

Surface-based Characterization of Intra-surgical Brain Deformations in Image-guided Surgery

Michel A. Audette

Dept. of Biomedical Engineering,
Faculty of Medicine,
McGill University, Montreal

August, 2002

A thesis submitted to the Faculty of Graduate Studies and Research in partial
fulfillment of the requirements of the degree of Doctor of Philosophy.

©Michel A. Audette 2002



National Library
of Canada

Bibliothèque nationale
du Canada

Acquisitions and
Bibliographic Services

Acquisitions et
services bibliographiques

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

ISBN: 0-612-85679-8

Our file Notre référence

ISBN: 0-612-85679-8

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

Canada

Contents

List of Figures	iv
List of Tables	xiii
1 Introduction	1
1.1 Image-guided Surgery	1
1.2 Multimodality Fusion	10
1.3 Accuracy Issues in Image-guided Surgery	15
1.4 A New Approach to Estimating Brain Shift	20
2 A Survey of Surface Registration Techniques for Medical Imaging	23
2.1 Introduction	23
2.2 Choice of Transformation	26
2.2.1 Rigid-body Transformation	26
2.2.2 Nonrigid Transformations	29
2.3 Surface Representation and Similarity	
Criterion	34
2.3.1 Feature-Based Representations	35
2.3.2 Point-Based Methods	39
2.3.3 Model-Based Representations	42
2.3.4 Techniques Based on Global Shape	44
2.4 Matching, Optimization and Transformation Computation	47
2.4.1 Discrete Feature Matching and Transformation	
Computation	48
2.4.2 Closest Point Finding and Global Optimization	50
2.4.3 Model-Based Motion Estimation	52
2.5 Conclusion	53
3 Laser Range-sensing and Calibration-based Range-MR Transformation	56
3.1 Introduction	56
3.2 Laser Range-sensing	59
3.3 Relating Range to MR Space by Probe-based Calibration, Sensor Base	
Tracking and IGS	
Patient Registration	64

3.4	Validation of the Range-MR Transformation Procedure	71
3.4.1	Materials and Methods	71
3.4.2	Validation Results	75
3.4.3	Discussion	75
3.5	Conclusion	80
4	Non-rigid Surface Displacement Estimation	82
4.1	Motivation and Overview of Non-rigid Cortical Motion Estimation	82
4.2	ICP Registration Technique: Implementation Details	85
4.2.1	ICP Homologous Pairs by Look-up: the Closest Point Map . .	85
4.2.2	Exploiting Regular Range Data Spacing: Surface Splines by Recursive Filtering	90
4.2.3	Special Considerations for Range-MR Non-rigid Surface Registration	94
4.3	Non-rigid Registration Validation	102
4.3.1	Brain Shift Simulation by Realistic Elastic Brain Phantom	102
4.3.2	Validation with an Analytical Displacement Applied to an An- thropomorphic Digital Brain Surface	107
4.4	Conclusions and Future Directions	115
5	Model-based Brain Surface Identification in MRI: Motivation, Model Overview and Surveys	117
5.1	Introduction: Motivation for Model-based Brain Surface Identification	117
5.2	An Intuitive Overview of Surface Evolution Models	121
5.3	A Concise Review of Model-based Surface Identification and Image Filtering	129
5.3.1	Model-based Surface Identification Techniques	129
5.3.2	Image Filtering Techniques	136
5.4	Conclusions	139
6	Model-Based Brain Surface Identification in MRI: Implementation Details	141
6.1	Surface Evolution Model	141
6.2	Integrating Anatomical Information into the Model: the WM/GM Dis- tance Function	147
6.3	Semi-automatic Voxel Classification	154
6.4	Numerical Implementation of the Surface Model	158
6.4.1	Computing Signed Distance with the Fast Marching Technique	158
6.4.2	Numerical Implementation of Surface Evolution Equation	161
6.5	Conclusions	163

