table 5.2 in order of increasing global registration quality (as determined by the metrics) and plotted in figure 5.8; the numbers in the left-hand column of the table correspond to the z-axis numbering in the plot.

These results indicate that increasing the number of basis functions in the SPM method does indeed improve the global registration quality as measured by each of the global registration metrics. Furthermore, although locally to each of the landmarks the manual TPS is assumed to be perfectly accurate, globally the registration is poor due to the sparseness of landmarks. The increased number of landmarks in the automated case, although the correspondences are only approximate, results in a greatly improved global registration quality.

Comparing the automated TPS with the fluid registration which generated the landmark correspondences, the global metrics show the deficiency in the TPS summary, with the full displacement field deformation of the fluid being superior. Finally, the fluid is shown to be more accurate than the SPM method - although both used the same similarity metric (sum of squared differences) applied to the full quantity of information in the whole pixel set. This is due to the ability of the fluid to achieve highly localised deformations.
5.2. Global registration metrics: comparison of the 2-D TPS, SPM and fluid registration

Figure 5.6: Differences between head3B and deformed head3A with registration (top left to right): none; linear; TPS (18 manually selected landmarks); TPS (170 landmarks derived from fluid); (bottom, left-right): SPM (4 basis functions); SPM (8 basis functions); SPM (12 basis functions); fluid registration.

Figure 5.7: Grid image (left) deformed during fluid registration of head3A to head3B and (centre) warped by the thin-plate spline using the 170 derived landmarks. The fluid deformations are localised at each image point whereas the spline smoothly interpolates the deformations between landmarks. (right) grid deformed by the manual TPS.
5.2. **Global registration metrics: comparison of the 2-D TPS, SPM and fluid registration**

Table 5.2: Global measurements of registration accuracy between the target and the image transformed by the registration method described in the left-hand column: sum of squared differences, correlation coefficient (cc) and mutual information (mi). For comparison, metrics for the comparison of the target to itself are given in the bottom row. Good registration gives high values of mutual information, and values close to 1 and 0 for the correlation coefficient and sum of squared differences respectively.

<table>
<thead>
<tr>
<th>registration method</th>
<th>sum sq diffs $\times 10^7$</th>
<th>cc</th>
<th>mi</th>
</tr>
</thead>
<tbody>
<tr>
<td>1  None (original source)</td>
<td>1.77</td>
<td>0.910</td>
<td>0.503</td>
</tr>
<tr>
<td>2  Linear</td>
<td>1.54</td>
<td>0.922</td>
<td>0.60</td>
</tr>
<tr>
<td>3  TPS (18 manual lmks)</td>
<td>0.822</td>
<td>0.954</td>
<td>0.795</td>
</tr>
<tr>
<td>3  TPS (170 fluid lmks)</td>
<td>0.420</td>
<td>0.977</td>
<td>0.899</td>
</tr>
<tr>
<td>4  SPM (4 basis fns)</td>
<td>0.420</td>
<td>0.977</td>
<td>0.863</td>
</tr>
<tr>
<td>5  SPM (8 basis fns)</td>
<td>0.377</td>
<td>0.980</td>
<td>0.922</td>
</tr>
<tr>
<td>6  SPM (12 basis fns)</td>
<td>0.350</td>
<td>0.981</td>
<td>0.949</td>
</tr>
<tr>
<td>7  Fluid registration</td>
<td>0.283</td>
<td>0.983</td>
<td>1.05</td>
</tr>
<tr>
<td>8  Gold standard: Target</td>
<td>0</td>
<td>1</td>
<td>6.06</td>
</tr>
</tbody>
</table>
5.2. Global registration metrics: comparison of the 2-D TPS, SPM and fluid registration

| Metric Value | Figure 5.8: Plot of metrics given in table 5.2. The correlation coefficient is plotted to the 8th power and the mutual information of target/target has been reduced to 2 for a clearer plot. For good registration, the sum of squared differences is reduced, whereas the correlation coefficient and mutual information are increased. |
5.3 Local and global registration metrics: comparison of scale-space methods

The registration metrics employed in the previous comparison gave an estimate of global registration quality; however it may be preferable to have a measure of registration in specific areas of the image. In this section we gauge the benefit of inserting fluid registration within a scale space strategy, by comparing the results of fluid registration in Gaussian and closure scale spaces, and of fluid registration where no scale space was used. The test image pair consisted of head1A (target) and head1B (source; figure 4.1). Head1A was a slice from a real MR data set and head1B had been created from it by deformation by a thin-plate spline with 20 landmarks (figure 5.9) of which 10 were near the boundary (scalp) and the remainder were internal. The true locations of a set of 20 corresponding landmarks were therefore known and were used as ground truths against which the positions of the source landmarks deformed by each fluid registration were compared. Registration at these landmarks was measured locally by comparing the deformed source locations to their known correspondences in the target.

Results of the fluid registration without scale space are given first in section 5.3.1, followed by registration in closure scale space in section 5.3.2 and Gaussian scale space in section 5.3.3. Section 5.3.4 compares the three sets of results.

Figure 5.9: Diagram showing locations of original (known) target landmarks in test image head1A: (left) ‘boundary’ and (right) ‘inner’. Original source landmarks are in the equivalent topological locations in head1B.

5.3.1 Fluid without scale space

Figures 5.10 and 5.11 show the progress of the fluid registration, without insertion within a scale space, at intervals of 15 timesteps together with the final deformation on termination; for each the displacement fields are also shown applied to a grid, showing more clearly the deformation. The termination criterion
was an increase in correlation coefficient between successive timesteps of less than 0.0001.

Figure 5.10: Progress of the fluid registration of head1B to head1A without scale space, showing iterations 15, 30, 45; left–right: Top row shows the deforming source, the lower row shows the same deformation applied to a regular grid.

Gross misregistration of cortex to scalp, as seen earlier in figure 4.3, is obvious. The final grid image (figure 5.11, lower right) highlights the distortion as a shift to the left the width of an entire grid square at the lowest edge of the misregistration, directly above an almost stationary gridline. The final landmark configuration (figure 5.12 centre right) shows the landmarks in the right anterior scalp region have hardly moved, while those in the left posterior region have moved towards their target locations.

\(^5\) radiological convention
Figure 5.11: Progress of the fluid registration of head1B to head1A without scale space, showing the 60th, 75th and final (83rd) iterations; left–right: Top row shows the deforming blurred source, the lower row shows the same deformation applied to a regular grid.

Figure 5.12: (left–right) Locations of original source landmarks; regions of deformed source landmarks at termination of the fluid registration without scale space; locations of centres of mass of these regions; locations of known target landmarks.
5.3.2 Fluid within closure scale space

The same data set was then registered using the fluid within closure scale space. A closure scale space of four levels was used. Mathematical morphological closure operations with disc-shaped structuring elements of radii 20, 4 and 3 pixel widths were applied to both source and target images (figure 5.13). The disc widths were chosen empirically, with the aim being to obtain a hierarchy of detail; the highest level aimed to retain only the outer image boundary and a rough division of the image into scalp and inside-scalp. The second level aimed to isolate three regions: scalp, skull (dark region between scalp and brain) and brain. The third level aimed to subdivide the brain region into the two hemispheres and to outline to some extent the major sulci. The lowest level of scale space contained the full-resolution image pair. Fluid registration was first applied to the source and target pair at the highest level of the scale space (disc width 20) until the termination criterion (as for the fluid without scale space) was satisfied. The resulting deformation field was then applied to a copy of the source in the next level of scale space and the fluid set in motion again. With the jump in scale level came a drop in correlation coefficient as image details were re-introduced which had not yet been matched. The process was repeated until termination in the highest resolution.

Figure 5.14 shows the displacements field at the end of each scale level, applied to the source in scale space, to the source at full resolution, and to a regular grid. Registration was almost complete at the outer boundary after the first level. The next level introduced slightly more detail in the deformation in the lower region (shown more clearly in the grid image). The third level showed little improvement on its predecessor; the final level added a little more detail in the posterior mediasagittal (centre) region. Figure 5.15 shows that generally a good match of deformed source to ground-truth landmarks has been achieved.

Figure 5.13: Closure with disc structuring elements of radii (left–right) 20, 4, 3 pixels applied to the 256 × 256 images head1A (top row) and head1B (bottom row).

*Sulci are involutions in the cortical surface of the brain.*
5.3. Local and global registration metrics: comparison of scale-space methods

Figure 5.14: Progress of the fluid registration of head1B to head1A at the end of each level in closure scale space; left–right: end of scale 3 (radius 20), 88th iteration; end of scale 2 (radius 4), 117th iteration; end of scale 1 (radius 3), 121st iteration; end of scale 0 (no closure), 136th iteration. Top row shows the deforming source in closure scale space, the lower rows show the same deformation applied to a regular grid and to the original-resolution source image.
Figure 5.15: (left–right) Locations of original source landmarks; regions of deformed source landmarks at termination of fluid in closure scale space; locations of centres of mass of these regions; locations of original target landmarks.
5.3.3 Fluid within Gaussian scale space

Finally we applied the fluid to the same data set within an isotropic Gaussian scale space.

Since we did not have a formula for determining the choice of levels in the scale space, we aimed to oversample, rather than undersample, the (continuous) scale space. Our reasoning was that since the registration terminated automatically within each level, oversampling would not be detrimental to the registration. If a particular level were superfluous, the registration would terminate within a few iterations and proceed to the next level. To check that this was indeed the case, we monitored the correlation coefficients at each iteration within each scale level and plotted the results in figure 5.22. Oversampling would be indicated by few iterations within the superfluous scale level.

The Gaussian scale space consisted of levels numbered $i = 1, \ldots, 10$, with the source/target pair being blurred at level $i$ by a Gaussian of spatial standard deviation $\sigma = 2i$. Figures 5.16 and 5.17 illustrate the scale space.

![Figure 5.16: Gaussian blurs of spatial standard deviation (left-right) 18, 16, 14, 12 applied to the 256 x 256 images head1A (top row) and head1B (bottom row).](image)

The scheme for terminating registration at each scale level and proceeding to the next was the same as that for closure scale space. Figures 5.18, 5.19 and 5.20 gives the progress of the fluid at the end of each level. The shape of the brain boundary has roughly reached that of the target at the end of the 6th level, whereas in comparison it was at the same stage of registration at the end of the second closure level. The final landmark configuration (figure 5.21) is similar to that obtained using closure scale space and appears to match the target well.
5.3. Local and global registration metrics: comparison of scale-space methods

5.3.4 Quantitative comparisons of the 3 fluid paradigms

Global metric

The correlation coefficient between target and deforming source were calculated at each timestep and those at the termination of each scale level (for the scale space cases) and at the 15-timestep intervals are listed in appendix B. In figure 5.22 the correlation coefficients are plotted for every timestep in the progress of each of three fluid registrations. For the fluid in Gaussian scale space, there are few iterations within scale levels 6, 5, 3 and 2, indicating that the scale space had been oversampled and that these levels were superfluous. The downward peaks in the scale space plots locate the inter-scale shifts where detail was re-introduced and had not yet been matched. The speed with which the new detail was registered is indicated by the width of these downward peaks. At each level of scale space (and without scale space) the plots show the correlation coefficients tending to a horizontal asymptote, ie registration was initially fast but that as the fluid progressed within that level the improvements were slower, indicating that a maximum in the similarity measure at that level was being reached. We note there is an increase in the magnitude in the drops in correlation coefficients (at the start of the next scale level) with increasing scale. This indicates that the quantity of detail re-introduced in the lower levels of scale space (ie at higher resolution) is higher.

Local metric

Figure 5.22 shows that closure gave slightly better registration than the Gaussian as measured by the global registration metric. Closure used more total iterations than Gaussian scale space, and also each level in closure scale space is slower to generate (for the higher levels where the structuring element is

Figure 5.17: Gaussian blurs of spatial standard deviation (left–right) 10, 8, 6, 4 applied to the 256 × 256 images head1A (top row) and head1B (bottom row).
5.3. Local and global registration metrics: comparison of scale-space methods

Figure 5.18: Progress of the fluid registration of head1B to head1A at the end of the first four levels in gaussian scale space; left–right: Top row shows the deforming blurred source, the lower rows show the same deformation applied to a regular grid and to the original-resolution source image.

larger) since Gaussian scale space is generated by convolution in the Fourier domain. Thus closure scale space is only preferable to the Gaussian if it gives better correspondences to the original landmark locations.

The spatial locations of the twenty known source and target landmarks (roughly indicated in figure 5.9) were plotted in figure 5.23, together with those tracked during each of the three fluid registrations using the technique described in section 5.2.1 (without the initial extraction step). The resulting figure gives an excellent graphical analysis of the local registration quality obtained by each of the fluid tests. Whereas in both the scale spaces registration appeared good (figures 5.14 and 5.20), figure 5.23 shows that at least one third of the landmarks (mainly in the lower region) had been insufficiently registered. For almost all landmarks, those deformed within either scale space are superimposed, showing that use of both scale spaces give equivalent registration quality in these locations.

We give a further example of local registration evaluation, using distances between known boundaries as well as known landmarks, in section 6.1. However we first resume our survey of registration algorithms by considering inhomogeneous methods, in the next chapter.
Figure 5.19: Progress of the fluid registration of head1B to head1A at the end of levels 5 to 2 in gaussian scale space; left–right: Top row shows the deforming blurred source, the lower rows show the same deformation applied to a regular grid and to the original-resolution source image.
5.3. Local and global registration metrics: comparison of scale-space methods

Figure 5.20: Progress of the fluid registration of head1B to head1A at the end of the final two levels in gaussian scale space; left–right: Top row shows the deforming blurred source, the lower rows show the same deformation applied to a regular grid and to the original-resolution source image.

Figure 5.21: (left–right) Locations of original source landmarks; regions of deformed source landmarks at termination of fluid in gaussian scale space; locations of centres of mass of these regions; locations of original target landmarks.
Correlation coefficient

Figure 5.22: Comparison of progress of the fluid registration through time of head1B to head1A in closure and Gaussian scale spaces and without scale space. Correlation coefficients are between the blurred target and deforming source images; the correlation coefficient drops at each change in scale level as more detail is re-introduced into the images, which needs matching. Within one scale level (or in the no-scale-space case) the correlation coefficient is monotonically increasing to a horizontal asymptote (approximates a negative exponential).
Figure 5.23: Plots of landmarks listed in tables B.2, B.4 and B.5: starting locations in the source image head1B, true target locations in head1A, and locations in deformed source produced by fluid registration without scale space and in Gaussian and closure scale spaces.
Chapter 6

Inhomogeneous Registration
6.1 Introduction

Single-level registration algorithms have been presented in chapters 2 and 3 in two groups: those which apply linear transformations and those which allow for higher order deformations. Generally higher order deformations are performed after an initial rough registration by a linear method, so a combination of linear and non-linear method are used sequentially. Chapter 4 examined the application of hierarchies of data, warp and model, where complexity increases temporally with the progress of registration. It is rare that an algorithm allows simultaneous or parallel application of both linear and non-linear models within one image, so that only selected areas of the image are allowed to deform. Many medical images contain regions representing both soft and hard tissue, and whereas the former require high order deformations to achieve a good fit to a target image, in an intra-subject study the hard tissue regions should remain rigid. Registration of such image pairs requires algorithms where the model varies spatially within the image domain, using, for example, prior information on the variation of tissue types within the deforming image. These are instances of inhomogeneous non-linear registration algorithms.

There is active research into the modelling of real tissue deformation and growth. The main focus is for application to real-time perioperative registrations and simulations of tissue growth; the models have yet to be applied extensively to the registration of two given static images. Published models include:

- Gibson et al. [1998] who describe attempts at real tissue deformation for surgical simulation, modelling interaction with surgical tools, and tissue cutting, tearing and suturing
- Troccaz et al. [1998] who describe models of “real” tissue deformation incorporated into interactive systems
- Schill et al. [1997] who model the falx cerebri (a membrane dividing the two hemispheres of the brain) as perpendicularly oriented elastic fibres
- Bro-Nielsen et al. [1997]; Andresen et al. [1998] who model bone deposition and resorption during growth of the mandible
- Griffin [1994] who study growth and development of the cortex.

An inhomogeneous model may be inserted into a segmentation scheme similar to that in [Kapur et al., 1996], to give a feedback-loop of registration and segmentation. Prior estimates of tissue stiffness can be obtained from an initial rough classifier, or from new MR imaging techniques [Manduca et al., 1996]. Inhomogeneous registration to a labelled atlas would then provide a refinement of the segmentation.

The second major application of inhomogeneous models is in cross-population studies, where the variability in shape varies for each structure in the anatomy. For example, the convolutions of the cortex vary more extensively across a normal population than do the shapes of deeper structures within the brain. Cross-population variability of individual structures substitutes for physical “stiffness” in the inhomogeneous registration model.
Finally, the deformation of a particular region may be temporarily restricted in order to allow other regions to register first.

This chapter classifies types of inhomogeneity in registration algorithm and reviews those available in the literature. We then present three modifications to the fluid registration algorithm which introduce inhomogeneities into its application. Section 6.4 describes inhomogeneities in applying the force field and in computing the velocity field, and section 6.5 presents the varying-viscosity fluid registration algorithm. Finally, sections 7.3 and 7.4.2 show results of application of these algorithms to 2- and 3-dimensional data respectively.

6.2 A classification of spatial inhomogeneities in registration algorithms

Chapter 4 classified temporal hierarchies of registration into those of data, those of descriptors of deformation or warp, and those of complexity of model. In a similar vein, we classify spatial inhomogeneities as those in the relevance of data, in the strength of deformation constraints, and in the application of model type.

6.2.1 Variable data influence

The first type of inhomogeneity varies the importance attached to information content in the domain of the image pair when computing the transformation required at each location in the source. In terms of a Bayesian approach, where the deformation is determined by a solution of a weighted sum of likelihoods and priors, (equation 3.10) the weight assigned to the likelihood is varied according to prior assumptions about the relevance of the data in different regions of the image. In terms of regularisation, where the equation solved is a weighted sum of driving forces and constraints on the deformation (equation 3.5), the influence of the driving force is weakened or strengthened relative to the deformation constraints.

Section 6.4.1 describes such an implementation, where after computing uniformly over the image pair the driving force at each pixel, a filter is passed over the force field, reducing the applied force to zero in selected areas of the image. We term such regions passive, where any resulting deformation is due solely to their proximity to actively-deforming regions. Thus we define:

**Definition 6.1 (Passive region)** Let \( \Omega_1 \subset \Omega \) be a region whose deformation is given by \( \vec{u}(\vec{x}) \) minimising the constrained optimisation

\[
J(\vec{u}) = \lambda \ R(\vec{u}(\vec{x})) + D(\vec{S}(\vec{u}(\vec{x})), T(\vec{x}))
\]

(6.1)

where \( D \) is the log likelihood and \( R(\vec{u}(\vec{x})) \) is the log prior constraint dependent on the deformation. We define \( \Omega_1 \) to be passive if for all pixels \( \vec{x} \in \Omega_1 \), the regularisation parameter \( \lambda \) weighting the log prior is infinitely large. Equivalently, within a passive region we solve for \( \vec{u}(\vec{x}) \) by minimising

\[
J(\vec{u}) = R(\vec{u}(\vec{x}))
\]

(6.2)
6.2. A classification of spatial inhomogeneities in registration algorithms

**Definition 6.2 (Actively-deforming region)** A region $\Omega_i$ is actively-deforming if for all pixels $\vec{z} \in \Omega_i$, the regularisation parameter $\lambda$ weighting the log prior is finite.

A typical medical application would be an inter-modality image pair of which the source contains additional structures (such as tumours) whose homologues are absent in the target but which may be confused, due to similarities of pixel intensity value, to other regions nearby. Registration protocol would consist of roughly locating such structures in the source using some segmentation method, labelling them as passive, and then applying an inhomogeneous algorithm such as the fluid with an inhomogeneous filter to the forces field to register the source to the target.

**6.2.2 Variable deformability**

The second inhomogeneity paradigm varies the strength of deformation constraint. Regions of the image are then classified as strongly- or weakly deformable. In the case of a registration modelled on the behaviour of a physical material, the deformability is described by one or more parameters of the material properties - the elasticity of an elastic medium (section 6.3.3) or the viscosity of a fluid (section 6.5). Allowing this parameter to vary spatially requires the derivations of modified PDE's to account for the parameter gradients.

**Definition 6.3 (Strongly and weakly deformable)** Let $\Omega_i \subset \Omega$ be a region whose deformation is given by $\tilde{u}(\vec{z})$ satisfying equation (6.1) where the prior smoothness constraint $R(\tilde{u}(\vec{z}), \mu(\vec{z}))$ is dependent on the deformation and on an independent parameter $\mu \in (0, 1]$ varying spatially within the image, such that $R(\mu) \rightarrow 0$ as $\mu \rightarrow 0$. We then term $\Omega_i$ to be strongly- or weakly-deformable according to the range of $\mu$ from 0 (strongest) to 1 (weakest).

The strength of the deformation parameter is supplied at every position in the source image, using prior information obtained from one of two sources: physical or statistical. In the first instance, prior information is available on the deformability of the physical tissues which the images represent. For example, Manduca et al. [1996] demonstrate that it may be possible to estimate tissue elasticity using certain scanning protocols, and there are straightforward assumptions such as that hard tissue remains rigid. Basing the variability of the deformation constraint on such physical information is valid only in intra-subject studies, where the registration of the source to the target attempts to reproduce actual physical movements within the tissues which the images represent. A second type of prior information is applicable to and derived from cross-population studies, where the variation in the deformation constraint is a function of the statistical cross-population variability in the shapes of the various structures within the images. Structures which have been found to display little variance in size and shape across a population of normals will be labelled weakly deformable, while other areas exhibiting greater variability will be strongly deformable and allowed greater deformations in registering to their homologues in the target or atlas image. For instance, Davatzikos [1997] allows high variability in ventricular and cortical fold regions, and low variability in subcortical structures, in his variable-elasticity algorithm (section 6.3.3).
Inhomogeneities in deformability can be extended to include *anisotropies* in the constraint parameters, such that there are preferential *directions* of deformation. Anisotropies in ease of deformation are common in physical tissue such as muscle. However, adapting the mathematical representation of the deformation of a physical medium (such as the viscous fluid) to allow anisotropies is more complex than only allowing for isotropic inhomogeneities and we leave such a possibility as a topic for future research.

### 6.2.3 Variable model type

Finally it is possible to vary the models or equations causing deformation within the image pair. This type of inhomogeneity is highly successful at achieving completely affine transformations within selected regions while deforming intervening or surrounding areas. Boundary conditions are set between model type regions such that a continuity of mapping is ensured across the image. Examples of such algorithms are the rigid body / mutiquadric spline combination (section 6.3.2), the three-component finite element model (section 6.3.1) and the fluid with additional boundary conditions within the velocity field (section 6.4.2). In the last two cases, the updating of nodal displacements (in the finite-element model) or of the pixel velocities (in the variant fluid) is prohibited within selected areas. This is an easy and effective method of ensuring these regions remain rigid, but furthermore they remain *motionless*.

We define here the concept of a rigid body within the deforming source, together with two paradigms of a rigid region.

**Definition 6.4 (Rigid region)** A region \( \Omega_i \subset \Omega \) is said to be rigid within a non-linear deformation of source \( S(\Omega) \) to target \( T(\Omega) \) if the transformation \( u(\Omega_i) \) is a linear transformation.

**Definition 6.5 (Motionless region)** A region \( \Omega_i \subset \Omega \) is said to be motionless if \( \forall \vec{z} \in \Omega_i \), the transformation \( u(\vec{z}) = 0 \). Where the registration \( S(u(\Omega_i, t)) \rightarrow T(R) \) is a function of time, \( \Omega_i \) is motionless if \( \frac{\partial}{\partial t} u(\Omega_i, t) = 0 \ \forall \vec{z} \in \Omega_i, \forall t \) and if \( u(\Omega_i, 0) = 0 \).

It may be desirable to have rigid but *independently-moving* regions:

**Definition 6.6 (Independently-moving region)** A region \( \Omega_i \subset \Omega \) is said to be independently-moving if \( \forall \vec{z} \in \Omega_i \), the transformation \( u(\vec{z}, t) = c(\vec{z}) \) is a non-zero linear function of \( \vec{z} \) where \( c(\vec{z}) = c(\vec{z}_j) \) \( \forall \vec{z}_j, \vec{z}_j \in \Omega_i \) at any time \( t \) and where \( \exists \vec{z} \in \Omega_i, \vec{z} \notin \Omega_i \) such that \( u(\vec{z}, t) = c(\vec{z}, t) \) satisfies the regularisation of a likelihood and prior.

One option, which we have not yet tested, is to supply uniform and constant but non-zero velocities within each rigid region. Alternatively, the rigid body / mutiquadric spline combination allows for any number of rigid bodies with an independent linear transformation for each. The main application of such algorithms will be in the modelling of tissue movements during surgery. Ideally they will attempt to faithfully reproduce movements of the physical tissues which the images represent, and will also be fast for real-time implementation.
6.3 Review of current inhomogeneous registration algorithms

6.3.1 Three-component finite element model

Edwards et al. [1997] is an example of an inhomogeneous finite-element model based on three tissue types, labelled rigid, deformable and 'fluid'. The deformations are driven by user-supplied landmark displacements and deformations of the deformable regions are constrained by three energy terms:

$$E_{\text{tension}}(N_i, N_j) = |N_j - N_i - N_{ij}^0|^2$$
$$E_{\text{stiffness}}(N_i, N_j, N_k) = |N_j + N_k - 2N_i|^2$$
$$E_{\text{fold}}(N_k, N_l, N_m) = \begin{cases} \frac{A^2}{2A_0^2} + \frac{\gamma^2 A^2}{A_0^2} & \text{if } A < \gamma \\ 2 & \text{otherwise} \end{cases}$$

where $N_{ij}^0$ is the original distance between the nodes $N_i$ and $N_j$ and the nodes $N_i, N_j$ and $N_k$ are collinear before deformation; nodes $N_k, N_l$ and $N_m$ form a triangle with initial area $A_0$ and deformed area $A$. The 'fluid' deformations are constrained only by $E_{\text{fold}}$ which prevents folding of the image. 'Rigid' regions are obtained by prohibiting the updating of their nodal displacements.

6.3.2 Incorporation of rigid structures in the multiquadric (MTQ) spline

A model permitting two tissue types is the modified multiquadric spline [Little et al., 1996, 1997], a variant of the thin-plate spline method which incorporates any number of user-segmented rigid bodies which undergo independent linear transformations only, while the remaining image regions are deformed using landmark-based spline warps with multiquadric basis functions. The method is applied to pre-segmented images, with regions classified as hard or soft tissue. Figure 6.1 illustrates the algorithm on a knees images with bones as rigid bodies. The hard tissue regions form a set $\mathcal{R}$ of $n$ rigid bodies $\{\Omega_i\}$, such that $\mathcal{R} = \bigcup_{i=1}^{n} \Omega_i$, where one 'body' can in fact consist of separate parts (which all undergo the same affine transformation) but no two 'bodies' can overlap*. The method requires use of a distance transform which is then used to weight differently the linear and non-linear components of the overall image mapping, such that the non-linear terms are smoothly reduced to zero as the rigid bodies are approached, and each rigid body is constrained to its own linear mapping while contributing to the underlying linear drift of the non-rigid areas.

The overall original spline mapping in 2D of any point $\tilde{x}_i = (x_i, y_i)$ to $\tilde{u}_i = (u_i, v_i) = (u_1, u_2)$ has been given in equation (3.15) as the set of maps

$$f_1(x_i, y_i) = a_{11} + a_{12}x_i + a_{13}y_i + \sum_{j=1}^{m} w_{1j}U(|P_j - (x_i, y_i)|)$$
$$f_2(x_i, y_i) = a_{21} + a_{22}x_i + a_{23}y_i + \sum_{j=1}^{m} w_{2j}U(|P_j - (x_i, y_i)|)$$

*Referring to definition 1.12, this would be a partition if $\mathcal{R} = \Omega$
6.3. Review of current inhomogeneous registration algorithms

Figure 6.1: Multiquadric deformation of knee image incorporating two rigid bodies (bones)

summarised as

$$\tilde{f}(\vec{z}) = A \begin{pmatrix} 1 \\ x_i \\ y_i \end{pmatrix} + \sum_{j=1}^{m} \tilde{w}_j U \left( |P_j - (x_i, y_i)| \right)$$  \hspace{1cm} (6.3)

where $A$ is a $2 \times 3$ matrix of coefficients of the linear transformation, and $\tilde{w}_j$ are rows from the $2 \times n$ matrix of coefficients of the non-linear transformation, with $n$ the number of fiducials or landmarks $\{P_j\}$. The affine coefficients $A$ were determined at the same time as the non-linear coefficients $\{\tilde{w}_j\}$, from the known mappings of the landmarks.

In the combination rigid/multiquadric spline the affine coefficient matrix $A$ is instead computed first, independently of the landmark mappings, from a linear combination of the predefined affine transformation matrices $A_i$ of the rigid bodies $\Omega_i$:

$$A = \sum_{i=1}^{n} \omega_i A_i$$  \hspace{1cm} (6.4)

The $\omega_i$ are Shepard weight functions which were introduced by Shepard [1968] for inverse distance weighted interpolation and are

$$\omega_i(\vec{z}) = \frac{q_i(\vec{z})}{\sum_{j=1}^{n} q_j(\vec{z})}$$

where $q_i(\vec{z}) = \frac{1}{D_i(\vec{z})^{-\alpha}}$ is a non-negative-valued distance transform giving a distance measure of point $\vec{z}$ from body $\Omega_i$. The choice of such distance transforms will be discussed later on in section A.1. The parameter $\alpha > 1$ is to ensure the continuity of first- (or higher-) order derivatives. The weights satisfy the following properties for all $\vec{z} \in \mathbb{R}^d$:

1. $\omega_i(\vec{z}) \geq 0$ since $D_i(\vec{z}) > 0$

2. Partition of Unity:

$$\sum_{i=1}^{n} \omega_i(\vec{z}) = \sum_{i=1}^{n} \frac{q_i(\vec{z})}{\sum_{j=1}^{n} q_j(\vec{z})} = \frac{\sum_{j=1}^{n} q_j(\vec{z})}{\sum_{j=1}^{n} q_j(\vec{z})} = 1$$

3. In the limit as $\langle \vec{z} \rangle$ approaches any pixel $\langle \vec{z}_i \rangle$ on the surface $\delta \Omega_i$ of rigid body $\Omega_i$, $D_i(\vec{z}) \to 0$, so $q_i(\vec{z}) \to \infty$, $\sum_{j=1}^{n} q_j(\vec{z}) \to q_i(\vec{z})$ and so $\omega_i(\vec{z}) \to 1$ as $\vec{z} \to \vec{z}_i \in \delta \Omega_i$
4. \( \omega_i(\bar{x}) = 0 \) if \( \bar{x} \in \Omega_j, j \neq i \)

which follows from the first three properties.

Properties 3 and 4 ensure that the total (affine) transformation \( A \) at object \( \Omega_i \) will equal its predetermined affine transformation, unaffected by contributions from the other objects, since for all points \( \bar{x}_i \in \Omega_i \):

\[
A_i(\bar{x}_i) = \sum_{j=1}^{n} \omega_j(\bar{x}_i)A_j(\bar{x}_i)
\]

\[
= \sum_{j=1}^{n} \omega_j(\bar{x}_i)A_j(\bar{x}_i) + \omega_i(\bar{x}_i)A_i(\bar{x}_i)
\]

\[
= \sum_{j=1}^{n} 0A_j(\bar{x}_i) + 1A_i(\bar{x}_i)
\]

\[
= A_i(\bar{x}_i)
\]

As points \( \bar{x} \) approach a given rigid body \( \Omega_i \), the relative contribution from the linear transformation \( A_i \) of \( \Omega_i \) increases; the underlying affine transformation in the image will vary smoothly between the rigid bodies (the smoothness controlled by \( \mu \)).

For the non-linear part of the mapping, the radial basis functions \( K(\bar{x}, \bar{x}_i) \) are modified by direct multiplication of the distance functions:

\[
\bar{K}(\bar{x}, \bar{x}_i) = D_0(\bar{x})D_0(\bar{x}_i)K(\bar{x}, \bar{x}_i)
\]

Little et al. [1996, 1997] use a variant of Hardy multiquadric basis functions. Recall from equation (3.43) that the circular hyperboloids are of the form

\[
K(\bar{x}, \bar{x}_i) = w (|\bar{x} - \bar{x}_i|^2 + C)^{-\frac{1}{2}}
\]

Increasing \( w \) increases the steepness of the asymptotic cone, and decreasing \( C \) pulls the surface apex into a point (the apex of the asymptotic cone). Little takes as \( C \) the squared distance from landmark \( \bar{x}_i \) to the closest rigid body, ie \( C = D_0(\bar{x}_i)^2 \). Hence the modified basis functions are

\[
\bar{K}(\bar{x}, \bar{x}_i) = D_0(\bar{x})D_0(\bar{x}_i) (|\bar{x} - \bar{x}_i|^2 + D_0(\bar{x}_i)^2)^{-\frac{1}{2}}
\]

Figure 6.2 shows the effects of the distance functions on the non-linear transformation. Displacements of landmarks further from the rigid bodies produce more 'elastic' contributions to the warps, in that their effect is more localised, whereas the effects of landmarks closer to the rigid bodies are smoother. The relative contributions of warp displacements to the transformation of a point increases with distance of a point from the set of rigid bodies. As a point approaches any rigid body its transformation becomes more affine; points in rigid bodies experience purely affine transformations, the warp contributions being zero.

The warp coefficients are found from the relative landmark displacements \( \tilde{u}_i \) by first subtracting the overall affine transformation \( T(\bar{x}_i) \) of each original landmark \( \bar{x}_i \) from its final position \( \bar{u}_i \). The affine trans-
6.3. Review of current inhomogeneous registration algorithms

Figure 6.2: Effect of the distance functions on the combined warp

Formulations are

\[ T(\tilde{x}_i) = \tilde{g}(\tilde{x}_i)^T A(\tilde{x}_i) \]  

where \( \tilde{g}(\cdot) \) is the polynomial basis for the affine transformation, usually \( \tilde{g}(\tilde{x}) = (1, x_1, x_2) \). This leaves the displacement of landmark \( i \) due to the warp

\[ \tilde{u}_i = \tilde{u}_i - T(\tilde{x}_i) \]

The warp coefficients \( W \) are obtained by inverting the matrix of modified warp functions applied to landmark pairs in the original image, \( \tilde{K} \), and applying it to the matrix \( \bar{u} - T : \)

\[ \bar{w} = \tilde{K}^{-1}(\bar{u} - T) \]

Having computed all the affine and warp coefficients, the total transformation \( \tilde{f} \) can then be applied to all points in the original image:

\[ \tilde{f}(\tilde{x}) = A(\tilde{x})\tilde{g}(\tilde{x}) + \sum_{i=1}^{n} W_i^T \tilde{K}(\tilde{x}, \tilde{x}_i) \]

where \( W_i^T \) is the \( i \)th row of the warp coefficients matrix \( W \).

\[ T = \begin{pmatrix} \tilde{g}(\tilde{x}_1)^T & A(\tilde{x}_1) \\ \tilde{g}(\tilde{x}_2)^T & A(\tilde{x}_2) \\ \vdots & \vdots \\ \tilde{g}(\tilde{x}_n)^T & A(\tilde{x}_n) \end{pmatrix} \]
6.3.3 Elastic registration with variable elasticity

Davatzikos [1997] presents an elastic registration model applied to images of the head where the elasticity parameter varies spatially within the image. The deformation is driven by distances between parametrically-defined pre-segmented cortical and ventricular surfaces in the source and target, and also incorporates a pre-strained elasticity term. The latter allows for voluntary growth in specified image areas, for example to model the growth of a tumour or to accelerate the expansion of ventricular regions to give them a better pre-registration to those in the target prior to the full registration. We give here an outline of his method.

First the brain tissue is segmented from the images, defining a binary mass function \( m(x) \in [0, 1] \) and a deformable surface (a two-dimensional snake) is applied to the mass functions of the source and target, giving for each a parametric description of the shape of the outer cortical surface, as follows: the surfaces are summarised in terms of two parameters \( u, v \) which are coordinates intrinsic to the surface; setting first \( u \) and then \( v \) as constant and varying the other, the surface is covered by a family of perpendicularly-intersecting curves, giving a grid closely fitting the underlying cortical shape. The two parametrised surfaces are then brought into register using an elastic deformation with forces derived from measures of curvature at the grid points (figure 6.3). A hierarchical approach is used where the deformable surface is composed of first a coarse and then a fine polygonal mesh. The latter gives a better fit to more intricate details in the cortical surface and is used to improve on the inter-surface match computed at the coarser scale. In the next stage of the algorithm, the ventricular boundaries are identified. At each point on the ventricular surface in the deforming source, a force is computed, consisting of the distance to the nearest point on the boundary of the target ventricular surface, weighted by the scalar product of outward normals to the ventricular surfaces at these points. These ventricular forces together with cortical forces obtained from the cortical surface matching provide a total external driving force field \( \hat{F} \) which is supplied to the variable-elasticity equation:

\[
\{ \hat{F} + \lambda \nabla^2 \hat{U} + (\lambda + \mu) \nabla (\nabla \cdot \hat{U}) \} + \\
\{ (\nabla \hat{U} + \nabla \hat{U}^T - 2 \hat{I}) \nabla \lambda + (\nabla \cdot \hat{U} - 3) \nabla \mu \} + \\
\{ \epsilon (2 \nabla \lambda + 3 \nabla \mu) + (2 \lambda + 3 \mu) \nabla \epsilon \} = 0 \quad (6.6)
\]
This consists of three terms. The first is the regularisation between driving force $\mathbf{F}$ and the constant-elasticity constraints on the vector of displacements $\mathbf{U}(z)$. The second term contains gradients in the elasticity parameters $\lambda$ and $\mu$, resulting allowing variation in the elasticity field. In his application [Davatzikos, 1997], the ventricular and cortical surface regions are set lower elasticity values, allowing for greater ease of deformation. The third term contains gradients in an additional strain tensor $\epsilon$ which forces extra expansion or contraction in pre-selected regions. For example, $\epsilon$ may be given a non-zero value in the ventricular areas if it has been estimated that these are abnormally sized to an extent that their surfaces are insufficiently close to those of the target to obtain a successful registration. The additional strain tensor in effect supplies an independent driving force to selected regions in the image, and hence the algorithm also contains inhomogeneities in activity, or data influence.

6.4 Modifications to the constant viscosity fluid algorithm

This section describes two methods of introducing inhomogeneities into the application of the regular fluid algorithm such that deformation is reduced or prohibited in areas specified as either passive or motionless. For ease of reference we will refer to these two modifications to the fluid algorithm as MF1 and MF2. The first applies a spatial filter to the force field, reducing forces to zero in the neighbourhood of selected areas; the second supplies additional boundary conditions to the computation of velocity fields such that velocities at the selected areas are always zero. We describe each method in the following subsections, and results of the application of these methods are given in subsections 7.3.1 and 7.3.4.

6.4.1 MF1: Inhomogeneous filter applied to the force field

The first modification, MF1, makes use of prior knowledge of regions whose intensity information we do not wish to contribute to driving the fluid registration. These regions will be passive in the registration process in that their deformation is only due to proximity to actively deforming regions. At each timestep in the fluid algorithm, forces at each pixel in the image are computed as before (using equation 3.68) for insertion into the fluid PDF, to drive the registration. A filter is then passed over the forces field driving the fluid deformation, so that deformations become zero as they approach a passive region. However this also has the affect of increasing the force field with increasing distance from the region, so that registration occurs much faster in regions far from the passive region. The net result is that altogether registration is slower since in each iteration the displacement at each pixel is scaled such that the maximum pixel displacement in the image is below a given threshold, typically 0.5 pixel units, to ensure the deformation fields change smoothly over time. Hence pixels furthest from passive regions will move by this maximum displacement while areas closer to the passive regions will hardly deform until later iterations, instead of all non-active areas registering
at the same time.

We therefore constructed a function (6.7) to weight the EDT before application to the force field.

\[
f(x) = \begin{cases} 
    0 & : x \leq a \\
    \frac{1}{2} \left(1 - \cos \left(\frac{\pi(x - a)}{b - a}\right)\right) & : a < x < b \\
    1 & : x \geq b 
\end{cases}
\]

(6.7)

This weighting function is uniformly zero inside and within a small margin \(a\) around the passive regions, is uniformly constant with value 1 beyond a given distance \(b\) from the passive regions, and smoothly increases from zero to 1 in the intermediate areas. In all following results we used \(a = 1\) and \(b = 13\).

Figure 6.4 illustrates the weighting function for \(a = 2\) and \(b = 10\). We found the 4EDT algorithm [Sonka et al., 1996] to be adequate to our needs as an approximation to the Euclidean distance. In our application, the passive regions were defined by their location, supplied as a binary image flagging each pixel as passive or active. Figure 6.5 shows the EDT applied to the three windows in the house image (figure 7.1).

Figure 6.5: Application of 4SED to windows in house image.

6.4.2 MF2: Extra boundary conditions in the solution of the velocity fields

The second modification, MF2, allows for specified regions to remain rigid by prohibiting their pixel movements. Such regions are identified within a supplied binary image or array whose pixels are labelled either motionless or mobile. This array is passed to the procedure (such as SOR) which solves the fluid
6.5. MF3: The inhomogeneous fluid model - fluid with spatially varying viscosity

PDE in each timestep. Only velocities at mobile pixel locations are updated; those labelled motionless remain at zero velocity. In the regular fluid algorithm, pixels lying on the boundary of the image are motionless.

A suggested extension to this concept is to specify pre-determined velocities in regions defined as independently moving. This may allow for rigid-body transformations in such areas within an overall fluid deformation. We have not proceeded with this option but leave it as a suggestion for future work.

6.5 MF3: The inhomogeneous fluid model - fluid with spatially varying viscosity

6.5.1 Introduction

We have developed a variant of the viscous fluid where the viscosity varies spatially over the image. We refer to it as MF3. It incorporates into the regular fluid model a similar idea to the inhomogeneous multi-quadratic spline (section 6.3.2) and the inhomogeneous elastic model (section 6.3.3), to permit very localised deformations, allowing for deformations of different tissue structures according to estimates of their rigidity or cross-population variability using prior knowledge suggested by a rough initial segmentation. In contrast to Little et al. [1997] and Edwards et al. [1997], the inhomogeneous fluid algorithm presented here allows for a continuum of strength of constraint of deformation. It depends on a modified viscous fluid PDE which was reformulated for us by Dr Kalvis Jansons of the Department of Mathematics, UCL. The theory may be related to the creation of inhomogeneous scale spaces using varying conductance in the heat diffusion equation [Whitaker, 1993].