7	Model-Based Surface Identification in MRI: Validation	165
7.1	Validation Tools: MRI Simulator and Realistic Digital Brain Phantom	165
7.2	Anisotropic Diffusion Validation	169
7.3	Validation of Semi-Automatic Voxel Classification	175
7.4	Fast Marching Distance Computation Validation	182
7.5	Surface Model Validation	186
7.5.1	Healthy Subject Validation: Synthetic Digital Phantom-based Data	186
7.5.2	Healthy Subject Validation: ICBM Real Subject Data	194
7.5.3	Pathological Subject Validation: Synthetic Digital Phantom-based Data	197
7.5.4	Pathological Subject Validation: Real Patient Data	197
7.6	Discussion and Conclusion	200
8	Conclusions and Future Directions	202
A	A Survey of Biomechanical Modeling of the Brain for Intra-surgical Displacement Estimation and Medical Simulation	207
A.1	Introduction	207
A.2	Preliminaries	208
A.2.1	Biomechanics	208
A.2.2	Numerical Estimation	212
A.3	Finite Element Modeling of the Brain	217
A.3.1	Solid Brain Models	217
A.3.2	Hybrid Solid-Liquid Brain Models	224
A.4	Summary	228
B	Derivation of Surface Evolution Model	229

List of Figures

1.1	Image-guided surgery: 3D image model of patient as presented in typical commercial IGS system, namely in tri-planar and 3D view formats, courtesy of Lahbib Soualmi of the MNI.	2
1.2	Frame-based localization: (a) the Olivier-Bertrand-Tipal (OBT) frame, (b) mounting arc for biopsy needles, (c) geometry of frame-based N-shaped localizer (adapted, with permission, from [423]).	5
1.3	Frameless localization devices: the FARO arm, during (a) probe calibration and (b) intraoperative use; (c) the Northern Digital Polaris optical tracking system.	6
1.4	Integration of MRI with stereoscopic DSA image pair [149].	11
1.5	An illustration of co-registered MRS and MRI data: superimposed choline (colour) and T1-weighted (gray scale) axial images, courtesy of Mark Preul [315].	14
1.6	Brain shift estimation framework, featuring the four stages comprising our non-rigid surface registration framework, which can be combined with a FE-based volumetric displacement estimation stage.	21
2.1	3D surface registration framework, featuring the choice of transformation between two anatomical surfaces (and of the search space for transformation parameters); the representation of the surface in terms of matching primitives (e.g. regions of consistent curvature); and finally, the matching of surface primitives and computation of the “best” transformation.	25
2.2	Illustration of the dual quaternion model for rigid transformation [408] (reproduced with the permission of R.A. Volz, copyright Academic Press 1991).	29
2.3	(a) Range image of femur epiphysis; (b) regions based on K/H sign combination. Both are courtesy of Gilbert Soucy of the McGill Centre for Intelligent Machines.	39

2.4	Iterative point-based registration of phantom face range data (courtesy of S. Lavallée, R. Szeliski and L. Brunie, copyright MIT Press 1986 [210]).	40
2.5	(a) An illustration spin-map geometry; (b) points distributed on the surface of a femur bone, shown with three point bases and their corresponding spin-maps, appearing as point sets and grey-scale images (courtesy of A.E. Johnson).	45
3.1	Range image of the left hemisphere of a human brain.	57
3.2	Range to MR transformation geometry, illustrating the range, probe and MR referentials and the sequence of procedures for relating them to each other.	60
3.3	Illustration of range-sensor and positioner configuration in the laboratory, with labelled components.	61
3.4	Biris laser range-sensing geometry: (a) triangulation; (b) defocusing	62
3.5	Range to MRI transformation computation. At $t = t_c$, the calibration procedure first relates range to probe space (a), then determines the sensor base points in range space (b). (c) Sensor base tracking determines sensor base points in probe space at $t = t_o$ and finds the range-probe transformation in the OR. (d) The final step uses this and the probe-MR transformation to map range to MR.	66
3.6	Precision machined tools for calibration: (a) positioner calibration plate and (b) “wedding cake” sensor calibration tool.	68
3.7	Illustration of calibration procedure: (a) range image of calibration plate, displayed as a grey-level; (b) intensity image overlapping range image, used to refine pit detection; (c) detected pits in range image; (d) range images comprising the 2-level calibration procedure.	69
3.8	(a) <i>Poupoune</i> , the exotic skin phantom, prior to validation study, and (b) in validation configuration. Illustration of least-squares IGS image-probe transformation and the 7 homologous point pairs used for left (c) and right (d) side studies: CT space (green) and probe space fiducials transformed to CT space (light blue), overlaid on outer skin phantom voxel.	72
3.9	Range-CT transformation procedure results (trial 1): (a) left side range surface data transformed to CT space, overlaid on CT skin phantom voxels; (b) original 7 CT space (green) and transformed range space (pink) fiducials of left side study; (c) transformed right side range surface data overlaid on CT skin voxels, as in (a); (d) CT space and transformed range space fiducials of right side study, as in (b).	74