6.5.2 Theory

In section 3.8.2 we saw that the fluid PDE is summarised by

\[ \nabla \cdot \tilde{\sigma} + \tilde{f} = 0 \]  (6.8)

where \( \tilde{f} \) is the driving force and \( \tilde{\sigma} \) is the stress tensor, given by

\[ \sigma_{i,j} = -p \delta_{i,j} + \mu \left( \frac{\partial v_i}{\partial x_j} + \frac{\partial v_j}{\partial x_i} \right) \]  (6.9)

and where \( \vec{v} \) is the velocity field, \( p \) is a pressure term and \( \mu \) is the viscosity constant. Since there is no explicit time-dependency in the PDE, the fluid has no inertia (conservation of momentum has not been enforced), and hence the fluid will be very viscous.

We now let the viscosity \( \mu \) vary spatially within the image and expand equation (6.8) to give the PDE for the variable viscosity fluid model:

\[ -\nabla p + (\nabla \mu \cdot \nabla \vec{v}) \vec{v} + \frac{\partial \mu}{\partial x_j} \nabla v_j + \mu \nabla^2 \vec{v} + \mu \nabla (\nabla \cdot \vec{v}) + \tilde{f} = 0 \]  (6.10)
In our application, the pressure gradients $\nabla p$ are neglected. The varying viscosity PDE bears some resem­bance to the varying elasticity PDE (equation 6.6), without the additional strain tensor terms and without the second elasticity parameter $\lambda$.

Since the partial differential operator now varies over the image, a fast solution by a basis function expansion or by convolution with filters derived from its eigenfunctions [Bro-Nielsen and Gramkow, 1996] is not longer possible; we use the SOR iterative method [Strikwerda, 1989], with the finite difference scheme detailed in appendix A.5.

As an aside, Gee et al. [1993] reported no significant difference if the elasticity parameter was altered globally in his elastic registration (see page 106). Likewise we found the fluid did not produce any noticeable difference in behaviour if the viscosity parameter $\mu$ was altered globally within the image. Hence, although $\mu$ may be varied spatially in MF3 to create an inhomogeneous algorithm, we suggest that the temporal parallel, being a hierarchical method based on the viscosity gradually decreasing with time, will be no different to the single level case. We do not, however, have experimental results to support this claim.

In the next chapter we will apply the regular and the three modified fluid algorithms in six tests to evaluate their success in locally inhibiting deformation. We will also demonstrate their application to assisting the detection of abnormalities and to modelling tissue types of varying rigidity. In addition, we compare the inhomogeneous fluid algorithms (and in particular MF3) to inhomogeneous registration using the combination MTQ described in section 6.3.2. Finally, in the last test, the regular fluid and MF3 are applied to full 3D data volumes.
Chapter 7

Evaluations of Inhomogeneous Registration Algorithms

7.1 Introduction

This section evaluates inhomogeneous registration algorithms in respect of their ability to locally inhibit the strength of deformation independently of the data in the image. In particular we assess the three modifications of the viscous fluid, MF1, MF2 and MF3 (varying viscosity) presented in the previous chapter. The evaluation consists of six tests. The first uses synthetic images and displays the ability of the fluid to achieve strong localised deformations, together with the varied success of the fluid modifications to restrict deformation in a selected region. The deformations are displayed on grid images and the deformations in the restricted regions are quantified using a bending metric computed by a matrix-based method.

In the second test, MF3 is applied to a real MR pre-/post-operative head image pair, demonstrating the reduction in deformation in an arbitrarily defined region. The extent of success is shown using grid images and local deformation metrics which will be described below.

The third test demonstrates the application to the modelling of hard tissue within the context of non-linear deformations. The test is not realistic since due to lack of adequate test data the target was artificially produced from the source using the combination MTQ. However it does permit additional ground truth evaluation of local registration quality due to the known correspondences of both landmark and boundary features. In this respect, the test is an extension of that in section 5.3.

The fourth test returns to the pre-/post-operative data set, this time comparing the regular fluid, MF2, MF3, TPS and modified MTQ in their ability to respect the rigidity of the scalp*. This test is of particular interest for its application in highlighting an abnormality in the post-operative target.

The final two tests are applied to 3-dimensional data sets. The fifth test compares the quality of the registration provided by the regular fluid, MF3 and SPM as well as validating MF3 for local inhibition in

*Actually it is the skull which remains rigid but the skull does not appear in MR images
the 3D case. The final test is of an inhomogeneous model-type hierarchy, where the combination MTQ is followed by MF3.

Prior to presenting the results of these tests, we first describe the metrics used to give a visual estimate of the deformation locally within the image domain.

### 7.2 Local metrics of rigidity

Results of the constant viscosity and variable viscosity fluid models are analysed using three local deformation metrics which highlight regions of high area expansion and bending. The resulting information exhibits the extent of validity of the variable viscosity model and may also be useful for automatically monitoring abnormalities during subject-atlas registration. The three local metrics are applied to the concatenated deformation fields \( \{(u_0(x_0, z_1), u_1(x_0, z_1), u_2(x_0, x_1, x_2))\} \) produced by both the fluid models. These are computed pixel-wise and displayed as images which can identify regions of high area expansion or bending.

The **elastic** differential operator is given by

\[
m_e(u_0(x_0, x_1, x_2), u_1(x_0, x_1, x_2), u_2(x_0, x_1, x_2)) = \sum_{i,j=0}^{2} \lambda \left( \frac{\partial u_i}{\partial x_i} \right) \left( \frac{\partial u_j}{\partial x_j} \right) + \mu \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right)^2
\]

Poisson’s ratio \( \sigma \):

\[
\sigma = \frac{\lambda}{2(\mu + \lambda)}
\]

gives the ratio of lateral shrink of the deforming medium to extensional strain on application of tensile forces. Setting \( \lambda \) to zero and \( \mu \) to 1, \( m_e \) is a metric of stretch or **region expansion**.

The **Laplacian** operator

\[
\Delta = \sum_{i=0}^{2} \frac{\partial^2}{\partial x_i^2}
\]

can be applied to each component \( u_i(x_0, x_1, x_2) \) of the deformation field. Taking the sum of squares of Laplacians of each component gives a **bending** metric:

\[
m_L(u_0(x_0, x_1, x_2), u_1(x_0, x_1, x_2), u_2(x_0, x_1, x_2)) = \sqrt{(\Delta u_0)^2 + (\Delta u_1)^2 + (\Delta u_2)^2}
\]

The **thin-plate bending** operator, often used in spline warps, includes mixed second-order derivatives.

The bending energy for the 2-D case was given in equation (3.13); for the 3-dimensional case we have:

\[
B = \left( \frac{\partial^2}{\partial x_0^2} \right)^2 + \left( \frac{\partial^2}{\partial x_1^2} \right)^2 + \left( \frac{\partial^2}{\partial x_2^2} \right)^2 + 2 \left( \frac{\partial^2}{\partial x_0 \partial x_1} \right)^2 + \left( \frac{\partial^2}{\partial x_0 \partial x_2} \right)^2 + \left( \frac{\partial^2}{\partial x_1 \partial x_2} \right)^2
\]

This gives our third deformation metric:

\[
m_B(u_0(x_0, x_1, x_2), u_1(x_0, x_1, x_2), u_2(x_0, x_1, x_2)) = \sqrt{(Bu_0)^2 + (Bu_1)^2 + (Bu_2)^2}
\]
7.3 2D Results

7.3.1 Synthetic labelled images

The aim of this test is to demonstrate normal and inhomogeneous fluid registrations for regions defined as normal, rigid, passive or motionless.

We created a set of four artificial images of size 256 × 256, labelled house, clown, house2 and clown2, illustrated in figure 7.1. All four contain five corresponding structures: roof (hat) of intensity 87; shadow (hair) of intensity 39; wall (face) of intensity 127; windows (eyes and mouth) of intensity 215; and background of intensity 255.

Figure 7.1: Artificial images: (1-r) clown, house, clown2, house2

To test the regular fluid algorithm, the house was deformed to the clown image with the following parameters: within any timestep, the number of SOR iterations was limited to a maximum of 70 and the SOR relaxation parameter \( \omega \) was fixed at 1.9. The deformed source (house) image was stored at every 50 timesteps, together with a grid image with the same deformation applied. At each timestep, the correlation coefficient between target and deforming source was computed; the termination criterion was a change in correlation coefficient less than 0.0001. Figure 7.2 shows the results. The starting correlation coefficient was 0.706; for the deforming source at timesteps 50, 100, 150, 200, 250, the correlation coefficients were 0.880, 0.956, 0.981, 0.987 and 0.991 respectively, with the aim being a correlation coefficient of 1.0 for a perfect registration. Deformation terminated at the 250th iteration since the change in correlation coefficient was computed at 6.25253e-05. Registration was not completed in the roof-shadow/hair region, but in all other respects the results demonstrate the power of the fluid algorithm to effect strong localised deformations. Identical results were obtained if at each timestep the SOR was allowed to iterated for a maximum of 4000 times.

The test was then repeated, registering house2 to the clown. Every 20th timestep was stored together with the deformation applied to a grid image; a maximum of 4000 SOR iterations were allowed, and the relaxation parameter \( \omega \) was fixed as before. The starting correlation coefficient was 0.645; for the deforming source at timesteps 20, 40, 60, 80, and 127 (termination), the correlation coefficients were 0.729, 0.798, 0.854, 0.887, 0.909, and 0.924 respectively. The deformed sources and grids at these timesteps are shown in figure 7.3. Termination was due to a drop in correlation coefficient (0.9242 to 0.9237).

Figure 7.4 shows the results of registering house2 to the clown image using varying viscosity (MF3), with
the door defined as rigid. The algorithm parameters were identical to those in the previous test, figure 7.3. A maximum of 4000 SOR iterations were allowed in each timestep, with the iterations terminating early if norms calculated from the residual dropped to less than 0.1% of the norms calculated at the start of each iteration cycle. For the first timestep, 656 iterations were required. Generally, around 1200-1400 SOR iterations were needed in each timestep. For the deforming source at timesteps 20, 40, 60, 80, and 118 (termination), the correlation coefficients were 0.722, 0.789, 0.846, 0.878, and 0.913 respectively. Termination was due to a drop in correlation coefficient (0.9127 to 0.9126). Note that in this case we do not require perfect registration (correlation coefficient 1.0) since we are not aiming for a deformation of the door into the shape of the clown’s mouth. Deformation in the door region has been severely inhibited, although the door has not remained entirely rigid.

The varying-viscosity fluid (MF3) was tested again, with house2 and clown2 as source and target; figure 7.5 shows the results. In this case, there was greater success in defining the door as rigid. For the first timestep, 2033 SOR iterations were required, and only 834 were required in the second timestep. Again, around 1200-1400 SOR iterations were needed in each timestep. For the deforming source at timesteps 20, 40, 60, 80, 100 and 129 (termination), the correlation coefficients were 0.612, 0.678, 0.731, 0.759, 0.780 and 0.794 respectively. Termination was due to a drop in correlation coefficient (0.7943 to 0.7939).

We then tested the constant viscosity fluid with reduced forces (MF1), figure 7.6, deforming house2 to clown2. The fluid termination criterion was altered to require a change in correlation coefficient of less than $1 \times 10^{-7}$. For the deforming source at timesteps 20, 40, 60, 80, 100 and 108 (termination), the correlation coefficients were 0.612, 0.677, 0.727, 0.753, 0.772 and 0.777. Termination was due to a drop in correlation coefficient (0.778 to 0.777). Deformation of the door is severely reduced compared to that in figure 7.3.

Finally figure 7.7 shows the deformation of house2 to clown2 with the house door defined as motionless (MF2). For the deforming source at timesteps 20, 40, 60, 80, 100 and 104 (termination), the correlation coefficients were 0.609748, 0.675, 0.724, 0.752, 0.771 and 0.773, with termination due to a drop in correlation coefficient (0.77331 to 0.77330). The door has successfully remained unaltered from the original house2 image.
In addition we noted the locations of sixteen grid points on or next to the door region of the house2 image in the undeformed grid, and of their locations in the the final deformations applied to grids in figures 7.3, 7.4, 7.5, 7.6 and 7.7. Figure 7.10 shows the door regions enlarged. For each of these five cases, we inserted the initial and final door grid-point locations into a thin-plate spline and computed its bending energy using the principal warps method given by equation (3.37); the values are given in table 7.1. The results show MF1 was more successful than MF3 in restricting deformation. However, figure 7.9 (top, centre) shows that MF3 achieves a better registration.
Figure 7.4: MF3 (Varying viscosity) house2 deformed to clown, with weakly-deformable door.

Figure 7.5: MF3 (Varying viscosity) house2 deformed to clown2, with weakly-deformable door.

Figure 7.6: MF1, house2 deformed to clown2, with door defined as a passive region.
7.3. 2D Results

Figure 7.7: MF2, house2 deformed to clown2 with door region defined as motionless.

Figure 7.8: Regular fluid (constant viscosity), house deformed to clown: deformed house subtracted from clown showing areas where registration was incomplete.

Figure 7.9: Subtraction images from deformed house2 images from (top) clown/clown2 and (bottom) original house2; (l-r) regular fluid, MF3 (weakly-deformable door; registered to clown), MF3 (weakly-deformable door; registered to clown2), MF1 (passive door), MF2 (motionless door).
7.3. 2D Results

Figure 7.10: The deformed door regions of the final grid images of each of figures (1-5) 7.3, 7.4, 7.5, 7.6 and 7.7

<table>
<thead>
<tr>
<th>Target</th>
<th>registration method</th>
<th>Bending energy</th>
</tr>
</thead>
<tbody>
<tr>
<td>clown</td>
<td>constant viscosity</td>
<td>1.018</td>
</tr>
<tr>
<td>clown</td>
<td>varying viscosity (MF3)</td>
<td>0.065</td>
</tr>
<tr>
<td>clown2</td>
<td>varying viscosity (MF3)</td>
<td>0.058</td>
</tr>
<tr>
<td>clown2</td>
<td>constant viscosity, reduced forces (MF1)</td>
<td>0.045</td>
</tr>
<tr>
<td>clown2</td>
<td>constant viscosity, zero velocities (MF2)</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 7.1: Bending energies of the deformation of the door grid points
7.3.2 Pre/post-operative head images, arbitrary weakly-deformable body

The aim of the next test was to validate the variable-viscous fluid model, that is, to visually estimate rigidity of a region labelled weakly-deformable, on more intricate data than the simple house/clown.

We used the data set head4 shown in figure 7.11, with head4A (figure 7.11, left) as the target. The images were supplied by Dr Louis Lemieux of the Institute of Neurology. They are coronal slices, at roughly the same location, of a pre-/post-operative data set, with the post-operative (figure 7.11 right) exhibiting hydracephalous and coning. The hydracephalous is apparent as gross distortions in the ventricular and left-cortical areas. Head4A had a wider intensity range than that of head4B. To avoid difference images between the two being dominated by the former as a result, pre-processing was required on head4A to rescale its intensity range to that of head4B. Additionally, we masked the images to remove noise in the background regions.

![Figure 7.11: Data set of (left) head4A (pre-operative) and (right) head4B (post-operative).](image)

We outlined an arbitrary butterfly-shaped structure in the neck region of the source head4B (highlighted in figure 7.12) which was set as weakly-deformable, having viscosity $\mu = 1.0$ in equation (3.67) and $\mu$ set to 0.1 in the remaining image area. The viscosities were supplied as an array giving $\mu$ at each point in the source image; this viscosity ‘image’ deforms concurrently with the deforming source during fluid registration. The variable viscosity model depends on derivatives of the viscosity image; as a regularisation to avoid excessively high derivatives in the case of a viscosity image with sharp edges, and to ensure that neighbouring fluid regions would not compromise the rigid boundaries, we partially extended the

![Figure 7.12: Source head4B with arbitrary weakly-deformable region highlighted and outlined.](image)

\[ \text{The rescaling was a simple linear mapping, as follows. Let } M_B \text{ and } m_B \text{ denote the maximum and minimum (respectively) pixel intensities in head4B, and } M_A \text{ and } m_A \text{ the maxima and minima respectively in head4A. Then the intensity } i_A \text{ of a pixel in head4A was remapped to an intensity } i_B + (i_A - m_A) \left( \frac{M_B - m_B}{M_B - m_A} \right). \]

\[ \text{using xdispunc tools.} \]
7.3. 2D Results

weakly-deformable regions using a sine-based function (equation 7.6) of the 4SED Euclidean distance transform (appendix A.1), with the parameters set empirically at $a = 1, b = 13$.

$$f(x) = \begin{cases} 
1 & : x \leq a \\
\frac{1}{4} \left(1 + \cos\left(\frac{\pi(x-a)}{b-a}\right)\right) & : a < x < b \\
0 & : x \geq b 
\end{cases} (7.6)$$

Figure 7.13 shows the results of fluid deformation after 40 timesteps, together with the results for the constant viscosity model with $\mu$ set to 0.1 for the entire image. Figure 7.14 shows the magnitudes of the resulting displacements; the deformations of the weakly-deformable region are shown again on the grid images in figure 7.15. The grid lines indicate that the arbitrary structure has deformed less using the

Figure 7.13: Deformed source images after 40 timesteps of fluid registration to the target using the (left) constant and (right) variable viscosity models.

Figure 7.14: Magnitude of displacements: (left) constant viscosity model; (right) variable viscosity. High values are coloured black.

Figure 7.15: Grid images showing the deformation of the arbitrary weakly-deformable region; (centre) constant viscosity model, (right) variable viscosity model.
variable viscosity model, than when the constant viscosity model was applied.

Figure 7.16 shows the application of the local elastic metric $m_e$ to the deformation field. Regions of high intensity indicate clearly large area expansion where the right brain hemisphere has been deformed to its pre-operative shape.

Figure 7.16: The elastic deformation metric: (left) constant viscosity, (right) varying viscosity (high values coloured white). Both clearly show regions of high area growth, indicating locations of abnormality in the source image.

Figures 7.17 and 7.18 show the Laplacian and thin-plate bending metrics respectively. Both show near-zero values within the region defined as weakly-deforming for the variable viscosity model, while indicating bending within the same region for the constant viscosity model. However, as the grid image shows that the weakly-deforming region does exhibit some deformation, it cannot be classified as rigid.

Figure 7.17: The Laplacian bending metric: (left) constant viscosity, (right) varying viscosity. The weakly-deformable region shows little to no bending in the varying viscosity model, in contrast to the constant viscosity model.
Figure 7.18: The thin-plate bending metric: (left) constant viscosity, (right) varying viscosity, showing more clearly than the Laplacian the difference in bending between the constant and variable viscosity models.
7.3.2D Results

7.3.3 Artificially-warped knee images

The previous experiment tested the bending of regions which were roughly at the same location in the source and target images. We then compared the constant and variable viscosity fluid models to an image pair where rigid bodies needed considerable displacement to register to their homologues in the target. The aim again was to reduce deformations in areas known to be rigid (hard tissue) and to see if this improved on the quality of registration. An additional feature of this particular test was that the exact locations of a set of corresponding landmarks and of the boundaries of the rigid bodies were known in both the source and target and thus the quality of registration locally could be assessed.

The data set consisted of the images shown in figure 6.1. The source (figure 6.1, left) was a real MR image of a knee, sized 256 × 256, and the target was created from it by a combination affine-multiquadric warp using 6 landmarks of which 3 remained fixed, and incorporating three independently-displaced rigid bodies (bones), labelled *patella*, *upper* and *lower*. Figure 7.20 illustrates the labelling of the source image.

Before registration, the source and target were each enclosed within a margin 56 pixels wide; figure 7.19 shows the source knee in its margin. This was to move the bones away from the boundary of the image where velocities (and thus also displacements) were fixed at zero as boundary conditions supplied to the fluid PDE. However, since the deformations in the margin region are of no interest, the images of the results and metrics are shown with the margin cropped.

The source was deformed to the target within a closure scale space, shown in figure 7.21, with the three rigid bodies allocated viscosities of 1.0 in the variable viscosity fluid, and the remaining pixels and all the pixels in the constant viscosity case having viscosities of 0.1. We did not optimise the choice of scale space.

Figure 7.22 shows the final results of registering source to target in the constant and variable viscosity cases respectively. The constant viscosity case shows an almost perfect registration, whereas the variable viscosity case shows incomplete registration around the upper half of the upper bone, at the lower bone and in the tissue to its right. The grid image for the variable viscosity case shows little distortion in the grid lines in the regions of the bones; the constant viscosity case shows higher frequency grid-line distortions in the lower half of the grid.

---

\[\text{we found that registration terminated prematurely without scale space - the correlation coefficient started to decrease.}\]

\[\text{in the conventional sense.}\]
Figure 7.20: Locations of the six landmarks and boundaries of three rigid bodies (labelled) in the source knee image.

The Laplacian, bending energy and elastic energy deformation metrics were then computed pixel-wise for both fluid models, together with the magnitudes of displacements; these are shown in figure 7.23. The three higher-order deformation metrics all show lower values in the regions of the bones for the variable viscosity case.

To quantify the relative extent of rigidity in the bones in the variable-viscosity case, table 7.2 lists the maximum energy value over the image for each of the deformation energy metrics for both fluid cases. We conclude that the variable-viscosity fluid is successful at reducing bending in pre-segmented regions.

<table>
<thead>
<tr>
<th>energy measure</th>
<th>maximum energy value</th>
<th>log_e of maximum energy value</th>
</tr>
</thead>
<tbody>
<tr>
<td>constant viscosity:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laplacian</td>
<td>16900</td>
<td>9.74</td>
</tr>
<tr>
<td>bending</td>
<td>1.83e+08</td>
<td>19.0</td>
</tr>
<tr>
<td>elastic</td>
<td>179000</td>
<td>12.1</td>
</tr>
<tr>
<td>(magnitude of deformations)</td>
<td>34.2</td>
<td>-</td>
</tr>
<tr>
<td>variable viscosity:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laplacian</td>
<td>3460</td>
<td>8.15</td>
</tr>
<tr>
<td>bending</td>
<td>8.08e+06</td>
<td>15.9</td>
</tr>
<tr>
<td>elastic</td>
<td>13900</td>
<td>9.54</td>
</tr>
<tr>
<td>(magnitude of displacements)</td>
<td>19.0</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 7.2: Maximum values of deformation metrics for the constant and variable viscosity cases of the knees test.
We then examined the quality of registration produced by each fluid model. In this test the exact locations of six homologous landmarks in the image pair were known and were compared to their locations in the deformed source images. Table 7.3 lists the results for the landmarks as numbered in figure 7.20.

<table>
<thead>
<tr>
<th>landmark number</th>
<th>original x</th>
<th>original y</th>
<th>regular fluid x</th>
<th>regular fluid y</th>
<th>MF3 x</th>
<th>MF3 y</th>
<th>target x</th>
<th>target y</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>255</td>
<td>0</td>
<td>253</td>
<td>4</td>
<td>251</td>
<td>5</td>
<td>255</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>255</td>
<td>6</td>
<td>260</td>
<td>4</td>
<td>259</td>
<td>0</td>
<td>255</td>
</tr>
<tr>
<td>4</td>
<td>224</td>
<td>224</td>
<td>267</td>
<td>198</td>
<td>259</td>
<td>206</td>
<td>250</td>
<td>206</td>
</tr>
<tr>
<td>5</td>
<td>224</td>
<td>173</td>
<td>242</td>
<td>162</td>
<td>242</td>
<td>166</td>
<td>235</td>
<td>168</td>
</tr>
<tr>
<td>6</td>
<td>228</td>
<td>127</td>
<td>199</td>
<td>142</td>
<td>201</td>
<td>145</td>
<td>209</td>
<td>137</td>
</tr>
</tbody>
</table>

Table 7.3: Coordinates of known source and target landmarks, with final locations of positions of the source landmarks after registration using both the regular and varying-viscosity fluid algorithms.

In addition, the fluid deformations were applied to the binary segmented bone images which had been supplied to the variable-viscosity fluid to define the regions of higher viscosity, and their boundaries outlined using the automatic contouring tools in the xdispuncell image display package. These were compared to the boundaries of the undeformed binary bone images. The combination rigid body/multiquadric which had created the target knee image from the source was applied to the source binary bone images and the boundaries of the resulting target bone regions outlined as well. Figure 7.24 shows the source, deformed source and target images with the known or deformed landmarks and region boundaries superimposed.

---

[See footnote on p130]
7.3. 2D Results

Figure 7.22: Registration of knee image to its warped version, using (top) constant viscosity, (bottom) variable viscosity fluid models: (1-r) deformed source; corresponding grid; difference image of deformed source from target.

The Chamfer metric of distance between two closed curves was applied to the sets of boundaries, using a program written by Dr Julia Schnabel. The Chamfer metric applies a distance transform (in this case the 5-7-11 Chamfer distance transform [Borgefors, 1986]) to boundary $A$ and lays boundary $B$ on the resulting field of distances; for each point $b \in B$, its distance from $A$ is then read off the underlying distance field. The mean, root mean squared (RMS) and maximum distances are computed for the set $\{b\}$.

Table 7.4 lists the distance metrics computed for the bone boundaries. The distances of deformed object boundaries to target object boundaries would equal zero for a perfect (and rigid) deformation. Furthermore, the distances of the deformed source boundaries to the undeformed source boundaries would equal the distances of source to target boundaries for a perfect registration. The Chamfer metrics show a better registration of bone boundaries using the constant viscosity model.

Finally, the registration of source to target landmarks and bone boundaries are summarised in figure 7.25. We conclude that in this test, although the variable-viscosity fluid was successful in respecting the rigidity of the bones, a better registration was obtained by the constant-viscosity fluid model. Although the bones were displaced vertically and horizontally towards the target by the variable-viscosity model, they were not supplied with the rotations required to bring them into register with the target.
Figure 7.23: Metrics of deformation of kneeB to kneeA, (l-r) Laplacian; elastic; bending; transformation, for (top) the constant viscosity and (bottom) variable viscosity cases.

Figure 7.24: Knee images with landmarks and bone boundaries marked: (l-r) source; source deformed using constant-viscosity fluid; source deformed by MF3 with viscosity set at 1.0 within bones and 0.1 elsewhere; target.
Table 7.4: Distance metrics applied to boundaries of bone regions.

<table>
<thead>
<tr>
<th>fluid type</th>
<th>comparison with</th>
<th>Chamfer: Mean</th>
<th>RMS</th>
<th>max</th>
</tr>
</thead>
<tbody>
<tr>
<td>lower bone:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>regular</td>
<td>target</td>
<td>0.572</td>
<td>0.0723</td>
<td>4.2</td>
</tr>
<tr>
<td>MF3</td>
<td>target</td>
<td>5.63</td>
<td>0.553</td>
<td>17.2</td>
</tr>
<tr>
<td>none</td>
<td>target</td>
<td>14.6</td>
<td>1.27</td>
<td>31.8</td>
</tr>
<tr>
<td>regular</td>
<td>source</td>
<td>13.5</td>
<td>1.24</td>
<td>31.8</td>
</tr>
<tr>
<td>MF3</td>
<td>source</td>
<td>10.2</td>
<td>0.892</td>
<td>17</td>
</tr>
<tr>
<td>upper bone:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>regular</td>
<td>target</td>
<td>0.583</td>
<td>0.0781</td>
<td>5</td>
</tr>
<tr>
<td>MF3</td>
<td>target</td>
<td>3.75</td>
<td>0.355</td>
<td>12.6</td>
</tr>
<tr>
<td>none</td>
<td>target</td>
<td>7.39</td>
<td>0.581</td>
<td>16</td>
</tr>
<tr>
<td>regular</td>
<td>source</td>
<td>7.29</td>
<td>0.660</td>
<td>17.6</td>
</tr>
<tr>
<td>MF3</td>
<td>source</td>
<td>5.27</td>
<td>0.484</td>
<td>13.6</td>
</tr>
<tr>
<td>patella:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>regular</td>
<td>target</td>
<td>0.455</td>
<td>0.0790</td>
<td>2</td>
</tr>
<tr>
<td>MF3</td>
<td>target</td>
<td>0.0756</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>none</td>
<td>target</td>
<td>5.86</td>
<td>0.767</td>
<td>12</td>
</tr>
<tr>
<td>regular</td>
<td>source</td>
<td>5.57</td>
<td>0.735</td>
<td>12.2</td>
</tr>
<tr>
<td>MF3</td>
<td>source</td>
<td>5.66</td>
<td>0.773</td>
<td>13.2</td>
</tr>
</tbody>
</table>
Figure 7.25: Source, known target and deformed source landmarks and boundaries of bone regions for the comparison of the regular fluid and MF3 on the knee test image.
7.3. 2D Results

7.3.4 Pre/post-operative head images, rigid scalp

The next experiment attempted to reduce deformation at an area where differences are known a priori in order to highlight other areas where there are unknown differences due to abnormalities. The source and target were again chosen to be head4A (target) and head4B (source), figure 7.11. Since these are slices through the same subject at approximately, but not exactly, the same location, they exhibit slight differences in the shape of the scalp. Although the same applies to the cortex and the internal structures of the brain, the main cause of their shape differences between source and target was the surgical procedure and its after-effects. It is these differences which would be classified as abnormal. We wished therefore to find the best method of registering the source to the target such that the abnormalities, rather than the known differences, were highlighted. We proposed the following method.

The target (pre-operative) image was to be used as a template or atlas image, defining the normal brain shape for that subject, and prior information of known abnormalities was to be supplied as regions where deformation was to be reduced using one of the inhomogeneity paradigms described in section 6.2. The source was to be registered to the target and the strength of deformation, as estimated from the magnitude image of the resulting displacements field and from images of the Laplacian/bending/elasticity energy metrics, would highlight regions of severe distortion of the source image from the normal brain shape. Assuming registration to the target in unknown regions (regions for which there was no prior information about their abnormality) was optimal, a further indication of abnormality would be the difference of the deformed source from the undeformed source, since this difference image would display differences in the unknown regions only between the shape of the source and of the normal shape as defined by the target. Differences in regions with known abnormalities would not be highlighted since their registration to the target would be suppressed. As a check that registration had been complete in the unknown regions and suppressed in the known regions, the deformed source would also be subtracted from the target. In this case, for an ideal registration, the difference image would be zero in unknown regions and non-zero in regions of known differences.

In this test we defined the scalp as weakly-deformable. The scalp region was segmented manually with the aid of the display tool xdispune. We compared constant-viscosity fluid registration; variable-viscosity fluid registration (MF3); registration using a single rigid body within a multiquadric spline, with landmarks derived using the variable-viscosity fluid registration; global thin-plate spline warp using landmarks derived using the constant viscosity fluid registration; and constant-viscosity fluid with zero-velocity boundary conditions set on the scalp (MF2). For the combination-multiquadric test, the automatic landmark-finding algorithm was modified so that landmarks lying on the labelled region (scalp) were not accepted for tracking by the fluid. This is because the version of the combination-multiquadric given in [Little et al., 1997] required that landmarks did not lie on any rigid body.

For each of the chosen registration paradigms, we collected results of the deformed source image, $S(\bar{u})$, of the same deformation applied to a regular grid image, $G(\bar{u})$, of the subtraction of the deformed source from the target, $T - S(\bar{u})$ and of the subtraction of the deformed source from the undeformed source,
The ideal inhomogeneous registration paradigm for this case would exhibit no deformation on the scalp region (painted white on the grid prior to deformation) in \( G(\bar{u}) \), would highlight differences in the scalp region in \( T - S(\bar{u}) \) and would highlight differences in the ventricular and left-cortical areas in \( S - S(\bar{u}) \).

The results are shown in figures 7.26, 7.27, 7.28, 7.29, 7.30 and 7.31. It is apparent that the constant fluid with zero-velocity boundary conditions (MF2) set for the scalp (figure 7.31) met the criteria best for this test. The centre right image shows the deformed source (post-operative image) subtracted from the target (the pre-operative image) showing differences in the scalp region, and little structure in the brain region. Hence registration has been successful within the brain region while the scalp has been constrained to remain rigid and so the scalp of the source has not deformed into the shape of that of the target. The right-hand image is of the deformed source (post-operative image) subtracted from the undeformed source. Since no deformation was allowed in the region of known difference (the scalp), the subtraction image shows only differences in regions for which there was no information known \textit{a priori}. In this case it clearly shows strong deformation in the ventricular and hydracephalous regions. We consider this figure to be an exciting result in providing known information to the varying-viscosity to identify more clearly abnormalities in a source image.

As an alternative to the grid images, figure 7.32 illustrates the deformations of the scalp and remaining regions for the variable-viscous fluid, TPS and combination-MTQ registrations. We painted a binary-intensities features image consisting of the scalp region and landmarks obtained using the corner-finding/grid algorithm (figure 7.32, left), applied to it the deformations produced by the constant viscosity fluid, by the TPS summary of the fluid and by the MTQ combination, and subtracted these from the undeformed binary image (figure 7.32, remaining images). This gave a clearer picture of the deformation of the scalp by each of these three registrations. The centre two images clearly display warping of the scalp, while the combination MTQ (as expected from its definition) shows virtually no deformation in this area.

Finally we compared the local deformation of the regular fluid to that of MF3 using the rigidity metrics in section 7.2. Figures 7.33 and 7.34 show the results, together with an image of the magnitudes of deformation at each pixel. For the Laplacian metric, we display the logs of the metric values. The difference in the bending in the scalp region between the regular fluid and MF3 is clear: in the MF3 result, the scalp region shows almost zero value for the metric. The elastic and deformation magnitude metrics highlight the strong deformations in the hydracephalous regions; we suggest that application of these metrics following a registration would be useful in identifying unusually strong deformations as a check for abnormalities in the images.

The final two evaluations were performed on 3-dimensional data. In the next section we detail the image acquisition and the pre-processing steps applied to it, and in the following two sections we describe the tests and analyse the results.
Figure 7.26: Constant viscosity fluid, post-op registered to pre-op. (l-r): deformed post-op; deformation applied to grid; difference image - deformed post-op subtracted from target (pre-op) showing areas where registration was incomplete, or remaining differences with the atlas; difference image - deformed post-op subtracted from post-op highlighting differences due to the deformation.

Figure 7.27: Constant viscosity fluid, post-op registered to pre-op, summarised/approximated by TPS. (l-r) TPS-deformed post-op; TPS deformation applied to grid; TPS-deformed post-op subtracted from pre-op; TPS-deformed post-op subtracted from post-op.

Figure 7.28: Varying viscosity fluid (MF3), post-op registered to pre-op. (l-r): deformed post-op; deformation applied to grid; deformed post-op subtracted from pre-op; deformed post-op subtracted from post-op.
Figure 7.29: Varying viscosity fluid, post-op registered to pre-op, summarised/approximated by TPS. (l-r) TPS-deformed post-op; TPS deformation applied to grid; TPS-deformed post-op subtracted from pre-op; TPS-deformed post-op subtracted from post-op.

Figure 7.30: Varying viscosity fluid, post-op registered to pre-op, summarised/approximated by rigid-MTQ. (l-r) rigid-MTQ-deformed post-op; rigid-MTQ deformation applied to grid; rigid-MTQ-deformed post-op subtracted from pre-op; rigid-MTQ-deformed post-op subtracted from post-op.

Figure 7.31: Boundary conditions (MF2): zero velocities on scalp. (l-r): deformed post-op, deformation shown on grid; deformed post-op subtracted from pre-op; deformed post-op subtracted from post-op.

Figure 7.32: Constant viscosity fluid, post-op registered to pre-op. (l-r) rigid body and selected markers (features image); fluid-deformed features image subtracted from original; TPS-warped features image subtracted from original; rigid/MTQ-warped features image subtracted from original.
Figure 7.33: Constant viscosity fluid, post-op registered to pre-op, deformation metrics (l-r): log of the Laplacian, bending, elastic, magnitude of transformation.

Figure 7.34: Varying viscosity fluid (MF3), post-op registered to pre-op, deformation metrics (l-r): log of the Laplacian, bending, elastic, magnitude of transformation.
7.4 3D Results

7.4.1 3D images

3-dimensional versions of non-linear, hierarchical and inhomogeneous registration algorithms were tested on MRI neck volumes. The original images were acquired at the Hammersmith hospital, courtesy of Professor Joseph Hajnal and Anjela Oatridge. Three full-3D neck volumes were provided, neckD, neckO and neckI. NeckI was of the chin down, neckD was with the head flexed backwards within the confines of the scanner bore, and neckO was of the head turned to the right; all were of the same volunteer subject. Table 7.5 details the image acquisition parameters.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>scanner</td>
<td>Picker 1.0T HPQ</td>
</tr>
<tr>
<td>acquisition protocol</td>
<td>a cervical-spine quadrature surface coil was used for reception</td>
</tr>
<tr>
<td>field of view (FoV)</td>
<td>30 cm (around neck vertebrae)</td>
</tr>
<tr>
<td>inter-slice spacing</td>
<td>2 mm</td>
</tr>
<tr>
<td>acquired image dimension</td>
<td>192 × 255 × 38</td>
</tr>
<tr>
<td>stored image dimensions</td>
<td>Fourier interpolated to 256 × 256 × 38</td>
</tr>
<tr>
<td>voxel dimensions</td>
<td>1.17188 × 1.17188 × 2 mm</td>
</tr>
<tr>
<td>data type</td>
<td>16-bit</td>
</tr>
</tbody>
</table>

Table 7.5: Parameters of image acquisition for the neck images.

Due to the acquisition method, the images exhibited a strong intensity ramp, with high values at the back of the neck and total signal loss in the face (figure 7.35 top left shows the central slice of neckD). We pre-processed the images to give a more uniform range of intensities in the anterior-posterior direction. Our method was to produce for each neck volume a reference image by blurring the neck image with a Gaussian of spatial standard deviation $\sigma = 5$. The neck volume intensities were then divided pixelwise by the intensities of the corresponding locations in the reference image (figure 7.35, top right). However this also enhanced background noise and so the images were masked semi-automatically using a contouring tool in the image display package xdi.spune, with manual corrections in the face area where there was weak contrast between tissue signal and background noise (figure 7.35, bottom left). Due to constraints of time and memory capacity, it was not feasible to apply the fluid registration algorithms of section 7.4.2 directly to the full-resolution data sets. The data sets were therefore downsampled, by blurring with a Gaussian of standard deviation $\sigma = 2$ and storing alternate pixels (figure 7.35, bottom right). The Gaussian blurring was performed in the Fourier domain, computed by an FFT which required image dimensions of powers of 2. Similarly the fluid algorithm required gradients of the image computed in the Fourier domain. Hence before downsampling, the data sets were padded by blank slices at $z = 1$–13 and $z = 51$–64, giving full resolution volumes of dimensions $256 \times 256 \times 64$ and downsampled images of dimensions $128 \times 128 \times 32$. We chose to pad with blank slices rather than duplicating the data slices since we did not wish to confuse the applied registration algorithms with additional data.
For the tests described in section 7.4.2 we required a segmentation of the spinal vertebrae from the neck volumes. We produced these slice-by-slice on the full-resolution images using the xdispunc display tools. Generally the contrast between vertebrae and intervening tissue was poor and so segmentation was performed manually. Figure 7.36 (top left) shows the segmentation of vertebrae from the central slice (slice 32). The spinous processes were omitted from the vertebrae segmented from neckD, but included in those of necks I and O after studying an atlas [McMinn et al., 1994]; due to time constraints it was deemed unfeasible to repeat the segmentation of neckD. After manually outlining each vertebra, it was masked and intensities in the vertebral region set to 256, producing a binary image of the spine (figure 7.36 top right shows the central slice). The binary spine volumes were downsampled using the same process.
as for the neck volumes. Figure 7.36 (bottom left) shows the central slice of the downsampled volume of neck D. Due to the partial-volume effect caused by the downsampling, the low-resolution spine volumes were no longer binary but contained a range of intensities, as shown in the 19th slice, (figure 7.36 bottom right).

Note that since the imaging field of view had been the spinal column specifically, the volumes did not extend to include the whole neck laterally. In addition, due to the attenuation of signal anteriorly, much of the facial structure was missing. This accounts for the unusual appearance of the volume-rendered representations of the data (produced using the Mayo package Analyze) in figures 7.42, 7.43 and 7.49.

Figure 7.36: Segmentation of vertebrae slice-by-slice from neckD. Top left: vertebrae outlined manually using the xdispunc regions tool (slice 32); right: resulting binary image after masking each vertebra and colour processing in xv; bottom left: the same spine slice downsampled by a half; slice 19 of the zero-padded and downsampled 128 × 128 × 32 spine volume
7.4.2 Comparison of fluid / variable-viscous fluid / SPM

We compared the regular fluid, variable-viscosity fluid and SPM registration on two pairs of the downsampled neck images: neckD to neckl and neckO to neckl. The variable-viscosity fluid used the downsampled spineD or spineO images to define the weakly-deforming regions. Our initial hypothesis was that constraining the deformations of the vertebrae, while allowing them to move, would improve on the quality of registration since this was a more accurate model of the underlying physical movements responsible for the differences between the source and target images.

The parameters set for all four fluid registration tests (constant/variable viscosity, neckD/O to neckl) were:

1. six-level scale space (Gaussian blurs of spatial standard deviation $2^i$ where $i = \{5, 4, 3, 2, 1, 0\}$.
2. within each timestep, a maximum of 50 iterations were allowed in the SOR solution of the fluid PDE (due to time constraints)
3. within each scale level, the fluid was set to perform at least three iterations, with an optional extra 100 iterations until the stopping criterion was met.
4. within each scale level, the stopping criterion was a reduction in correlation coefficient of less than $+0.0001$.

The initial three iterations (parameters 3) were to ensure that some change had indeed occurred especially in the first level of scale space. At all six levels of scale space, extra looping did occur in all four tests after the initial three iterations, indicating that the initial loops had not been redundant.

The 3D SPM registrations were provided for comparison by John Ashburner of the Functional Imaging Laboratory, having been converted to the Mayo Analyze image format and stored as 8-bit pixels. This had implications for the global pixel-wise measures of registration when comparing the results to those of fluid registration: the SPM-deformed volumes returned for comparison had an intensity range of 0–32, whereas the fluid-deformed volumes had an intensity range of 0–255, and hence a sum of squared differences** was inappropriate for the comparison; we give instead the correlation coefficient. These are given in table 7.6.

<table>
<thead>
<tr>
<th>source image</th>
<th>registration</th>
<th>fluid</th>
<th>var-fluid</th>
<th>SPM</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>0.830</td>
<td>0.962</td>
<td>0.963</td>
<td>0.925</td>
</tr>
<tr>
<td>O</td>
<td>0.865</td>
<td>0.961</td>
<td>0.960</td>
<td>0.929</td>
</tr>
</tbody>
</table>

Table 7.6: Correlation coefficients results of registering neckD to neckl (top row) and neckO to neckl (bottom row).

The correlation coefficients indicate that both fluid models perform better than SPM on these two test cases. The results are inconclusive for the regular/variable fluid comparison: the variable-viscosity fluid

**Ideally the sum of squared differences should be used since this was used by both algorithms to derive the likelihood
outperforms the constant-viscosity fluid for the first test, whereas the latter performs slightly better in the second test.