4.1	Elastic PVA-C brain phantom: (a) liquid stage during construction (b) final result after freezing and thawing.	85
4.2	Elastic brain phantom surface from MRI scan (rotated for better visibility). The outline of the plate used to implement deformation can be discerned within the boundary of the phantom.	86
4.3	(a) Distance and (b) Closest Point Maps computed from the surface of the elastic phantom (white), and shown in tri-planar views, mapped to a grey level. In (b) ambiguous voxels, possessing more than one closest surface point, are shown in black (see discussion in § 7.4).	87
4.4	Symbolic description of non-rigid surface registration. Top: the labels of the identified brain surface voxels are symbolized by colour. Row 2: the brain surface labels are propagated to all non-surface voxels. Row 3: the range data, of initially arbitrary orientation, are transformed to MR space on the basis of the sensor-base tracking procedure. Bottom: each range point inherits a Closest Point Map label from the MR voxel with which it coincides, thereby specifying its MR homologous point.	88
4.5	Frequency of response of the recursive approximating spline filter.	92
4.6	Illustration of recursive spline filter response, given an input consisting of a sum of a ramp and 7 harmonic sinusoids, as λ decreases from 128 to 0.5.	93
4.7	Range data featuring unreliable, low-signal pixels which are identified and eliminated from further consideration.	95
4.8	Dealing with many-to-one range-MR matches arising from unequal data density (MR points in white, range points in black). Top: true surface displacement $\mathbf{D}_S$; second row: displacement field implicit in many-to-one homologous pairs; third row: disambiguation by pruning all but mutually closest homologous pairs; and bottom: distance-weighted propagation.	96
4.9	Multiple or spurious MR closest point: disambiguation from continuous displacement assumption, by neighbourhood estimate.	98
4.10	Generating the MR-range transformation $\mathbf{T}_M^R$ from $\mathbf{T}_R^M$. Top: overall rigid range-MR transformation $\mathbf{T}_{R,r}^M$; middle row: overall non-rigid range-MR displacement $\mathbf{T}_{R,n}^M$. Bottom: the non-rigid and rigid components of the transformation which maps the MR brain surface to the range image are as follows: $\mathbf{T}_{M,n}^R$ is the negative of $\mathbf{T}_{R,n}^M$, and $\mathbf{T}_{M,r}^R$ is the transpose of $\mathbf{T}_{R,r}^M$	100

4.11	Elastic brain phantom surface from MRI scan with roughly registered range image overlaid, decimated by a factor of 2 (both rotated for improved visibility).	103
4.12	Statistics of distances between homologous points, assuming an overall transformation $\mathbf{T}_R^M$ (based on expressions 4.8 evaluated up to iteration $K = k$ or $L = k$) applied to range points: (a) $\log_{10}$ of distance average; (b) $\log_{10}$ of distance standard deviation.	104
4.13	Evolution of forward range-MR transformation over iterations k and progressively more local characterization. The colour scheme of the range data indicates a homologous MR point within a certain distance d of the range surface: red: $d < 0.5mm$, yellow: $0.5 \leq d < 2.0mm$, green: $d \geq 2.0mm$.	105
4.14	Evolution of the inverse MR-range transformation over iterations k and progressively more local characterization. The distance between homologous points is shown as a coloured point in MR space. The colour is determined by the logarithmic “hot” colour scale on the right. Black corresponds to -1.0 (0.1mm), saturated red to -0.17 (0.67mm), saturated yellow to 0.67 (4.78mm), and white to 1.3 (20mm).	108
4.15	Evolution of inverse MR-range transformation versus iteration k and progressively more local characterization, for the <i>slice</i> $y_R = 79.2$ ($j = 192$), featuring homologous range points and MR brain points mapped back to range space according to $\mathbf{T}_{M,k}^R$ and projected to the $x_R - z_R$ plane.	109
4.16	Influence of smoothness parameter value for forward transformation: (a) distance statistics and (b) final forward transformation results, for $\epsilon_{N,m} = \bar{\epsilon}_N + \sigma_{\epsilon_N}$ (top); $\epsilon_{N,m} = \bar{\epsilon}_N + 2 \times \sigma_{\epsilon_N}$ (middle row); $\epsilon_{N,m} = \infty$ (bottom). Compare with results for $\epsilon_{N,m} = \bar{\epsilon}_N + 3 \times \sigma_{\epsilon_N}$ appearing in figures 4.12 and 4.13 (f).	110
4.17	Influence of smoothness parameter value for inverse transformation: (a) final inverse transformation colour map and (b) 2D slice results, for $\epsilon_{N,m} = \bar{\epsilon}_N + \sigma_{\epsilon_N}$ (top); $\epsilon_{N,m} = \bar{\epsilon}_N + 2 \times \sigma_{\epsilon_N}$ (middle row); $\epsilon_{N,m} = \infty$ (bottom). Compare with results for $\epsilon_{N,m} = \bar{\epsilon}_N + 3 \times \sigma_{\epsilon_N}$ appearing in figures 4.13 (f), 4.14 (f) and 4.15 (f).	111
4.18	Two synthetic range images overlaid on original phantom brain surface data, from which they were obtained by interpolation.	112