To compare visually the registration quality of the SPM and fluid algorithms, we show the central slice of each deformed volume (figures 7.37 and 7.40), their difference from the central slice of the target (figures 7.38 and 7.41) and the outer surfaces of the volumes (figures 7.42 and 7.43). The last two figures were obtained by converting the source, target and fluid-registered source from UNC format to the Mayo Analyze format and using the Analyze volume rendering tools: we chose the option of voxel gradient shading, with a light source at a vertical angle 30°, horizontal angle 25°. The rendered images were captured using a snapshot tool, converted back to UNC format and then were all cut to equal size with the midpoints of their diameters at the image centres. This process was preferred to using an automatic cropping tool which may have given images of different sizes; insertion into this document created by \LaTeX \ would then have scaled each individually to fit into uniform output images, preventing a fair comparison of the volumes.

The rendered volumes show both fluid models have registered the external boundaries of the volumes better than the SPM (for example, look at the chin regions in the profiles and at the ‘facial’ details in the central columns of figures 7.42 and 7.43); the SPM has introduced unnecessary deformations on the outer surfaces evident in the left columns.

The central slice images (figures 7.37, 7.40, 7.38 and 7.41) show that internally, a better match has been achieved by SPM, especially in the trachea; in these figures we see again the fluid is good at registering the skin surface at the back of the neck and the chin, but note also that unnatural deformations have occurred in the brain posteriorly in figure 7.37.

We conclude that the fluid registrations matched the strong outer boundaries well but matching of internal structures was poor and SPM was superior in this respect; in this test the internal structures were of most interest and the outer boundaries were irrelevant since they were arbitrary, depending on the manually-assisted contouring in the pre-processing of the images and the extent of the region of interest (ROI) imaged by the scanner, and were mainly not outer anatomical surfaces.

Figure 7.37: Slice 16 (of the 32-slice downsampled volumes) of the images, left to right: D; fluid D/I; variable-viscous fluid D/I; SPM D/I; I

We upsampled the displacement fields obtained from both fluid registrations of neckD to neckI and applied them to the original neckD images, to give full-resolution deformations. These are shown in figure 7.39.
7.4. 3D Results

Figure 7.38: Slice 16 (of the 32-slice downsampled volumes) of the difference images, left to right: D/I; fluid D/I; variable-viscous fluid D/I; SPM D/I

Figure 7.39: Central slices of full resolution neckD (left) and neckl (right). Upsampled displacement fields obtained from (centre-left) constant-viscosity and (centre-right) varying-viscosity fluid registration of downsampled neckD to downsampled neckl.

Figure 7.40: Slice 16 (of the 32-slice downsampled volumes) of the images, left to right: O; fluid O/I; variable-viscous fluid O/I; SPM O/I; I

Figure 7.41: Slice 16 (of the 32-slice downsampled volumes) of the difference images, left to right: O/I; fluid O/I; variable-viscous fluid O/I; SPM O/I
Figure 7.42: (Top to bottom) NeckD; fluid registration of neckD to neckl; SPM registration of neckD to neckl; variable-viscosity fluid registration of neckD to neckl; neckl.
Figure 7.43: (Top to bottom) NeckO; fluid registration of neckO to neck1; SPM registration of neckO to neck1; variable-viscosity fluid registration of neckO to neck1; neck1. Rendered voxel gradient shading, using the Analyze package, with light source at vertical angle 30°, horizontal angle 25°.
We applied the transformation fields produced by all four fluid tests to the initial spine volumes segmented from neckD and neckO (spineD and spineO respectively) and rendered these using the Analyze volume-rendering tools. The results are shown in figures 7.44 and 7.46. Figure 7.45 shows the upsampled deformation fields from the fluid deformations of neckD to neckI applied to the full-resolution segmented spineD. The constant viscosity fluid registration shows an extension of the upper two vertebrae of the spines on comparison to the original segmentation from neckD. The spines have both moved slightly towards that of neckI, although the segmentations were only rough so the comparison is not exact.
Figure 7.44: (top) spineD segmented from neckD; (upper middle) spineD with regular fluid deformation of neckD to neckl; (lower middle) spineD with variable-viscosity fluid deformation of neckD to neckl with vertebrae ‘rigid’. For comparison, spine is shown segmented from neckl (bottom). The 3D images were generated from sets of 2D slices by voxel gradient shading volume rendering using the Analyze package.
Figure 7.45: (left-right) central slice of binary spine image segmented from neckD; central slice of deformed spine segmented from neckD computed by applying upsampled deformation field from constant viscosity fluid registration of neckD to neck1; and from variable viscosity fluid registration; for reference only: binary spine image segmented roughly from neck1 (this segmentation took no part in the fluid registrations)
Figure 7.46: (top) spineO segmented from neckO; (upper middle) spineO with regular fluid deformation of neckO to neckI; (lower middle) spineO with variable-viscosity fluid deformation of neckO to neckI with vertebrae ‘rigid’. For comparison, spineI is shown segmented from neckI (bottom). The 3D images were generated from sets of 2D slices by voxel gradient shading volume rendering using the Analyze package.
Figures 7.47 and 7.48 show the logs of the Laplacian and elasticity energies as bending metrics on the deformation fields obtained from both types of fluid registration on neckD and neckO respectively registered to neckI, and also shows the magnitude of the displacement fields. The latter show no clear difference to differentiate the fluid models; however the bending metrics clearly show a strong reduction in distortion in the vertebral regions in the variable-viscosity fluid.

Figure 7.47: Slice 16 (of the 32-slice downsampled volumes) of the metric images, D to I for (top row) constant viscosity fluid, left to right: log of Laplacian; log of elasticity energies; magnitude of transformations. Bottom row shows the equivalent for the varying viscosity case.

Figure 7.48: Slice 16 (of the 32-slice downsampled volumes) of the metric images, O to I for (top row) constant viscosity fluid, left to right: log of Laplacian; log of elasticity energies; magnitude of transformations. Bottom row shows the equivalent for the varying viscosity case.
7.4.3 A model-type hierarchy - variable-viscous fluid following rigid-MTQ

The final test demonstrates a hierarchy of varying model type where both models used are spatially inhomogeneous, and its failure due to each model aiming to register a different feature set.

The UMDS group at Guys and St Thomas’ Hospital registered neckI to neckD with the segmented vertebrae of spineD defined as rigid bodies within a 3D multiquadric warp using manually selected landmarks. We then applied the variable-viscous fluid to their deformed neckI to test whether the fluid would improve on their registration to neckD. We note first that the data sets neckI and neckD contain overlapping but not identical structures since the scans were of a region of interest centered on the spine but not extending to cover the entire head. The rigid/multiquadric deformation matches the rigid bodies and manually-specified biologically-corresponding structures whereas the fluid registration has no such knowledge but matches nears overlapping isocontours. In the case of fluid registration within Gaussian scale space, the outer image boundaries are registered first; these may not necessarily represent corresponding biological structures.

Figure 7.49: (Top to bottom) NeckI; registration of neckI to neckD by Graeme Penney using rigid vertebrae inside 3D multiquadric; neckD. Rendered voxel gradient shading, using the Analyze package, with light source at vertical angle 30°, horizontal angle 25°.

The multiquadric combination matched exactly the vertebral surfaces, and manually-selected homologous landmarks within the volumes - that is, matched the internal structure, and allowed the external boundaries to freely deform. This is because a spline is strictly an interpolant and does not constrain the mapping to register regions outside the envelope of the landmark locations. Hence after registration by the combination multiquadric, the outer boundaries of the deformed source I were very different from the original outer boundary of I and from the outer boundary of the target D.
Figure 7.50: Registration of 1 to D by Graeme Penney. Shown are (l-r) slices 14, 16, 19 and 32 of the 256 × 256 × 64 full-resolution volume. Slices 14 and 19 show parts of the volume outside of the envelope bounding the control points, which have undergone deformations extrapolating from the registration within the envelope, constrained only by the bending energy of the spline but not constrained by any similarity criteria.

We then tested the application of a registration hierarchy of model type. We took the result of combination multiquadric registration of neck 1 to neck D and followed it with a variable-viscous fluid registration to D. Since the spinal vertebrae of the original source had been registered exactly by the MTQ combination to those in D, we used the binary spinal segmentation of D as the array determining the viscosities to supply to the fluid.

The MTQ combination had been performed on the full resolution 256 × 256 × 64 neck volumes. However due to time constraints and memory capacity of the workstation used, it was not feasible to proceed with the fluid registration on the full volume images, and so they were first downsampled to volumes of size 128 × 128 × 32. On termination of the fluid registration, the deformation field was then upsampled to full resolution and applied to the full resolution result of the multiquadric registration.

The fluid attempts to match boundaries in the image pair. As a result, the fluid produced deformations at the outer surfaces to bring them into register, which were propagated using the fluid model throughout the deforming source volume. Since this registration was performed initially at a low level of scale space (to avoid misregistrations at local minima), much of the internal structure was missing, and hence there were no restrictions on the fluid pushing the internal structure out of register. Figure 7.51 (left and centre left) show the half- and full-resolution central slices of the neck volume 1 after deformation by the fluid registration to D, following the combination multiquadric registration to D. Comparison of this slice to the central slice of the target D (figure 7.51, right) and to the central slice of 1 registered to D by the combination multiquadric (figure 7.50, right), several differences are obvious due to the displacement of the internal structures. The location of slice 32 of the multiquadric-registered volume 1 was moved by the fluid to slice 34 (figure 7.51 centre right).

This completes our evaluation of inhomogeneous registration algorithms and the main body of the dissertation. The final chapter summarises the material presented and suggests topics for future research.
Figure 7.51: Registration of I to D by variable-viscous fluid following registration by Graeme Penney (fig. 7.49. Shown are (l-r) the central slice (slice 16) of the downsampled registered volume; the same central slice (slice 32) at the highest resolution using the upsampled displacement field; slice 34 of the highest resolution registered source volume; slice 32 of the full-resolution volume D (target).
Chapter 8

Conclusions and Future Work
8.1 Summary

Our contributions to the body of medical image registration knowledge consists of three parts: an extensive analysis of algorithms, modifications to the fluid algorithm and a small evaluation of non-linear methods.

8.1.1 Analysis of algorithms

This dissertation has provided an extensive structured analysis of the wide variety of registration algorithms developed for use in medical image applications. A summary of the review is given in figure 8.1.

The main grouping was according to their level of sophistication, or their ability to reproduce in a robust manner deformations required to register any likely source and target image pair encountered in a clinical or medical-research situation. Thus we differentiated between linear and non-linear methods, and devoted a separate chapter to non-linear algorithms that discriminate between regions in respect to their assumed deformability.

The least flexible of the algorithm types, those applying linear transformations only, may nevertheless be ideal in certain intra-subject applications. In reviewing these algorithms we have concentrated on the feature spaces extracted and their relevance to inter-modality applications, and on attempts to guarantee correspondences between the selected features in each image. These two problems are likewise features of all other groups of algorithms.

In analysing single-level non-linear algorithms we have concentrated on understanding the underlying physical and/or statistical models and their relationship with the resultant mathematical model on which the computational implementation is based. We have studied variants of the four most popular non-linear registration models: those of interpolating splines, statistical radial-basis function expansions, and algorithms modelling the behaviour of elastic and viscous-fluid materials. In particular, we have taken care to clarify the mathematical implementation of the TPS as an extension of linear matrix transformations, and have described many spline variants.

Following Gee and Peralta [1995], we have described the relation between defining the registration problem in terms of a regularisation problem, where the transformation optimises a cost function consisting of driving forces and smoothness constraints, and in Bayesian-statistical formulations, where the transformation maximises a posterior probability derived from likelihoods and prior assumptions. In addition we have discussed two aspects of their implementation: interpolation at both the macro- and micro-level, and representations of deformations as full deformation fields, their dual - namely coefficients to truncated basis function expansions, and combination representations by radial basis functions and finite elements.

Issues of robustness and computational efficiency have been dealt with in the chapter reviewing algorithms of the third level of sophistication, namely hierarchical algorithms. These have been classified as those varying temporally in data complexity, in warp complexity and in model complexity. In each
8.1. Summary

In case, the strategy for initially matching on a global scale, with gradual re-introduction of complexity and localisation, has been discussed in respect to its ability to avoid a common problem encountered in medical image registration, namely that of local minima traps which may cause misregistrations, unwanted deformations and premature termination of the registration process. Additional benefits of shorter computational solution times have been mentioned where relevant.

The final analytical chapter has dealt with the most flexible and demanding of registration paradigms, namely where the method of registration is to be varied spatially within the image domain. Paralleling the analysis of temporal variations, the algorithms have been classified according to their type of inhomogeneity: those where the variance is in the influence of the data, or dependence on the likelihood, those where the variance is in the deformability, or the role played by priors, and those where the assumptions of the underlying physical or conceptual model change with location. The ultimate goal in this group of algorithms is to deal with two increasingly relevant applications in research and clinical situations: cross-population studies of the varying deformability of individual structures and the modelling of deformations of anatomies consisting of different tissue types.

Concerning the latter issue, there have been published implementations of approximations to real tissue deformations for surgical simulations and for modelling growth. We foresee that such models will eventually be incorporated explicitly into knowledge-based algorithms for the inter-subject registration of medical images.

We have coined a set of terms to describe the types of restriction on deformability. First we classify regions as being either passive or actively-deforming depending on the influence of the data locally, or, equivalently, on whether the log likelihood is included in the cost function. We then refer to strongly- and weakly-deformable regions as determined by the value of an argument of the log prior. Finally we differentiate between two types of regions exhibiting complete rigidity: those which are motionless and those which are independently moving.

For each type of deformability restriction we have developed a variant on the highly non-linear viscous fluid algorithm. These form the second major contribution of this work which we describe in the next section.

8.1.2 Inhomogeneous non-linear registration and the varying-viscosity fluid model

The second major contribution of this thesis has been in the development of modifications to the viscous-fluid registration algorithm, to allow for the ease of deformability to vary spatially within the deforming source image. In particular we have employed a reformulation of the viscous-fluid PDE to allow the successful spatial variation in viscosity of the fluid model. The varying-viscosity fluid falls short of the ideal which would allow for a full range of rigidity through to fluidity. However it does to some extent allow the differentiation between strongly- and weakly-deforming regions.
For comparison, we have implemented two other modifications of the fluid, where we vary the influence of data-driven forces, and restrict the movement of selected regions by imposing boundary conditions on their velocity fields. These permit passive versus actively-deforming regions, and motionless regions respectively.

We have evaluated the success of these models using visual displays of locally-computed deformation metrics.

Results have shown that deformation of pre-defined regions can be successfully reduced by all three variants. Further work is needed in testing across a wider spectrum of viscosities and on more suitable test data.

8.1.3 Comparative evaluation of non-linear registration methods

In addition to testing the fluid variants developed, we have demonstrated a comparative analysis of non-linear methods, using both global and local registration metrics.

One of our most interesting results is figure 5.23 which illustrates graphically the local registration accuracy of landmarks to known target locations by the fluid in Gaussian and closure scale spaces compared to that by the fluid on the full resolution data only. It shows the benefit of scale space methods in avoiding misregistrations and the almost identical results of the fluid in either type of scale space.

Returning to the fluid variants, we found their comparison using artificial data sets to be illuminating, and in particular figure 7.9 (lower right) and figure 7.10 (far right) shows the complete rigidity of MF2. Figure 7.9 (top centre) shows the ability of the varying viscosity fluid (MF3) to restrict deformation while still achieving a degree of registration (although in the test shown in figure 7.22 a better registration is given by the unmodified fluid).

Our most interesting application has been that in section 7.3.4 where we used restricted deformability to highlight abnormalities in the source. Figure 7.31 (right) clearly displays the regions of abnormality in the upper half of the image.

8.2 Future Work

To conclude we suggest three topics for future work. The first is a unification of three variants to the thin-plate spline to give an automated hierarchical inhomogeneous model. The second is an evaluation of the application of the varying-viscosity fluid to cross-population studies, and the third is a further extension of the varying-viscosity fluid to allow anisotropies of deformability.

8.2.1 The automated hierarchical inhomogeneous spline

We suggest a unification of three adaptations to the TPS registration which are:

1. The automated TPS by Kim et al. [1996] which iteratively repositions initially-selected landmarks
8.2. Future Work

to minimise mutual information in the source-target pair. We suggest that first the this be extended to fully automatic landmark selection using the modified corner-finding algorithm we present in section 5.2.1, with a sparse set of source and target landmarks initially in the same location.

2. The hierarchical TPS presented by Feldmar et al. [1997] which inserts additional landmarks in areas identified as being poorly registered.

3. The combination MTQ of Little et al. [1996, 1997], which allows the incorporation of any number of independently moving rigid bodies. This would also benefit from a prior extension to the automatic registration of the identified corresponding rigid bodies.

In addition to combining these into the automated, hierarchical, inhomogeneous multiquadric spline, we suggest a comparison of the first and second methods by applying both to an image pair with identical initial landmark configurations and measuring the registration quality on termination by one of the global similarity measures.

8.2.2 Evaluation of MF3 in cross-population studies

The second suggestion for future work is an investigation of the benefits of applying the varying viscosity fluid to cross-population neuroanatomy studies. The project is conceived as consisting of two stages:

1. Estimating the cross-population variability of individual structures within the brain.

2. Supplying this information as an estimate of statistical deformability in MF3 for registration of additional data sets to the mean of the population used in part 1.

We hypothesise that the registration will be faster or more reliable using MF3 with these statistical estimates but a thorough evaluation would be required to support the hypothesis.

8.2.3 Anisotropic fluid

In addition to physical tissues varying in strength of deformability, many also exhibit a preferred direction of deformation. A classic example is that of muscles which may stretch locally in one given direction preferentially. Whereas the algorithms modelling inhomogeneous deformations are relatively few, the set of anisotropic models is almost non-existent. While the inhomogeneous fluid, even in its inhomogeneous version, is a poor approximation to the behaviour of real tissue deformations, it would nevertheless be interesting to add to the number of anisotropic registration models by constructing a fluid variant whose viscosity is allowed to vary with orientation as well as location. As with the inhomogeneous model, this would initially require a re-working of the PDE to accommodate such variation.

Finally, the varying-viscosity fluid is interesting as a theoretical experiment and as a topic of future research we suggest comparisons and parallels to the generation of inhomogeneous scale spaces. Additionally there may be possible non-medical applications in image morphing for computer graphics. Another
interesting exercise would be to research viscosity varying due to intrinsic features of the image pair and see how this compares to curve evolution.
<table>
<thead>
<tr>
<th>LIKELIHOOD</th>
<th>ALGORITHM</th>
<th>PRIORS</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIMILARITY</td>
<td>FEATURE SPACES</td>
<td>ALLOWABLE FUNCTIONS</td>
</tr>
<tr>
<td>METRICS</td>
<td>principal axes</td>
<td>linear</td>
</tr>
<tr>
<td></td>
<td>geometric features</td>
<td>piecewise linear</td>
</tr>
<tr>
<td></td>
<td>landmarks</td>
<td>landmark-based interpolation</td>
</tr>
<tr>
<td></td>
<td>curves</td>
<td>finite element interpolation</td>
</tr>
<tr>
<td></td>
<td>surfaces</td>
<td></td>
</tr>
<tr>
<td></td>
<td>intensity features</td>
<td></td>
</tr>
<tr>
<td></td>
<td>pixel intensities</td>
<td></td>
</tr>
<tr>
<td></td>
<td>gradients</td>
<td></td>
</tr>
<tr>
<td>SSD</td>
<td>single or iterated LINEAR transformations</td>
<td></td>
</tr>
<tr>
<td>correlation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MI</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**HIERARCHICAL** algorithms
- temporally increasing complexity of:

<table>
<thead>
<tr>
<th>DATA WAVR</th>
<th>MODEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>DATA INFLUENCE</td>
<td>DEFORMABILITY</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix A

Appendices of Theory

A.1 Distance transforms

A distance transform is a convolution mask for rapidly computing the approximate Euclidean distance of pixels from a set of objects in the image. The image is first binarised, with pixels in the object set assigned the value 0 and the remaining pixels assigned an arbitrary high value. The mask is then passed over the image. The mask is a template covering only a small region of the image at one time, identifying a ‘central’ pixel and a neighbourhood. The distance transform mask adds a distance value \( c(k, l) \) to each pixel it covers. The ‘central’ image pixel value is then reset to the minimum pixel value within the region.

\[
v_{i,j}^m = \min_{(k,l) \in \text{mask}} \left( v_{i+k,j+l}^{m-1} + c(k, l) \right)
\]

Borgefors [1986] assesses several distance transforms (DT). Her conclusion is that the best integer approximation to the Euclidean distance is given by a chamfer DT. Increasing the neighbourhood size of the mask increases the accuracy of the approximation but takes longer to compute. There are many variants of the chamfer mask, with increasing values. She recommends the chamfer 3–4 or the chamfer 5–7–11, fig. A.1. Distance transforms may be applied using either sequential or parallel propagation. In parallel propagation the whole mask is passed many times over the image until the distance image is no longer altered. In sequential propagation the mask is split into two, the forwards mask and the backwards mask, both containing the ‘central’ pixel. The middle of figure A.1 shows these for the chamfer 3–4. Each is passed over the image only once, the forwards mask from top left to bottom right and the backwards mask in the reverse direction.