4.19	Analytical deformation results: (a) top study and (b) left study. Top row: initial deformation $A + B\cos(2\pi f(x_1 - x_{1,c}))\cos(2\pi f(x_2 - x_{2,c}))$, where $A = 15mm$ and $B = 5mm$; middle row: final rigid registration; bottom row: final non-rigid registration. Red signifies that the deformation is estimated with error $\leq 1.0 mm$, yellow: $1.0 mm < \text{error} \leq 2.0 mm$, and green: error $> 2.0 mm$	114
5.1	Imbedding function with 2D contour example, with inward-moving zero-level isocontour, based on synthetic image featuring 2 elliptical bodies: (a) zero-level isocontour (white), overlaid on image at $t = 0$, at $t = 15$ sec., and at $t = 60$ sec.; (b) imbedding function Ψ , with level set $\{\Psi = 0\}$ coinciding with isocontour in (a), at $t = 0$, at $t = 15$ sec., and at $t = 60$ sec.; (c) narrow band implementation of contour evolution model, with identical situation as in (b): the computation of Ψ is restricted within a narrow band near the evolving zero-level isocontour.	122
5.2	Identification of elliptical boundaries as in figure 5.1, with a physically-based model without topological adaptation, courtesy of Johan Montagnat [86]: initial configuration, 100th iteration, final iteration. . . .	123
5.3	Demonstration of difference in tensile properties of evolution and physical models, on 2D invaginated circle: (a) contour evolution model at initialization ($t=0$), at 10 sec., and at 30 sec; (b) physical model at initialization, at 8th iteration, and at 30th iteration, courtesy of Johan Montagnat [86].	125
5.4	Contour identification with outward-moving isocontour: zero-level isocontour (white), overlaid on invaginated circle image as in figure 5.3; at $t = 0$, at $t = 10$ sec. and at $t = 60$ sec.. . . .	125
5.5	3D surface identification with synthetic volume featuring (a) a pair of ellipsoids and (b) an invaginated sphere: multiplanar and rendered views of $\Psi(x, y, z) = 0$ at initial, intermediate and final instants. . . .	126
5.6	2D illustration of each term at work: (a) diffusive: at 0 sec., at 10 sec., and at 120 sec.; (b) hyperbolic: at 0 sec., at 10 sec., and at 30 sec.; (c) speed and doublet: the model behaviour is shown above in the absence of these two terms (compare with fig. 5.1): at 0 sec., at 15 sec., and at 30 sec.. (d) Illustration in 2D of doublet term: 1. raw 2D signal $I(x, y)$ with contour $\mathcal{C}$ approaching boundary; 2. speed function $F(x, y)$ inversely related to the gradient magnitude of the filtered signal; and 3. doublet term $(-\nabla F \cdot \mathbf{N}) \mathbf{N}$: the factor $-\nabla F$ acts as a sink near the minimum, attractive where $(-\nabla F \cdot \mathbf{N})$ is positive, repulsive where $(-\nabla F \cdot \mathbf{N})$ is negative.	127

5.7	Left: illustration of 2-simplex mesh, shown in black, with the dual triangulation shown in grey; Right: simplex representation of a brain surface, shown for three levels of meshing resolution (courtesy of Johan Montagnat) [271].	131
6.1	Algorithmic description of the tissue class-based surface evolution model framework presented in this thesis.	145
6.2	Pictorial description of the tissue class-based surface evolution model framework presented in this thesis.	146
6.3	Axial and coronal views of evolving cortical surface model with its image term solely based on gradient magnitude. (a) With a low gradient threshold (allowing points of low gradient magnitude to halt the zero-level surface), the surface may converge to the WM boundary rather than the outer brain boundary. (b) With a higher threshold, the surface can bleed through the outer brain boundary, and never actually converge to a solution.	149
6.4	WM/GM Distance Function: (a) illustration of parameters used in computation, featuring ellipsoid roughly equated with WM data and used for normalizing $\varepsilon_{WM/GM}$, as well as inner and outer thresholds, overlaid on simplified sagittal plane of anatomy; (b) WDF mapped to a grey-level intensity with WM contours overlaid in white; (c) speed function based on membership of voxels classified as WM or GM, using T1 alone without a contrast agent (note that there is little discrimination from T1 alone between GM and muscle, skin or sagittal sinus voxels); (d) speed function based on WM/GM voxels, as in (c), but excluding GM voxels which exceed the WDF shown in (b).	152
6.5	Illustration of voxel classification, for T1-, T2- and PD-Weighted feature space, based on the middle coronal slice of simulated data with 5% noise (see discussion in § 7.1), preprocessed with anisotropic diffusion ($t = 0.5sec$).	155
6.6	Simplified illustration of partitioning of parameter space by (a) the Minimum Distance classifier, where the centroids and feature space partition is computed from training points, and (b) the semi-supervised Fuzzy-C Means classifier, which iteratively refines the MD-partitioned feature space on the basis of both training and putatively classified data points. Training points are shown in solid colours, data points are shown as white-centered shapes and centroids of each class are marked with an X.	157

6.7	Computation of Distance and Closest Point Map for 2D contour $\mathcal{C}$, featuring the “arrival time” of the contour (whose value is symbolized by a shade of blue: the darker the blue, the lesser the arrival time): $\mathbf{p}_0$ is the closest contour point for $\mathbf{p}_1$ and $\mathbf{p}_2$	159
6.8	Applications of Distance and Closest Point Map, for 2D contour example: (a) the sign computed for the signed distance map is determined by the sign of the dot-product between the normal and the vector from the closest surface point to the current point of interest; (b) the restriction of the computation of the evolution equation to a narrow band close to the zero-level also exploits the Fast Marching distance computation; (c) the extension of image information from a contour point to other points closest to it: the contour point $\mathbf{p}_0$, being on a anatomical boundary, stabilizes the value of $\Psi(\mathbf{p}_1)$ and $\Psi(\mathbf{p}_2)$	162
7.1	An overview of the simulation process (courtesy of G. Bruce Pike [198]).	167
7.2	(a)-(d) 3D digital healthy brain phantom and simulated MR data (noise 3%, inhomogeneity 0%): (a) phantom, (b) T1-, (c) T2- and (d) PD-weighted volumes. (e)-(g) Analogous phantom as in (a), with synthetic astrocytoma inserted: (e) phantom, (f) T1- and (g) T2-weighted volumes, based on data in Bottomley [43]. No PD information on brain pathologies was found in the literature.	168
7.3	Illustration of the effect of the scale parameter, for (a) simulated T1-weighted volume (with null noise and inhomogeneity values) on (b) Gaussian and (c) anisotropic diffusion.	170
7.4	2D Illustration of measurements for validation and choice of scale of anisotropic diffusion: (a) phantom class data, with 1D profile along normal direction, through known boundary point; (b) filtered simulated T1-weighted MR data, with 1D profile as in (a); (c) magnitude of the gradient of filtered T1 data, with 1D profile as (a) and (b), with measurements shown.	172
7.5	Measurement statistics for validation and choice of scale of anisotropic diffusion, under different noise and 0% RF inhomogeneity conditions. Average and standard deviation of: 1. local maximum gradient magnitude under (a) Gaussian filtering (GF) and (b) anisotropic diffusion (AD); 2. localization error of local gradient maximum from known boundary position under (c) GF and (d) AD; 3. shoulder (80% of gradient magnitude maximum) width value along the boundary under (e) GF and (f) AD.	174