Danielsson [1980] presents two sequential algorithms for computing the true Euclidean distance, the 4SED and the 8SED. The algorithms compute two separate distance maps, one \( L_x(i, j) \) of the distance of each pixel in the \( x \)-direction from the object set, and the other \( L_y(i, j) \) of the \( y \)-direction distances. Together these give a vector array \( \vec{L}(i, j) \) whose size for each pixel is the Euclidean distance \( \| \vec{L}(i, j) \| = \sqrt{L_x^2 + L_y^2} \). The 4SED and the 8SED algorithms are given in full on pages 230 and 234 of Danielsson [1980]. He shows that they are error free other than for a few sparsely scattered points, with absolute
A.2 Derivation of the elasticity prior from the physical model

This section follows Christensen [1994] in deriving the elastic PDE (3.66). For further details the reader is referred to any good text on elasticity theory or continuous media, such as Timoshenko and Goodier [1970].

An elastic medium has a resistance to deformation supplied by cohesive forces acting on and between its constitutive volume elements. These forces ensure the continuity of the medium, ensuring it does not tear or fold under the action of reasonable external forces.

Strain is defined as the deformation of a volume element of the elastic material. There are two types of strain: extensional and shear. The former are stretching or compression of the material caused by tensile forces applied normal to a surface of the volume element. Shear strain are deformations within the (limiting) plane of the surface of the element at the point of application of the forces.

As mentioned in section 3.7.1, Christensen [1994] applied the theory to the backwards transformation, finding the mapping from the final transformed source image $S'$ (modelled as consisting of deformed volume elements) to the original source image $S$ (modelled as consisting of cuboidal volume elements).

Consider a point $P$ in the original source image and its corresponding point $p$ in the final transformed image, figure (A.2).

Any line element $d\vec{r}$ emanating from $P$ will have been mapped (in the forwards transformation) from an (undeformed) line element $d\vec{R}$ originating at $P$. Extensional strain describes the difference in length...
A.2. Derivation of the elasticity prior from the physical model

Figure A.2: Point and volume element in deformed and original target image

between $d\vec{R}$ and $d\vec{r}$, figure (A.3). In addition the angle between two non-coincident line elements $d\vec{r}_1$ and $d\vec{r}_2$ at $p$ may have changed from the angle between the corresponding line elements $d\vec{R}_1$ and $d\vec{R}_2$ at $P$; this difference is a measure of the shear strain.

Coordinate system

Points $\{\vec{X}\}$ in the original target image are mapped to $\{\vec{x}\}$ in the final transformed source image $S'$. The backwards mapping gives the coordinates of $S'$ in terms of those in $S$ and hence we have

$$\vec{x} = \tilde{u} (\vec{X})$$

Both $\{\vec{x}\}$ and $\{\vec{X}\}$ lie in the Cartesian coordinate system $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$. Point $p$ is at location $\vec{r}$ where

$$\vec{r} = \sum_{i=1}^{3} x_i \vec{e}_i$$

and $P$ is at

$$\vec{R} = \sum_{i=1}^{3} X_i \vec{e}_i$$
Thus for the infinitesimal line elements $d\vec{r}$ and $d\vec{R}$ and by the chain rule for differentiation,

$$
\begin{align*}
\sum_{i=1}^{3} dX_i \frac{\partial \vec{R}}{\partial X_i} &= \sum_{i=1}^{3} dX_i \vec{e}_i \\
\sum_{i=1}^{3} dX_i \frac{\partial \vec{r}}{\partial X_i} &= \sum_{i=1}^{3} dX_i \vec{e}_i
\end{align*}
$$

Also, the mapping between $S'$ and $S$ will be given by a displacement field $\vec{u}$ specifying a displacement from every pixel $\vec{x}$ in the original image, giving

$$\vec{r}'(\vec{x}) = \vec{R}(\vec{x}) - \vec{u}(\vec{x}) \quad \text{(A.1)}$$

Figure A.4: Displacements $\vec{u}$ and decompositions of line elements $d\vec{R}$ and $d\vec{r}$ (following Christensen)

**Eulerian strain tensor**

The strain tensor $\bar{\eta}(\vec{x}) = \{\eta_{nm}\}$ is a symmetric matrix describing the strains at a point $\vec{x}$ in the elastic body. For example, referring to figure A.4, $\eta_{3,3}$ gives the extensional strains along the direction $\frac{\partial \vec{r}}{\partial x_3}$ and the off-diagonal term $\eta_{1,2}$ gives the shear strains within the $(\frac{\partial \vec{r}}{\partial x_1}, \frac{\partial \vec{r}}{\partial x_2})$ plane.

The strain tensor is defined by

$$
\sum_{i=1}^{3} 2\eta_{nm} dX_m dX_n = (d\vec{R}.d\vec{R}) - (d\vec{r}.d\vec{r})
$$

$$
= \left( \sum_{i=1}^{3} dX_i \frac{\partial \vec{R}}{\partial X_i} \right) \cdot \left( \sum_{i=1}^{3} dX_i \frac{\partial \vec{R}}{\partial X_i} \right) - \left( \sum_{i=1}^{3} dX_i \frac{\partial \vec{r}}{\partial X_i} \right) \cdot \left( \sum_{i=1}^{3} dX_i \frac{\partial \vec{r}}{\partial X_i} \right)
$$

giving (Christensen [1994])

$$
\eta_{nm} = \frac{1}{2} \left( \frac{\partial \vec{R}}{\partial X_m} \cdot \frac{\partial \vec{R}}{\partial X_n} - \frac{\partial \vec{r}}{\partial X_m} \cdot \frac{\partial \vec{r}}{\partial X_n} \right)
$$

$$
= \frac{1}{2} \left( \delta_{nm} - \frac{\partial \vec{r}}{\partial X_m} \cdot \frac{\partial \vec{r}}{\partial X_n} \right)
$$
A.2. Derivation of the elasticity prior from the physical model 217

where the latter term does not equal \( \delta_{nm} \) since the volume element at \( \mathcal{F} \) is not orthogonal.

Then substituting from equation A.1,

\[
\eta_{nm} = \frac{1}{2} \left( \frac{\partial \bar{R}}{\partial X_m} \frac{\partial \bar{R}}{\partial X_n} - \frac{\partial}{\partial X_m} (\bar{R} - \bar{\bar{u}}) \frac{\partial}{\partial X_n} (\bar{R} - \bar{\bar{u}}) \right)
\]

\[
= \frac{1}{2} \left( \frac{\partial \bar{R}}{\partial X_m} \frac{\partial \bar{R}}{\partial X_n} - \left( \frac{\partial \bar{R}}{\partial X_m} - \frac{\partial \bar{\bar{u}}}{\partial X_m} \right) \left( \frac{\partial \bar{R}}{\partial X_n} - \frac{\partial \bar{\bar{u}}}{\partial X_n} \right) \right)
\]

\[
= \frac{1}{2} \left( \frac{\partial \bar{R}}{\partial X_m} \frac{\partial \bar{R}}{\partial X_n} - \frac{\partial \bar{\bar{R}}}{\partial X_m} \frac{\partial \bar{\bar{R}}}{\partial X_n} + \frac{\partial \bar{\bar{u}}}{\partial X_m} \frac{\partial \bar{\bar{u}}}{\partial X_n} \right)
\]

\[
= \frac{1}{2} \left( \frac{\partial u_n}{\partial X_m} + \frac{\partial u_m}{\partial X_n} - \sum_{r=1}^{3} \frac{\partial u_r}{\partial X_m} \frac{\partial u_r}{\partial X_n} \right)
\]

since

\[
\bar{R} = \sum_{i=1}^{3} X_i \bar{e}_i \rightarrow \frac{\partial \bar{R}}{\partial X_n} = \bar{e}_n \rightarrow \frac{\partial \bar{\bar{u}}}{\partial X_m} \bar{e}_n = \frac{\partial u_n}{\partial X_m}
\]

For small deformations, the products \( \frac{\partial u_r}{\partial X_m} \frac{\partial u_r}{\partial X_n} \) are negligible, so the strain tensor is reduced to a simplified version

\[
\epsilon_{nm} = \frac{1}{2} \left( \frac{\partial u_n}{\partial X_m} + \frac{\partial u_m}{\partial X_n} \right)
\]  

(A.2)

![Figure A.5: Internal forces \( \delta \mathcal{F} \) acting on elemental area \( \delta A \)](image)

Stress \( \bar{\sigma} \) is the cohesive force per unit area, and is defined by

\[
\bar{\sigma} = \lim_{\delta A \to 0} \left( \frac{\delta \mathcal{F}}{\delta A} \right) = \frac{d \mathcal{F}}{d A}
\]

The force balancing equation is now given by

\[
\frac{\partial \sigma_{1i}}{\partial X_1} + \frac{\partial \sigma_{2i}}{\partial X_2} + \frac{\partial \sigma_{3i}}{\partial X_3} + b_i = 0 \quad i = 1, 2, 3 \]  

(A.3)

i.e.

\[
\nabla \cdot \bar{\sigma} + \bar{b} = 0
\]  

(A.4)

For an isotropic material, elasticity theory gives the stress-strain relationship as

\[
\sigma_{nm} = 2\mu \epsilon_{nm} + \lambda \delta_{nm} \sum_{r=1}^{3} \epsilon_{rr}
\]  

(A.5)
where \( \lambda \) and \( \mu \) are Lamé’s elasticity constants. Substituting (A.5) and (A.2) into (A.3) gives the linear elastic PDE (3.66).

### A.3 Eigenfunctions of the elasticity operator

Christensen [1994] gives the eigenvectors of (3.66) for both the 2D and 3D cases.

In the 2D case, given *sliding* boundary conditions

\[
\begin{align*}
\frac{\partial u_1}{\partial X_1} \bigg|_{(0,X_2)} &= \frac{\partial u_1}{\partial X_2} \bigg|_{(X_1,0)} = \frac{\partial u_2}{\partial X_1} \bigg|_{(0,X_2)} = \frac{\partial u_2}{\partial X_2} \bigg|_{(1,X_3)} = 0
\end{align*}
\]

the eigenvectors are

\[
\begin{align*}
\tilde{e}^a_{nm} &= \frac{2}{\sqrt{m^2 + n^2}} \left( m \sin m \pi X_1 \cos n \pi X_2, n \cos m \pi X_1 \sin n \pi X_2 \right) \\
\tilde{e}^e_{nm} &= \frac{2}{\sqrt{m^2 + n^2}} \left( -n \sin m \pi X_1 \cos n \pi X_2, m \cos m \pi X_1 \sin n \pi X_2 \right)
\end{align*}
\]

with eigenvalues

\[
\begin{align*}
\lambda^1_{nm} &= -\pi^2 (2\mu + \lambda) (m^2 + n^2) \\
\lambda^2_{nm} &= -\pi^2 \mu (m^2 + n^2)
\end{align*}
\] (A.6)

For *bending* boundary conditions

\[
\begin{align*}
\frac{\partial u_1}{\partial X_1} \bigg|_{(X_1,0)} &= \frac{\partial u_1}{\partial X_2} \bigg|_{(X_1,1)} = \frac{\partial u_2}{\partial X_1} \bigg|_{(0,X_2)} = \frac{\partial u_2}{\partial X_2} \bigg|_{(1,X_3)} = 0
\end{align*}
\]

the eigenvectors are

\[
\begin{align*}
\tilde{e}^a_{nm} &= \frac{2}{\sqrt{m^2 + n^2}} \left( m \cos m \pi X_1 \sin n \pi X_2, n \sin m \pi X_1 \cos n \pi X_2 \right) \\
\tilde{e}^e_{nm} &= \frac{2}{\sqrt{m^2 + n^2}} \left( -n \cos m \pi X_1 \sin n \pi X_2, m \sin m \pi X_1 \cos n \pi X_2 \right)
\end{align*}
\]

with eigenvalues the same as A.6.

In the 3D case, the bending boundary conditions are

\[
\begin{align*}
u_1(0, X_2, X_3) &= u_1(1, X_2, X_3) = u_2(X_1, 0, X_3) = u_2(X_1, 1, X_3) = u_3(X_1, X_2, 0) = u_3(X_1, X_2, 1) = 0
\end{align*}
\]

and

\[
\begin{align*}
\frac{\partial u_1}{\partial X_3} \bigg|_{(X_1,0,X_3)} &= \frac{\partial u_1}{\partial X_2} \bigg|_{(X_1,1,X_3)} = \frac{\partial u_1}{\partial X_3} \bigg|_{(X_1,X_2,0)} = \frac{\partial u_1}{\partial X_3} \bigg|_{(X_1,X_2,1)} = 0 \\
\frac{\partial u_2}{\partial X_3} \bigg|_{(0,X_2,X_3)} &= \frac{\partial u_2}{\partial X_1} \bigg|_{(1,X_2,X_3)} = \frac{\partial u_2}{\partial X_3} \bigg|_{(X_1,X_2,0)} = \frac{\partial u_2}{\partial X_3} \bigg|_{(X_1,X_2,1)} = 0 \\
\frac{\partial u_3}{\partial X_3} \bigg|_{(0,X_2,X_3)} &= \frac{\partial u_3}{\partial X_2} \bigg|_{(1,X_2,X_3)} = \frac{\partial u_3}{\partial X_3} \bigg|_{(X_1,X_2,0)} = \frac{\partial u_3}{\partial X_3} \bigg|_{(X_1,1,X_3)} = 0
\end{align*}
\]
A.4 Numerical solutions of the viscous fluid PDE

Giving eigenvectors

\[ \bar{e}_{klm}^1 = \alpha_1 (k \sin k \pi X_1 \cos l \pi X_2 \cos m \pi X_3, l \cos k \pi X_1 \sin l \pi X_2 \cos m \pi X_3, m \cos k \pi X_1 \cos l \pi X_2 \sin m \pi X_3) \]

\[ \bar{e}_{klm}^2 = \alpha_2 (-l \sin k \pi X_1 \cos l \pi X_2 \cos m \pi X_3, k \cos k \pi X_1 \sin l \pi X_2 \cos m \pi X_3, 0) \]

\[ \bar{e}_{klm}^3 = \alpha_3 (km \sin k \pi X_1 \sin l \pi X_2 \cos m \pi X_3, lm \sin k \pi X_1 \cos l \pi X_2 \sin m \pi X_3, -(k^2 + l^2) \cos k \pi X_1 \cos l \pi X_2 \sin m \pi X_3) \]

where

\[ \alpha_1 = \sqrt{\frac{8}{k^2 + l^2 + m^2}} \]
\[ \alpha_2 = \sqrt{\frac{8}{k^2 + l^2}} \]
\[ \alpha_3 = \sqrt{\frac{8}{(k^2 + l^2)(k^2 + l^2 + m^2)}} \]

with eigenvalues

\[ \lambda_{klm}^1 = -\pi^2 (2 \mu + \lambda)(k^2 + l^2 + m^2) \]
\[ \lambda_{klm}^2 = \lambda_{klm}^3 = -\pi^2 \mu (k^2 + l^2 + m^2) \]

For the bending boundary conditions

\[ u_1(X_1, 0, X_3) = u_1(X_1, 1, X_3) = u_1(X_1, X_2, 0) = u_1(X_1, X_2, 1) = 0 \]
\[ u_2(0, X_2, X_3) = u_2(1, X_2, X_3) = u_2(X_1, X_2, 0) = u_2(X_1, X_2, 1) = 0 \]
\[ u_3(0, X_2, X_3) = u_3(1, X_2, X_3) = u_3(X_1, 0, X_3) = u_3(X_1, 1, X_3) = 0 \]

\[ \frac{\partial u_1}{\partial X_1}(0, X_2, X_3) = \frac{\partial u_1}{\partial X_1}(1, X_2, X_3) = \frac{\partial u_2}{\partial X_2}(X_1, 0, X_3) = \frac{\partial u_2}{\partial X_2}(X_1, 1, X_3) = \frac{\partial u_3}{\partial X_3}(X_1, X_2, 0) = \frac{\partial u_3}{\partial X_3}(X_1, X_2, 1) = 0 \]

the eigenvectors are

\[ \bar{e}_{klm}^1 = \alpha_1 (k \cos k \pi X_1 \sin l \pi X_2 \sin m \pi X_3, l \sin k \pi X_1 \cos l \pi X_2 \sin m \pi X_3, m \sin k \pi X_1 \sin l \pi X_2 \cos m \pi X_3) \]
\[ \bar{e}_{klm}^2 = \alpha_2 (-l \cos k \pi X_1 \sin l \pi X_2 \sin m \pi X_3, k \sin k \pi X_1 \cos l \pi X_2 \sin m \pi X_3, 0) \]
\[ \bar{e}_{klm}^3 = \alpha_3 (km \cos k \pi X_1 \sin l \pi X_2 \sin m \pi X_3, lm \sin k \pi X_1 \cos l \pi X_2 \sin m \pi X_3, -(k^2 + l^2) \sin k \pi X_1 \sin l \pi X_2 \cos m \pi X_3) \]

with eigenvalues as for the sliding boundary conditions.

A.4 Numerical solutions of the viscous fluid PDE

Finite differencing within SOR

Solution of the viscous fluid PDE (3.67) or of the elasticity PDE (3.66) in the spatial domain for given forces \( f \) using finite difference approximations to the spatial derivatives can be found using the iterative successive over-relaxation method (SOR) with checkerboard update [Press et al., 1995; Strikwerda,
1989]. Successive approximations to the solutions \( v_{ij}^{(1)}, v_{ij}^{(2)} \) of internal (non-boundary) pixels are computed using the schemes in equations A.8 and A.11, where all velocities \( v_{ij}^{(1)}, v_{ij}^{(2)} \) with even-sum indices are updated first, followed by those with indices of odd-sum.

For constant viscosity (see also Christensen [1994]) the residual \( r_{ij} \) is computed by equation (A.7) and then used to update the velocities using the scheme of (A.8). The current pixel velocity is updated using a checkerboard (or red-black) update [Press et al., 1995; Strikwerda, 1989]. That is, velocities are updated at alternate grid points in a given iteration, with those at the complement updated on the following iteration, so the velocity at a given grid point is updated at every other iteration. The residual is also used to compute a norm for the termination, given in (A.9).

\[
\begin{align*}
\mathbf{r}_{ij}^{(1)} &= (f_{ij}^{(1)} + (\lambda + 2\mu)(v_{i+1,j}^{(1)} + v_{i-1,j}^{(1)}) + \mu(v_{i,j+1}^{(1)} + v_{i,j-1}^{(1)}) \\
&+ \frac{\lambda + \mu}{4} (v_{i+1,j+1}^{(2)} + v_{i-1,j+1}^{(2)} - v_{i+1,j-1}^{(2)} - v_{i-1,j-1}^{(2)})/(6\mu + 2\lambda) - v_{ij}^{(1)}) \quad (A.7) \\
v_{ij}^{(1)*+1} &= v_{ij}^{(1)*-1} + \omega r_{ij}^{(1)} \quad (A.8) \\
a_{\text{norm}}^{(1)} &= \sum_{ij} \left| r_{ij}^{(1)} \right| \quad (A.9)
\end{align*}
\]

\[
\begin{align*}
\mathbf{r}_{ij}^{(2)} &= (f_{ij}^{(2)} + (\lambda + 2\mu)(v_{i,j-1}^{(2)} + v_{i,j+1}^{(2)}) + \mu(v_{i+1,j}^{(2)} + v_{i-1,j}^{(2)}) \\
&+ \frac{\lambda + \mu}{4} (v_{i+1,j+1}^{(1)} + v_{i-1,j+1}^{(1)} - v_{i+1,j-1}^{(1)} - v_{i-1,j-1}^{(1)})/(6\mu + 2\lambda) - v_{ij}^{(2)}) \quad (A.10) \\
v_{ij}^{(2)*+1} &= v_{ij}^{(2)*-1} + \omega r_{ij}^{(2)} \quad (A.11) \\
a_{\text{norm}}^{(2)} &= \sum_{ij} \left| r_{ij}^{(2)} \right| \quad (A.12)
\end{align*}
\]

Termination is determined by the norms \( a_{\text{norm}}^{(1)} \) and \( a_{\text{norm}}^{(2)} \) falling below specified thresholds \( t_1 \) and \( t_2 \), computed from the driving forces:

\[
\begin{align*}
t_1 &= c \sum_{ij} \left| f_{ij}^{(1)} \right| \\
t_2 &= c \sum_{ij} \left| f_{ij}^{(2)} \right|
\end{align*}
\]

with \( c = 10.0^{-5} \).

\( \omega \) is the relaxation parameter, which for overrelaxation takes a value in the range (1,2). Chebyshev acceleration [Press et al., 1995] updates \( \omega \) on each iteration using the spectral radius of the Jacobi iteration matrix for the finite difference solution; we show the Jacobi iteration matrix for our problem in figure A.6. The iteration matrix in our case consists of four blocks. The upper two blocks correspond to the computation of the velocities \( v_{ij}^{(1)} \) in the \( x \)-direction, the lower two blocks correspond to computation of
A.4. Numerical solutions of the viscous fluid PDE

The velocities $v_{ij}^{(2)}$ in the $y$-direction:

\[
\begin{pmatrix}
0 & 0 \\
0 & 0 \\
\vdots & \vdots \\
0 & 0
\end{pmatrix}
\begin{pmatrix}
0 \\
0 \\
\vdots \\
0
\end{pmatrix}
= \begin{pmatrix}
0 \\
0 \\
\vdots \\
0
\end{pmatrix}
\]

The breaks in the diagonals in the matrix are due to the boundary conditions of fixed zero velocities.

Figure A.6: Digital representation of the Jacobi iteration matrix for finite-difference solution of the elasticity PDE

We did not find that Chebyshev acceleration significantly improved on the speed of convergence within
the upper limit of 4000 SOR iterations we set per timestep since initial iterations converge more slowly
due to a lower value of \( \omega \) used. We abandoned its application to our finite difference solution and solved
it using a fixed value of \( \omega = 1.9 \).