7.6	Illustration of sensitivity of MD classification to noise levels, to pre-processing method and scale, and to number of modalities: percentage of correct voxels, 1. based on T1-weighted scan alone, at various noise levels, under (a) Gaussian filtering (GF) and (b) anisotropic diffusion (AD); 2. based on T1- and T2-weighted scans, under (c) GF and (d) AD; 3. based on T1-, T2- and PD-weighted scans, under (e) GF and (f) AD.	177
7.7	Spatial distribution of misclassified voxels, based on accumulators in the $y - z$ (sagittal) plane. The brighter the pixel, the larger the number of misclassified voxels having that x position. Results for (a) T1-weighted scan alone and (b) T1- and T2-weighted scans.	178
7.8	Percentage of correct voxels of white and grey matter of MD classifier, at various noise levels, scales and combinations of modalities (using anisotropic diffusion): (a) WM and (b) GM results based on T1-weighted scan alone; (c) WM and (d) GM results based on T1- and T2-weighted scans; and (e) WM and (f) GM results based on T1-, T2- and PD-weighted scans.	179
7.9	Illustration of sensitivity of FCM classification to noise levels, to pre-processing method and scale, and to number of modalities: percentage of correct voxels based on (a) T1-weighted scan alone, (b) on T1- and T2-weighted scans, and (c) on T1-, T2- and proton density-weighted scans, all after anisotropic diffusion; (d) illustration of migration of centroids over successive iterations caused by border effects at noise=5% and scale $t=2.0$ sec., where a "o" illustrates the position of the class centroid at the first iteration (coinciding with MD classification) and an "x" indicates the final (25th) iteration. Note the final position of MUSCLE/SKIN class centroid nearly coinciding with the initial GM centroid position.	180
7.10	(a) Distance Map to brain surface produced by Fast Marching (FM) technique, with distance values ranging in the volume from 0 to 100 pixels are mapped to grey-levels 0 to 192, and brain boundary appearing in white; (b) Distance Map to brain surface produced by Brute Force (BF) technique, with same distance to grey-level mapping as in a; (c) difference between FM and BF Distance Maps, shown as grey-level image: a voxel error ranging from 0.0 to 2.0 voxels maps to 0-255; (d) difference between FM and BF, averaged over voxels of similar distance to the brain surface: the average FM-BF difference value for each bin appears in the top histogram, whereas the number of points summed in each bin is shown in the lower histogram.	183

7.11	Closest Point Maps computed by 3D (a) FM and (b) BF techniques, for the brain boundary, labelled white, with surface labels 1-207459 mapped to grey levels 64-192, and with ambiguously labelled voxels shown as black (0); in (c), a bar graph illustrates the percentage of labels in agreement, as a function of distance to surface, based on the same voxel populations presented in figure 7.10 (d); (d) FM Closest Point Map with ambiguities due to digitization artefacts eliminated with a distance threshold of 2 voxels.	185
7.12	(a) Tissue classes produced by MD classifier and (b)-(e) brain surface identified with surface model, from simulated T1-weighted scan: (b) tri-planar view; (c) rendered view; (d) distribution of distance between brain surface in scan and phantom brain boundary; (e) error between brain surface identified in scan and underlying brain boundary in phantom, shown as colour plot: avg.= 0.48 ± 0.71 mm.	188
7.13	Tissue classes produced by MD classifier, as in figure 7.12, for (a) T1- and T2-weighted data (avg.= 0.53 ± 0.74 mm.), and (b) T1-, T2- and PD-weighted data (avg.= 0.50 ± 0.65 mm.).	190
7.14	Sensitivity to noise level: plots of statistics of final surface to brain boundary error (left) and number of surface voxels (right) against noise level; (b)-(f) tri-planar and rendered views of final surface, plus 3D colour plot of error, for (b) noise= 0%, (c) noise=3%, (d) noise=7%, (e) noise=10%, and (f) noise=20%.	191
7.15	Sensitivity to model parameters (T1): plots of statistics of final surface to brain boundary error (left) and number of surface voxels (right) versus a variation of the following WDF thresholds: (a) outer cortex; (b) frontal lower brain; (c) cerebellum; and (d) inner brain.	192
7.16	Sensitivity to model parameters (T1): rendered views of final surface, plus 3D colour plot of error, for variations of the following WDF thresholds, about a standard value indicated by an asterisk (*): (a) outer cortex; (b) frontal lower brain; (c) cerebellum; and (d) inner brain. . .	193
7.17	Evolution of the surface model, with real T1-weighted data: (a) ICBM subject 00100; (b) ICBM subject 00101; (c) ICBM subject 00102. . .	195
7.18	Application to ICBM subjects (a) 00103 and (b) 00104, based on the model acting on T1 data alone, on T1 and T2, and on T1, T2, and PD. Top: tissue classes based on each feature combination; bottom: tri-planar and rendered views of the corresponding final surface. . . .	196