A.5 Finite difference scheme for the numerical solution of the
varying-viscosity fluid

In the heat diffusion case we apply finite differences on \( \kappa \phi' \), writing

\[
(\kappa \phi')' = \kappa_{i+\frac{1}{2}} \nabla \phi|_{i+\frac{1}{2}} - \kappa_{i-\frac{1}{2}} \nabla \phi|_{i-\frac{1}{2}}
\]

(A.14)

and then substitute

\[
\kappa_{i+\frac{1}{2}} \nabla \phi|_{i+\frac{1}{2}} = \left( \frac{\kappa_{i+1} + \kappa_i}{2} \right) \left( \frac{\phi_{i+1} - \phi_i}{h} \right)
\]

and

\[
\kappa_{i-\frac{1}{2}} \nabla \phi|_{i-\frac{1}{2}} = \left( \frac{\kappa_i + \kappa_{i-1}}{2} \right) \left( \frac{\phi_i - \phi_{i-1}}{h} \right)
\]

Here we use averaging to obtain

\[
\kappa_{i+\frac{1}{2}} = \frac{\kappa_i + \kappa_{i-1}}{2}
\]

and finite differences to approximate

\[
\nabla \phi|_{i-\frac{1}{2}} = \left( \frac{\phi_i - \phi_{i-1}}{h} \right)
\]

\( h \) is the spacing between the nodes or points used in the approximation; in this particular case, \( h = 1 \).

Altogether, equation (A.14) when expanded out gives the finite difference scheme

\[
\begin{pmatrix}
\vdots & \vdots & \vdots \\
\frac{\kappa_{i-1} + \kappa_i}{2h} & -\frac{\kappa_{i-1} + 2\kappa_i + \kappa_{i+1}}{2h} & \frac{\kappa_i + \kappa_{i+1}}{2h} \\
\vdots & \vdots & \vdots
\end{pmatrix}
\begin{pmatrix}
\phi_{i-1} \\
\phi_i \\
\phi_{i+1} \\
\vdots
\end{pmatrix}
\]

(A.15)

whereas the first version would have given

\[
\begin{pmatrix}
\vdots & \vdots & \vdots \\
\frac{\kappa_{i-1} - \kappa_{i+1} + 2\kappa_i}{4h} & -\frac{\kappa_{i+1} + 2\kappa_i}{4h} & \frac{\kappa_{i+1} - \kappa_{i-1} + 2\kappa_i}{4h} \\
\vdots & \vdots & \vdots
\end{pmatrix}
\begin{pmatrix}
\phi_{i-1} \\
\phi_i \\
\phi_{i+1} \\
\vdots
\end{pmatrix}
\]

(A.16)
A.5. Finite difference scheme for the numerical solution of the varying-viscosity fluid

Now we will apply the same idea to the fluid PDE.

Instead of expanding out (6.8) to (3.67), we will write it (in Einstein summation) as

\[
\frac{\partial \sigma_{11}}{\partial x_1} \frac{\partial \sigma_{11}}{\partial x_2} = \left( \frac{\partial \sigma_{11}}{\partial x_1} \right)_{i-j,k}^{(6.8)} + \frac{\partial \sigma_{11}}{\partial x_2} \frac{\partial \sigma_{11}}{\partial x_2} \left( \frac{\partial \sigma_{11}}{\partial x_2} \right)_{i-k,j}^{(6.8)}
\]

where \( \bar{V} \) is the velocity field we are solving.

In other words, this gives us for every pixel location \((i, j)\), the vector

\[
\left( \frac{\partial \sigma_{11}}{\partial x_1} + \frac{\partial \sigma_{11}}{\partial x_2} \right)_{i-j,k}^{(6.8)} = \left( \frac{\partial \sigma_{11}}{\partial x_1} \right)_{i-j,k}^{(6.8)} + \frac{\partial \sigma_{11}}{\partial x_2} \frac{\partial \sigma_{11}}{\partial x_2} \left( \frac{\partial \sigma_{11}}{\partial x_2} \right)_{i-j,k}^{(6.8)}
\]

Applying finite differencing and half-pixel averaging we have:

\[
\frac{\partial}{\partial x_1} \frac{\partial \sigma_{11}}{\partial x_2} \bigg|_{i-j,k}^{(6.8)} = \left( \frac{\partial \sigma_{11}}{\partial x_2} \right)_{i+\frac{1}{2},j,k}^{(6.8)} - \left( \frac{\partial \sigma_{11}}{\partial x_2} \right)_{i-\frac{1}{2},j,k}^{(6.8)}
\]

where lines A.18 and A.20 employ finite differencing and line A.21 uses nearest-pixel averaging to evaluate \( \mu \) and \( \sigma_{11}^{(6.8)} \) at \( \frac{1}{2} \)-pixel locations.

Similarly we have

\[
\frac{\partial}{\partial x_1} \frac{\partial \sigma_{11}}{\partial x_1} \bigg|_{i-j,k}^{(6.8)} = \left( \frac{\partial \sigma_{11}}{\partial x_1} \right)_{i+\frac{1}{2},j,k}^{(6.8)} - \left( \frac{\partial \sigma_{11}}{\partial x_1} \right)_{i-\frac{1}{2},j,k}^{(6.8)}
\]
A.5. Finite difference scheme for the numerical solution of the varying-viscosity fluid

\[
\left(\frac{\mu_{i+1,j,k} + \mu_{i,j,k}}{2}\right) \left[\frac{v_{i+1,j,k}^{(2)} - v_{i,j,k}^{(2)}}{2}\right] - \left(\frac{\mu_{i,j,k} + \mu_{i-1,j,k}}{2}\right) \left[\frac{v_{i,j,k}^{(2)} - v_{i-1,j,k}^{(2)}}{2}\right] = 0
\]

Hence the finite-difference expansion of the \( x \)-direction part of equation 3.67 is given by:

\[
0 = \left[\left(\frac{\mu_{i+1,j,k} + \mu_{i,j,k}}{2}\right) \left[\frac{v_{i,j,k+1}^{(1)} - v_{i,j,k}^{(1)}}{2}\right] - \left(\frac{\mu_{i,j,k} + \mu_{i,j,k-1}}{2}\right) \left[\frac{v_{i,j,k}^{(1)} - v_{i,j,k-1}^{(1)}}{2}\right]\right] + \left(\frac{\mu_{i+1,j,k} + \mu_{i,j,k}}{2}\right) \left[\frac{v_{i,j,k+1}^{(3)} - v_{i,j,k}^{(3)}}{4}\right] - \left(\frac{\mu_{i+1,j,k} + \mu_{i,j,k}}{2}\right) \left[\frac{v_{i,j,k}^{(3)} - v_{i,j,k-1}^{(3)}}{4}\right] + \frac{f_{i,j,k}}{4}
\]

Isolating the central pixel \( v_{i,j,k}^{(1)} \):

\[
v_{i,j,k}^{(1)} = \left[8\mu_{i,j,k} + 2 \left(\mu_{i+1,j,k} + \mu_{i-1,j,k}\right) + \mu_{i,j,k+1} + \mu_{i,j,k-1} + \mu_{i,j,k+1} + \mu_{i,j,k-1}\right]
\]

\[
+ 2 \left[\left(\frac{\mu_{i+1,j,k} + \mu_{i,j,k}}{2}\right) v_{i+1,j,k}^{(1)} + \left(\frac{\mu_{i,j,k} + \mu_{i-1,j,k}}{2}\right) v_{i-1,j,k}^{(1)}\right] + \left(\frac{\mu_{i+1,j,k} + \mu_{i,j,k}}{2}\right) v_{i,j,k+1}^{(1)} + \left(\frac{\mu_{i,j,k} + \mu_{i,j,k-1}}{2}\right) v_{i,j,k-1}^{(1)} + \frac{f_{i,j,k}}{4}
\]

(A.25)
Solving using SOR (successive over-relaxation), we compute the residual

\[
\text{residual} = \left[ 2 \left( \mu_{i,j+1,k} + \mu_{i,j,k} \right) v_{i,j+1,k}^{(1)} v_{i+1,j+1,k}^{(1)} + \left( \mu_{i,j+1,k} + \mu_{i-1,j,k} \right) v_{i-1,j,k}^{(1)} v_{i,j,k}^{(1)} \right] + \\
\left( \mu_{i,j+1,k} + \mu_{i,j,k} \right) \left[ v_{i,j+1,k}^{(1)} + \frac{1}{4} \left( v_{i,j+1,k}^{(2)} + v_{i,j+1,k}^{(2)} - v_{i-1,j,k}^{(2)} - v_{i+1,j,k}^{(2)} \right) \right] + \\
\left( \mu_{i,j+1,k} + \mu_{i,j,k} \right) \left[ v_{i,j+1,k}^{(1)} + \frac{1}{4} \left( v_{i-1,j,k}^{(2)} + v_{i+1,j,k}^{(2)} - v_{i,j+1,k}^{(2)} - v_{i,j,k}^{(2)} \right) \right] + \\
\left( \mu_{i,j+1,k} + \mu_{i,j,k} \right) \left[ v_{i,j+1,k}^{(1)} + \frac{1}{4} \left( v_{i,j+1,k}^{(3)} + v_{i,j+1,k}^{(3)} - v_{i+1,j,k+1}^{(3)} - v_{i+1,j,k-1}^{(3)} \right) \right] + \\
\left( \mu_{i,j+1,k} + \mu_{i,j,k} \right) \left[ v_{i,j+1,k}^{(1)} + \frac{1}{4} \left( v_{i,j+1,k}^{(3)} + v_{i,j+1,k}^{(3)} - v_{i+1,j,k-1}^{(3)} - v_{i+1,j,k-1}^{(3)} \right) \right] + \\
2 (\mu_{i,j+1,k} + \mu_{i,1,j,k}) + (\mu_{i,j+1,k} + \mu_{i,j,k}) + \\
\mu_{i,j,k+1} + \mu_{i,j,k} + \mu_{i,j,k} + \mu_{i,j,k} - 1 \right] \\
\mu_{i,j,k+1} + \mu_{i,j,k} + \mu_{i,j,k} + \mu_{i,j,k} - 1 \right] + \\
2 f_{i,j,k}^1 / [8 \mu_{i,j,k} + 2 (\mu_{i,1,j,k} + \mu_{i,j,k}) + \mu_{i,j,k} + \\
\mu_{i,j,k} + \mu_{i,j,k} + \mu_{i,j,k} + \mu_{i,j,k} - 1 \right] \right] \\
(A.26)
\]

and then the current pixel velocity is updated using a checkerboard (or red-black) update [Press et al., 1995; Strikwerda, 1989] as in the constant viscosity case. Hence for the iteration at time \( t + 1 \), the \( x \)-component of the velocity is updated by

\[
v_{i,j,k}^{(1),t+1} = v_{i,j,k}^{(1),t} \left( 1 - \omega \right) + \omega \cdot \text{residual} \quad (A.27)
\]

For the \( y \)-direction we have

\[
0 = \left( \frac{1}{2} \mu_{i,j+1,k} + \mu_{i,j,k} \right) \left[ \frac{v_{i,j+1,k}^{(2)}}{4} v_{i+1,j+1,k}^{(2)} - \frac{v_{i,j+1,k}^{(2)}}{4} v_{i,j,k}^{(2)} \right] - \left( \frac{1}{2} \mu_{i,j+1,k} + \mu_{i-1,j,k} \right) \left[ \frac{v_{i,j+1,k}^{(2)}}{4} v_{i,j,k}^{(2)} - \frac{v_{i,j,k}^{(2)}}{4} v_{i,j+1,k}^{(2)} \right] + \\
\left( \frac{1}{2} \mu_{i,j+1,k} + \mu_{i,j,k} \right) \left[ \frac{v_{i,j+1,k}^{(2)}}{4} v_{i,j+1,k}^{(2)} - \frac{v_{i,j+1,k}^{(2)}}{4} v_{i,j,k}^{(2)} \right] - \left( \frac{1}{2} \mu_{i,j+1,k} + \mu_{i,j,k} \right) \left[ \frac{v_{i,j+1,k}^{(2)}}{4} v_{i,j,k}^{(2)} - \frac{v_{i,j,k}^{(2)}}{4} v_{i,j+1,k}^{(2)} \right] + \\
\left( \frac{1}{2} \mu_{i,j+1,k} + \mu_{i,j,k} \right) \left[ \frac{v_{i,j+1,k}^{(2)}}{4} v_{i+1,j+1,k}^{(2)} - \frac{v_{i,j+1,k}^{(2)}}{4} v_{i,j,k}^{(2)} \right] - \left( \frac{1}{2} \mu_{i,j+1,k} + \mu_{i,j,k} \right) \left[ \frac{v_{i,j+1,k}^{(2)}}{4} v_{i,j,k}^{(2)} - \frac{v_{i,j,k}^{(2)}}{4} v_{i,j+1,k}^{(2)} \right] + \\
\left( \frac{1}{2} \mu_{i,j+1,k} + \mu_{i,j,k} \right) \left[ \frac{v_{i,j+1,k}^{(2)}}{4} v_{i,j+1,k}^{(2)} - \frac{v_{i,j+1,k}^{(2)}}{4} v_{i+1,j+1,k}^{(2)} \right] - \left( \frac{1}{2} \mu_{i,j+1,k} + \mu_{i,j,k} \right) \left[ \frac{v_{i,j+1,k}^{(2)}}{4} v_{i+1,j+1,k}^{(2)} - \frac{v_{i,j+1,k}^{(2)}}{4} v_{i,j,k}^{(2)} \right] + \\
\left( \frac{1}{2} \mu_{i,j+1,k} + \mu_{i,j,k} \right) \left[ \frac{v_{i,j+1,k}^{(2)}}{4} v_{i,j+1,k}^{(2)} - \frac{v_{i,j+1,k}^{(2)}}{4} v_{i+1,j+1,k}^{(2)} \right]
\]

\[
giving the residual
\]

\[
\text{residual} = \left( \mu_{i,j+1,k} + \mu_{i,j,k} \right) \left[ \frac{v_{i,j+1,k}^{(2)}}{4} v_{i+1,j+1,k}^{(2)} + \frac{1}{4} \left( v_{i,j+1,k}^{(2)} + v_{i+1,j+1,k}^{(2)} - v_{i,j+1,k}^{(2)} - v_{i,j+1,k}^{(2)} \right) \right] + \\
\left( \mu_{i,j+1,k} + \mu_{i,j,k} \right) \left[ \frac{v_{i,j+1,k}^{(2)}}{4} v_{i,j,k}^{(2)} + \frac{1}{4} \left( v_{i,j+1,k}^{(2)} + v_{i,j+1,k}^{(2)} - v_{i,j+1,k}^{(2)} - v_{i,j+1,k}^{(2)} \right) \right] + \\
2 \left( \mu_{i,j+1,k} + \mu_{i,j,k} \right) v_{i,j+1,k}^{(2)} + \left( \mu_{i,j+1,k} + \mu_{i,j,k} \right) v_{i,j+1,k}^{(2)}
\]

\[
(A.28)
\]
A.5. Finite difference scheme for the numerical solution of the varying-viscosity fluid

\[
+ (\mu_{i,j,k+1} + \mu_{i,j,k}) \left[ v_{i,j,k+1}^{(2)} + \frac{1}{4} \left( v_{i,j+1,k+1}^{(3)} + v_{i,j-1,k+1}^{(3)} - v_{i,j-1,k}^{(3)} - v_{i,j+1,k}^{(3)} \right) \right] + \\
(\mu_{i,j,k} + \mu_{i,j,k-1}) \left[ v_{i,j,k-1}^{(2)} + \frac{1}{4} \left( v_{i-1,j,k+1}^{(3)} + v_{i-1,j,k-1}^{(3)} - v_{i-1,j,k}^{(3)} - v_{i+1,j,k}^{(3)} \right) \right] + \\
2f_{i,j,k}^{3} / [8\mu_{i,j,k} + 2(\mu_{i,j,k+1} + \mu_{i,j,k-1}) + \\
(\mu_{i,j,k+1} + \mu_{i,j,k-1} + \mu_{i,j,k+1} + \mu_{i,j,k-1})]
\]

(A.29)

and then the current pixel \( y \)-component of the velocity is updated using

\[
v_{i,j,k}^{(2)+1} = v_{i,j,k}^{(2)+1} (1 - \omega) + \omega \cdot \text{residual}
\]

(A.30)

Finally, the \( z \)-direction component of the variable-viscous PDE is expanded using the finite difference scheme as:

\[
0 = \left( \frac{\mu_{i+1,j,k} + \mu_{i,j,k}}{2} \right) \left[ v_{i,j,k+1}^{(3)} - v_{i,j,k}^{(3)} \right] - \left( \frac{\mu_{i,j,k} + \mu_{i-1,j,k}}{2} \right) \left[ v_{i,j,k}^{(3)} - v_{i,j,k-1}^{(3)} \right] + \\
\left( \frac{\mu_{i,j,k} + \mu_{i,j,k+1}}{2} \right) \left[ v_{i,j,k+1}^{(3)} - v_{i,j,k}^{(3)} \right] - \left( \frac{\mu_{i,j,k} + \mu_{i,j,k-1}}{2} \right) \left[ v_{i,j,k}^{(3)} - v_{i,j,k-1}^{(3)} \right] + \\
\left( \frac{\mu_{i,j,k} + \mu_{i,j,k+1} + \mu_{i,j,k}}{2} \right) \left[ v_{i,j,k+1}^{(3)} + v_{i,j,k-1}^{(3)} - v_{i,j,k-1}^{(3)} - v_{i,j,k+1}^{(3)} \right] + \\
+2 \left( \frac{\mu_{i,j,k+1} + \mu_{i,j,k}}{2} \right) \left[ v_{i,j,k+1}^{(3)} - v_{i,j,k}^{(3)} \right] - \left( \frac{\mu_{i,j,k} + \mu_{i,j,k}-1}{2} \right) \left[ v_{i,j,k}^{(3)} - v_{i,j,k-1}^{(3)} \right] 
\]

(A.31)

giving the residual

\[
\text{residual} = \left[ (\mu_{i+1,j,k} + \mu_{i,j,k}) \left[ v_{i,j,k+1}^{(3)} + v_{i,j,k}^{(3)} \right] + \left( \frac{\mu_{i,j,k} + \mu_{i,j,k+1}}{2} \right) \left[ v_{i,j,k}^{(3)} + v_{i,j,k+1}^{(3)} \right] + \left( \frac{\mu_{i,j,k} + \mu_{i,j,k+1} + \mu_{i,j,k}}{2} \right) \left[ v_{i,j,k+1}^{(3)} + v_{i,j,k-1}^{(3)} \right] + \\
+2 \left( \frac{\mu_{i,j,k+1} + \mu_{i,j,k}}{2} \right) \left[ v_{i,j,k+1}^{(3)} - v_{i,j,k}^{(3)} \right] - \left( \frac{\mu_{i,j,k} + \mu_{i,j,k}-1}{2} \right) \left[ v_{i,j,k}^{(3)} - v_{i,j,k-1}^{(3)} \right] 
\]

(A.32)

and the update given by

\[
v_{i,j,k}^{(3)+1} = v_{i,j,k}^{(3)+1} (1 - \omega) + \omega \cdot \text{residual}
\]

(A.33)
Appendix B

Tables of fluid registration in scale spaces from section 5.3

This appendix tabulates the correlation coefficients of iterations in the fluid registrations tested in section 5.3, and the locations of the source, target and deformed source landmark locations.

<table>
<thead>
<tr>
<th>image number</th>
<th>start</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>iteration</td>
<td>N/A</td>
<td>15</td>
<td>30</td>
<td>45</td>
<td>60</td>
<td>75</td>
<td>83</td>
</tr>
<tr>
<td>corr. coeff</td>
<td>0.544</td>
<td>0.657</td>
<td>0.751</td>
<td>0.834</td>
<td>0.898</td>
<td>0.933</td>
<td>0.939</td>
</tr>
</tbody>
</table>

Table B.1: Increase in the correlation coefficient metric during progress of fluid without scale space for the test in section 5.3 for the images shown in figures 5.10 and 5.11. Looping terminated because the change in correlation coefficient was $-0.000898$, less than the termination threshold of 0.0001
Table B.2: The coordinates of the 10 boundary and 10 internal landmarks, shown in figure 5.12, selected for testing the local registration quality of the fluid model without scale space. The known target landmarks within head1A were inserted into a thin-plate spline together with the landmarks listed here as 'source' to produce the deformed image head1B. Head1B was then registered back to head1A using the fluid model in closure scale space. The resulting deformation when applied to the source landmarks produced small regions whose centres of mass are listed here as 'f(source)'. The columns show alternately x- and y-coordinates.