7.19	Surface model behaviour with synthetic pathological data: (a) original pathological phantom and (b)-(e) brain surface identified with surface model, from simulated T1-weighted scan: (b) tri-planar view; (c) rendered view; (d) distribution of distance between brain surface in scan and phantom brain boundary; (e) error between brain surface identified in scan and underlying brain boundary in phantom, shown as colour plot: avg.= 0.68 ± 0.51 mm.	198
7.20	Surface model behaviour with real pathological T1-weighted data; (a)-(c) patient 1: (a) classification results, featuring 7 classes; (b) tri-planar view; (c) rendered view; (d)-(f) patient 2: (d) classification results, featuring 7 classes; (e) tri-planar view; (f) rendered view.	199
A.1	Meshing used at Wayne State University, courtesy of C. Zhou [425]. .	210
A.2	A hierarchical FE mesh, courtesy of O. Astley [10]: (a) initial parent mesh; (b) definition of volume of interest for child mesh (c); (d) refinement of child mesh, then integrated with parent mesh.	215
A.3	Brain tissue rheological studies: (a) illustration of uniaxial compression apparatus; (b) layout with coordinate axes. Components: 1- specimen and loading platens, 2- load cell to measure axial force, 3- micrometer to measure axial displacement, and 4- laser to measure radial displacement. Courtesy of Karol Miller and Kiyoyuki Chinzei [262].	222
A.4	Brain tissue rheological studies: illustration of uniaxial compression results vs. , under three different loading velocities: (a) 500 mm/min; (b) 5 mm/min; (c) 0.005 mm/min, courtesy of Karol Miller and Kiyoyuki Chinzei [262].	224

List of Tables

3.1	Trial results for calibration and sensor base tracking procedure, for range images of the left and right sides of the head. Rigid body transformation by least-squares and by robust regression, with and without ICP refinement based on the skin identified in CT and range spaces. The value ϵ_{rms} represents the RMS error (in mm) between CT fiducial positions and their homologues in the range image transformed to CT space by $\mathbf{T}_{range}^{CT}$. The value d_{surf} is the RMS distance (in mm) between the transformed range points and the skin surface in CT space, as read from a precomputed distance map. The symbols $\ \mathbf{t}_c^o\ $ and θ_c^o indicate the magnitude of the translation and rotation between the sensor base positions at t_c and t_o	76
4.1	Analytical deformation statistics. Left column: deformation parameters (in mm, mm and mm^{-1}), as per expression (4.11). Middle column: magnitude of vector difference between estimated and analytical displacement (at the start, after the final rigid iteration and after the final non-rigid iteration). Right column: distance between mutually closest points of the synthetic range and MR surfaces.	113

Abstract

Image-guided surgery (IGS) is a technique for localizing anatomical structures on the basis of volumetric image data and for determining the optimal surgical path to reach these structures, by the means of a localization device, or probe, whose position is tracked over time. The usefulness of this technology hinges on the accuracy of the transformation between the image volume and the space occupied by the patient anatomy and spanned by the probe. Unfortunately, in neurosurgery this transformation can be degraded by intra-surgical brain shift, which often measures more than 10 mm and can exceed 25 mm. I propose a method for characterizing brain shift that is based on non-rigid surface registration, and can be combined with a constitutively realistic finite element approach for volumetric displacement estimation.

The proposed registration method integrates in a unified framework all of the stages required to estimate the movement of the cortical surface in the operating room: model-based segmentation of the pre-operative brain surface in magnetic resonance image data, range-sensing of the cortex in the OR, range-MR rigid transformation computation, and range-based non-rigid brain motion estimation. The brain segmentation technique is an adaptation of the surface evolution model. Its convergence to the brain boundary is the result of a speed term restricted to white and grey matter voxels made explicit by a classifier, and the final result is post-processed to yield a *Closest Point Map* of the brain surface in MR space. In turn, this Closest Point Map is used to produce the homologous pairs required to determine a highly efficient, 2D spline-based, Iterative Closest Point (ICP) non-rigid surface registration. The baseline for computing intra-operative brain displacement, as well as the initial starting point of the non-rigid ICP registration, is determined by a very good rigid range-MR transformation, produced by a simple procedure for relating the range coordinate system to that of the probe, and ultimately to that of the MR volume.