<table>
<thead>
<tr>
<th>landmark no.:</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>source</td>
<td>54</td>
<td>53</td>
<td>116</td>
<td>28</td>
<td>174</td>
</tr>
<tr>
<td>f(source)</td>
<td>60.2</td>
<td>53.8</td>
<td>115.4</td>
<td>36.8</td>
<td>164.0</td>
</tr>
<tr>
<td>target</td>
<td>81</td>
<td>81</td>
<td>128</td>
<td>49</td>
<td>162</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>landmark no.:</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>source</td>
<td>154</td>
<td>187</td>
<td>101</td>
<td>194</td>
<td>87</td>
</tr>
<tr>
<td>f(source)</td>
<td>131.7</td>
<td>192.1</td>
<td>111.5</td>
<td>187.7</td>
<td>96.0</td>
</tr>
<tr>
<td>target</td>
<td>162</td>
<td>174</td>
<td>132</td>
<td>194</td>
<td>108</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>internal:</th>
<th>landmark no.:</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>source</td>
<td>86</td>
<td>162</td>
<td>178</td>
<td>94</td>
<td>92</td>
<td>81</td>
</tr>
<tr>
<td>f(source)</td>
<td>95.4</td>
<td>169.8</td>
<td>172.0</td>
<td>94.7</td>
<td>100.9</td>
<td>87.6</td>
</tr>
<tr>
<td>target</td>
<td>104</td>
<td>173</td>
<td>170</td>
<td>90</td>
<td>109</td>
<td>98</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>landmark no.:</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>source</td>
<td>115</td>
<td>51</td>
<td>75</td>
<td>55</td>
<td>168</td>
</tr>
<tr>
<td>f(source)</td>
<td>112.0</td>
<td>45.0</td>
<td>84.9</td>
<td>66.2</td>
<td>168.9</td>
</tr>
<tr>
<td>target</td>
<td>127</td>
<td>68</td>
<td>99</td>
<td>80</td>
<td>170</td>
</tr>
</tbody>
</table>
Table B.3: The increase in the correlation coefficient (cc) metric during the progress of the fluid within closure scale space. The image numbers in bold font are those shown in figure 5.14. There are many iterations in the first scale level partly because the threshold of maximum number of SOR iterations was set at 50 so the PDE was not fully solved initially; as the fluid progressed, the computed velocity field became closer to the true velocity field for a given iteration. The second reason is that only fine adjustments (small displacements) were needed for the higher resolution levels since the images were reasonably well-registered by that stage.
Table B.4: The coordinates of the 10 boundary and 10 internal landmarks, shown in figure 5.15, selected for testing the local registration quality of the fluid model within closure scale space.
Table B.5: The coordinates of the 10 boundary and 10 internal landmarks, shown in figure 5.21, selected for testing the local registration quality of the fluid model within Gaussian scale space.
Table B.6: Increase in correlation coefficient during progress of fluid within Gaussian scale space (scale levels 9 to 5). Image numbers in bold are those shown in figure 5.10.
<table>
<thead>
<tr>
<th>Scale Level</th>
<th>Image Number</th>
<th>Total Iteration</th>
<th>Iterations in Level</th>
<th>Correlation Coefficients</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>11</td>
<td>N/A</td>
<td>12</td>
<td>0.822 0.987</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Δcc in final itn: 9.56e-05</td>
</tr>
<tr>
<td>3</td>
<td>12</td>
<td>N/A</td>
<td>3</td>
<td>0.778 0.985</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Δcc in final itn: 3.86e-05</td>
</tr>
<tr>
<td>2</td>
<td>13</td>
<td>N/A</td>
<td>4</td>
<td>0.710 0.986</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Δcc in final itn: -9.06e-05</td>
</tr>
<tr>
<td>1</td>
<td>14</td>
<td>N/A</td>
<td>9</td>
<td>0.623 0.990</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Δcc in final itn: -0.000195</td>
</tr>
<tr>
<td>0</td>
<td>15</td>
<td>N/A</td>
<td>6</td>
<td>0.544 0.979</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Δcc in final itn: -0.000212</td>
</tr>
</tbody>
</table>

Table B.7: Increase in correlation coefficient during progress of fluid within Gaussian scale space (scale levels 4 to zero). Image numbers in bold are those shown in figure 5.11.


Luis Alvarez and Jean-Michel Morel. Morphological approach to multiscale analysis: from principles to equations. In ter Haar Romeny [1994].


Author Index

Aldroubi et al. [1996], 102, 234
Alpert et al. [1990], 48, 50, 234
Alpert et al. [1997], 123, 234
Alvarez and Morel [1994], 122, 234
Alvarez et al. [1992], 77, 234
Amir et al. [1991], 100, 102, 105, 106, 116, 234
Alpert et al. [1990], 48, 50, 234
Alpert et al. [1997], 123, 234
Alvarez and Morel [1994], 122, 234
Alvarez et al. [1992], 77, 234
Andresen et al. [1998], 153, 234
Arad et al. [1994], 84, 90, 93, 234
Arata et al. [1995], 50, 234
Ashbumer [1997], 100, 234
Atkins and Mackiewich [1996], 34, 234
Ayache [1995], 63, 234
Aykroyd and Mardia [1996], 116, 125, 234
Babaud et al. [1986], 116, 234
Bajcsy and Broit [1982], 102, 234
Bajcsy and Kovacic [1989], 102, 105, 118, 234
Bajcsy et al. [1983], 102, 105, 235
Baneijee et al. [1995], 54, 63, 235
Barber et al. [1995], 61, 235
Barrodale et al. [1993], 84, 235
Belohlavek et al. [1996], 35, 234
Berger [1985], 107, 235
Black et al. [1998], 34, 235
Bookstein and Green [1992], 84, 92, 235
Bookstein and Green [1993], 84, 235
Bookstein and Green [1994], 84, 235
Bookstein [1989], 83, 84, 88-90, 235
Borgefors [1986], 181, 213, 235
Brandt and Mikulinsky [1985], 110, 235
Brigger and Unser [1997], 124, 235
Bro-Nielsen and Gramkow [1996], 110, 165, 235
Bro-Nielsen et al. [1997], 153, 235
Brown and Semelka [1995], 29, 235
Brown [1992], 28, 43, 235
Bucholz et al. [1997], 35, 236
Buiten and van Putten [1997], 28, 29, 236
Burton [1981], 122, 236
Bydder [1995], 73, 127, 236
Cai et al. [1996], 52, 236
Chen et al. [1994], 45, 236
Christensen et al. [1993], 102, 106, 236
Christensen et al. [1994], 106, 122, 127, 236
Christensen et al. [1995], 108, 118, 122, 127, 236
Christensen et al. [1994], 78, 102, 106, 108-110, 122, 214, 216, 218, 220, 236
Chui [1992], 102, 236
Collins et al. [1994], 63, 117, 236
Crossman and Neary [1995], 26, 133, 236
Crowley [1981], 122, 236
Dale and Sereno [1993], 34, 236
Danielsson and Hammerin [1992], 43, 237
Danielsson [1980], 214, 236
Davatzikos [1996], 34, 237
Davatzikos [1997], 155, 161, 162, 237
Davis et al. [1997], 94, 95, 237
Decker et al. [1997], 127, 237
Denney and Krzywicki [1967], 82, 237
Desmond et al. [1998], 35, 237
Deubler and Olivo [1997], 124, 237
Dhawan et al. [1995], 48, 50, 237
Downie et al. [1996], 125, 126, 237
Dulaney [1984], 26, 237
Dubreule [1984], 94, 237
Duchon [1976], 84, 237
Dyn and Wahba [1982], 91, 237
Edwards et al. [1997], 157, 164, 237
Ellis et al. [1994], 133, 237
Evans and Spruck [1991], 77, 237
Faber and Stokely [1988], 47, 237
Feldmar et al. [1997], 126, 210, 237
Florack [1993], 116, 120, 237
Flusser [1992], 126, 238
Fogel [1996], 28, 29, 238
Foley et al. [1992], 71, 238
Fonseca and Manjunath [1996], 28, 238
Friston et al. [1995], 83, 96, 118, 238
Gadian [1995], 29, 238
Gamboa-Aldeco and Chen [1986], 47, 238
Gauch [1992], 34, 238
Gee and Perella [1995], 76, 79, 104, 107, 207, 238
Gee et al. [1993], 102, 105, 118, 165, 238
Gee et al. [1994a], 105, 106, 238
Gee et al. [1994b], 105, 106, 123, 126, 238
Gee et al. [1995a], 107, 238
Gee et al. [1995b], 80, 104, 105, 107, 238
Gee et al. [1997], 104, 238
Gelnar et al. [1998], 35, 238
Geman and Geman [1984], 78, 82, 238
Gibbs [1960], 81, 238
Gibson et al. [1998], 35, 153, 238
Gilhuijs and van Herk [1993], 52, 238
Gonzalez and Woods [1992], 55, 238
Goshtasby [1986], 126, 239
Goshtasby [1987], 126, 239
Goshtasby [1988a], 28, 29, 239
Goshtasby [1988b], 84, 239
Grainger and Allison [1997], 26, 239
Grenander et al. [1992], 102, 239
Griffin and Colchester [1995], 118, 239
Griffin [1994], 153, 239
van den Elsen et al. [1993], 63, 244
Subject Index

actively-deforming, 155, 162
adjacent, 36
aliasing, 67
Alzheimer's, 35
Anisotropic fluid, 210
appodising function, 71
ART, 30
atlas, 34, 50
anatomical, 35
probabilistic surface, 34
atrophy, 43
attenuation, 29, 51
back-sampling, 68
basis function, 22, 95–102, 108
dual of displacement field, 66
selection, 100–102
truncations, 77
wavelet, 102
basis functions, 80
radial basis functions, see spline, radial basis function
Bayes' Theorem, 79
Bayes' theorem, 80–82, 96
Bayesian, 76, 78
bending, 77, 78, 83, 88–90, 101, 166, 167, 170, 173, 179, 185, 186, 202
biharmonic equation, 88, 90
boundary conditions, 101, 110
brain, 34, 43, 50, 60, 117
average, standard, 63
blood perfusion, 30
deep structures, 34, 50
function, 31, 33, 73
masked, 60
misregistration, 118
surface, see cortical surface, 51
volume, 35
Cauchy, 109
crown, clown2, 168
complexity
warp, 87
computed tomography, see CT
concatenation, 47, 108
constraint
memory, 65
smoothness, 22, 76–79, 83, 84, 91, 101
time, 65
continuous medium, 77, 108
contrast
enhancement, 34
image, 32

intensity, 32
mapping, 22, 38
medium, 29
tissue, 29, 45, 55, 59, 79
convex hull, 54, 68
convolution, 110
correlation coefficient, 22
correspondence, 44, 45, 51–56, 59, 63, 84, 95
cortical surface, 34, 51
covariance, 22, 101
cross-correlation, 44, 55, 59
cross-modality, see inter-modality
cross-population, 26
variability, 26
CSF, 21
CT, 21, 29, 30, 50, 58–60
curvature, 36, 52, 77
extremal, 53
Gaussian, 53
mean, 53
data
analysis, 34, 35
incomplete, 50, 60
manipulation, 34
medical image, 29–33
deformation, 38, 65
ideal, 65
interpolation for, 72
representations, 66
basis function expansion, 66, 77, 95
combination, 67
displacement field, 66, 108
severe, 108
density
function, 48, 49
of isotopes, 30
of nuclei, 32
of rigid body, 48
pixel intensity, 73, 110
probability density, 57
proton, see proton density
tissue, 26, 29, 30
Dirichlet, 92
dissimilarity measure, see similarity measure, 95
distance function, see also similarity measure, 22, 36, 43, 51, 77
least squares, 51, 77
mean square, 52
distance measure, 48
distance transform, 21
chamfer, 52
Euclidean, 21
Subject Index 250

domain, 22
  image, 48, 58
  source, 38
  spatial, 82
  spectral, 82

EBS, see elastic body spline
edge detection, 28, 51, 54
edge strengths, 77, 78
edgels, 92
EDT, see distance transform
elastic, 67, 78, 94, 102–108
  variable elastic registration, see elasticity
elastic body spline, 21, 94
elasticity
  parameter, 155
  tissue, 155
  variable, 155, 161, 165
energy functional, 22, 80, 82
  conditional, 61
Euclidean distance transform, see distance transform
feature matching, 44, 51
  crest-lines, 52–54
  involutions, 54
  landmark, 51
  surface, 47, 51–52, 77
feature space, 36, 43
FFI, 21
FID, 31, 32
fiducials, see landmarks
finite differences, 82, 89, 108, 165, 219–221
finite element, 156, 157
finite elements, 67, 126
finite support, 89
fluid, 65, 66, 83, 102, 107–110, 154–156, 204
  MF1, 155, 162, 165
  MF2, 156, 162, 165, 166, 169
  MF3, 164, 166, 168
  PDE, 164, 165
fMRI, 21, 29
forces
  balance, 77
  constraining, 108
  driving, 77, 77, 108
  filter, 162
Fourier transform, 71
data acquisition, see FID
deficiency
  of basis functions, 77
  radio, 31–33
  resonance, 31
function
  basis, see basis function
  binary, 48
  cost, see energy functional
  distance, see distance function
  weight, 22
functional
  imaging, 30, 33, 35
  magnetic resonance, see fMRI
Gaussian, 72
  curvature, 53
  distribution, 80, 80, 100, 101
  noise, 96
  radial basis function, 93
  scale space, 118
  smoothing, 117
  geometric invariants, 44
  geometric primitives, 28
  Gibbs distribution, 81, 82
  gradient, 45
  Green's functions, 82
  head1A,head1B, 113
  head2A,head2B, 113
  head3A,head3B, 131
  head4A,head4B, 174
  hierarchy, 83
  data, 112, 116–125, 153, 154
  model, 127, 153, 154, 167, 203, 204
  warp, 112, 125–126, 153, 154
homogeneous
  coordinates, 45, 84
  density, 48
  homologous, 38, 51
  homology, 26
  homomorphic, 77
  Hotelling transform, 101
  house,house2, 168
  hydracephalous, 174
  hyperparameter, 77, 79
image acquisition, 29
image processing pipeline, 34
image-guided surgery, 35
inertia tensor, see tensor
inhomogeneous
  field in scanner, 60
inter-modality, see also registration, inter-modality, 20, 26, 34, 38, 43, 60
inter-subject, 25
interpolation, 67, 77
  B-spline, 72
  back-sampling, 73
  between-slice, 94
  cubic, 70
  holes, 73
  kriging, 94
  linear, 69, 72
  macro-level, 67, 68, 90, 92
  micro-level, see interpolation,pixel-level
  nearest neighbour, 69, 72
  pixel-level, 68, 92
  relaxed, 90
  sinc, 70, 74
  spline, 67, 83, 84
  versus extrapolation, 67, 89
intra-subject, 25, 43
  isoparametric shape functions, 67
Kartunen-Loève, 101
Kullback-Leibler, 57
$L_1, L_2$ norms, 77
Lamé elasticity constant, 78
landmarks, 36, 51, 60, 67, 77, 83–95, 113
  location uncertainty, 90, 94
Laplacian, 77, 91, 101
likelihood, 79, 79, 82, 96
  differential of, 78
multimodal distribution, 80
local minima, 43, 45, 82, 83
magnetic resonance, see MR
MAP, 21, 80, 80, 82
mapping
  contrast, see mapping, pixel intensity
  domain, see transformation
  intensity, 45, 174
  inverse, 73
  pixel intensity, 22, 45, 59–60, 96
  reverse, 67
Markov chain Monte Carlo, see MCMC
Markov random field, 82
masked, 130
mass creation, 109
matching
  feature, see feature matching
  pixel intensity, see pixel matching
mathematical morphology, 51
matrix
  elemental, 45
  linear transformation, 45–47
matrix
  covariance, 22
maximum a posteriori, see MAP
MCMC, 21
MF1, 21
MF2, 21
MF3, 21
minimisation algorithm, 43
misregistration, 59, 60, 108, 118
MMSE, 21, 80
modality, 29, 31, 34, 38, 43, 59
model
  mathematical, 65, 76, 76
  physical, 47, 48, 65, 76, 77, 77, 78, 83, 90, 108
  statistical, 47, 65, 76, 78, 79
underlying, 68
MR, 21, 29, 33, 40, 50, 58, 60, 71, 113
multi-sensor, 28
multi-temporal, 28
multimodal, see inter-modality
multiresolution pyramid, see resolution, levels
mutual information, 21, 22, 45, 55, 56, 56–57, 59–91
Navier-Stokes, 108
noise, 34, 55, 61, 62, 79, 80, 174
Gaussian, 96
removal, 77
occlusions, see data, incomplete
optical flow, 40, 44, 45, 61–62
optimisation, 59, 77, 78, 82
  registration problem, 76, 82
  search algorithm, see search algorithm
outliers, 90
over-relaxation, see SOR
PA, see principal axes
parameter
  elasticity, 22
  regularisation, see regularisation
  viscosity, 22
partial differential equation, see PDE
partition, 37, 38
passive, 154, 162
pattern matching, 28
PDE, 67, 76, 77, 164
  elastic, 104
  fluid, 108, 109, 178, 193, 219
  variable elasticity, 161
  variable viscosity fluid, 164
perioperative, 26
PET, 21, 29, 31, 33
pipeline, 34
pixel matching, 44, 55
pliable sheet, 83, 88, 90
point-spread function, 71
Poisson's ratio, 95
population, 26, 34, 63
positron emission tomography, see PET
posterior, 80–82, 102
practical implementation, 65
pre-/post-operative, 26, 43, 108
pre-processing, 34, 34, 44, 45, 59, 174, 191, 194
precess, 31–33
pressure, 109
principal axes, 21, 44, 47, 47–51, 117
  direction, 53, 54
  moments of inertia, 49
  warps, 89
principal components analysis, 101
prior, 79, 79, 82, 96
prior knowledge, 52
proton density, 32
proton emission tomography, see PET
pyramid, 125
  Gaussian, 122
  Laplacian, 122, 124
  spline, 123
  wavelet, 124
radial basis functions, 67
radiation
  electromagnetic, 31
  exposure, 50
  treatment planning, 35
region, 22, 37, 38
search, 118
registration, 36
  cross-population, 26, 63
  dynamic, 26
  general, 28
  hierarchical, 118
  inhomogeneous, 36, 153–204
  inter-modality, 26, 44, 45, 48, 51, 55, 58, 60
  inter-subject, 25, 26
Subject Index

intra-modality, 45
intra-subject, 25
linear, 36, 43–63
non-linear, 38, 38, 65, 68
perioperative, 26
piecewise-linear, 38, 38, 63
subject-atlas, 26, 34
time studies, see time studies
regridding, 108, 109
Bayesian, 78
cubic, 70
elastic body, 94
kriging, 94
radial basis function, 90, 92, 92–93, 126
Gaussian, 93
quadrics, 92
shifted log, 93
smoothing, 90
thin-plate, 21, 67, 82–95, 108, 113
automated, 91
Dirichlet, 92
distortion, 34, 60

sampling, 29, 39, 65, 67, 72
anisotropic, 51
in-plane, 51
saturation pulse, 31, 32
scale space, 116–125
closure, 121
Gaussian, 116
inhomogeneous, 118
morphological, 121
scanner, 34, 38, 60
CT, 29, 30
distortion, 34, 60
MR, 31–33, 70
PET, 31
ultrasound, 31
scanning
perioperative, 33
schizophrenia, 35
search algorithm, 43, 51, 60, 61
directed search, 44, 52, 61
genetic, 116
gradient descent, 82
Powell descent, 61
stochastic search, 52, 116
search space, 43, 61
segmentation, 34, 34, 35, 45, 51, 59, 153, 155, 164, 191, 192, 198, 204
shape
cortical, 34
signal-to-noise ratio, 33
similarity measure, see also distance function, 22, 43, 43, 55, 76, 77, 83, 90
correlation coefficient, 55

Subject Index 252
cross-correlation, 63
least squares, 51
mutual information, 91
sum of squared differences, 21, 78, 80, 96
simulated annealing, 82
smoothness, see constraint
solution
analytic, 44, 50
fast (fluid), 110
iterative, 44, 50, 55, 91, 108
SOR, 21, 82, 108, 110, 163, 168–169, 219
source, 22, 45, 50
spacing
inter-pixel, 29
spinal chord, 35
spline
automated, 209
cubic, 70
elastic body, 94
kriging, 94
multiquadric, MTQ, 21, 157–160, 165, 166, 186, 203, 204
radial basis function, 90, 92, 92–93, 126
Gaussian, 93
quadrics, 92
shifted log, 93
smoothing, 90
thin-plate, 21, 67, 82–95, 108, 113
automated, 91
Dirichlet, 92
distortion, 34, 60

statistical parametric mapping, see SPM
stretching, 77
strongly deformable, 155, 208
structures, 155, 194, 204
abnormal, 26, 153
brain, 34, 50, 155, 185
knee, 35
relative positions, 77
rigid, 26, 157, 162, 164, 203
within image, 34
sum of squared differences, see similarity measure, 22
superimpose, 35
surgery
image-guided, 35
planning, 35
simulation, 35

T1, 32
T2, 32, 33
target, 22, 45, 50
tensor
inertia, 22, 48, 49, 50
strain, 162, 165
stress, 109, 164
thermodynamic system, 81
thin-plate
analogy for PA, 49
spline, see spline, 78
time studies, 26, 43
timestep, 108
tomography, 29, 30
TPS, see spline, thin-plate
transformation, 22, 36
   affine, 29, 47, 85
   continuous, 77
   homomorphic, 77
   inverse, 73
   linear, 43, 45–47
      parameters, 45, 47
      parameter
         estimated, 60
      perspective, 47
      polynomial, 29
      rigid, 29
      singular, 108
      spline, 29
trend similarity, 59
tumour, 35, 155, 161
      volume, 35
ultrasound, 29, 33
UNC, 21, 130
University of North Carolina, see UNC
velocity field, 108, 154, 156, 162
viscosity, 155
   variable, 110, 154, 164
viscous fluid, see fluid, 108
visualisation, 35
VOI, see volume of interest
volume dilation, 109
volume of interest, 21, 47, 48
wavelet, 125
   basis function expansions, 102
   pyramid, 124, 125
weakly deformable, 155, 171, 174–176, 185, 193, 208
X-ray CT, see CT
xdispunc, 130
xdispunc, 130