Finally, this thesis also features a review of image guidance for brain surgery, a published survey of anatomical surface registration, and a literature review of biomechanical modeling of brain tissue.

Résumé

La chirurgie guidée par image est une technique pour localiser des structures anatomiques à partir d'images volumétriques et pour déterminer le trajet chirurgical optimal pour s'y rendre, par le biais d'un outil de localisation dont la position est suivie dans le temps. L'utilité de cette technologie dépend de la précision de la transformée entre l'image volumétrique et l'espace occupé par l'anatomie du patient et arpenté par la sonde. Toutefois, cette transformation peut être compromise par une déformation corticale intra-chirurgicale, dont l'amplitude dépasse fréquemment 10 mm et peut atteindre 25 mm. Je propose une méthode pour caractériser cette déformation, qui est basée sur un recalage non-rigide de surfaces, et qui peut être combinée avec une méthode par éléments finis pour estimer le déplacement volumétrique, comportant des propriétés matérielles réalistes.

Le méthode de recalage proposée intègre dans une approche unifiée toutes les étapes requises pour estimer le mouvement non-rigide de la surface corticale dans la salle d'opération: la segmentation de la surface externe du cerveau dans le volume IRM, la capture de la morphologie 3D du cortex par triangulation au laser, le calcul de la transformée rigide entre les espaces capteur-laser et IRM, et l'estimation du mouvement non-rigide à partir de scans 3D du capteur-laser. La technique utilisée pour la segmentation est une adaptation du modèle dit "évolution de surface". Sa convergence vers la surface externe du cerveau résulte d'une fonction spatiale de vitesse qui restreint cette évolution aux voxels de matière blanche et grise déterminés de façon explicite par un classificateur. Le résultat final de la segmentation est traité par la suite pour produire une Carte des Points Les Plus Près issue de la surface corticale dans le volume IRM. Par ailleurs, cette carte est exploitée pour définir les paires de points homologues qui détermineront un recalage non-rigide de surfaces basé sur des splines 2D et de type "itératif et à partir des points les plus près". Enfin, la référence pour le calcul du mouvement intra-opératoire, et le point de départ pour le recalage non-rigide, sont obtenus par une très bonne transformée entre capteur-laser et IRM, produite par une simple procédure pour apparier l'espace capteur-laser à

l'espace sonde, et ultimement à l'espace IRM.

Enfin, cette thèse comporte une revue de la chirurgie guidée par image, une revue publiée du recalage de surfaces anatomiques, et un survol de la modélisation biomécanique du cerveau.

Dédicace

A la mémoire de mon père défunt, Paul Michel Audette, et à ma mère Louise, à mes soeurs Renée et Marie-Nicole, à mon frère Louis-Philippe, et à leurs conjoints respectifs, en reconnaissance de leur soutien indéfectible.

Acknowledgements

First, I would like to thank my supervisors T.M. Peters and K. Siddiqi for their guidance and encouragement throughout this thesis research. Prof. Peters made it possible for me to make a transition from full-time professional engineer to half-time engineer and student, and eventually to full-time student, and was frequently a voice of reason who helped rein in some of my more speculative ideas. Prof. Siddiqi made a timely intervention in my research and helped me understand surface evolution models and overcome previously insurmountable obstacles. I am also indebted to my other thesis committee members: F.P. Ferrie of the McGill Center for Intelligent Machines (CIM), my unofficial third supervisor, for his time and support with the range-sensor and surface registration implementations; R.E. Kearney of the McGill Biomedical Engineering Dept., for his conscientious approach to improving the rigour of my research and keeping its scope reasonable; and J. Nemes of the McGill Mechanical Engineering Dept., for his interest in my research and genuine willingness to help with the continuum mechanics.

Second, I am indebted to many collaborators and colleagues, for their support at the implementation level. In no particular order, E. Johnstone of the McGill Biomedical Engineering Dept. for machine shop assistance; former MNI students R. Comeau, R. Kasrai, Ph. St-Jean, L.R. Ostergaard, J. Atkinson and F. Van Meer for their help on every aspect of my research ranging from software implementation to machine shop assistance; G.B. Pike, D.L. Collins, A.C. Evans and A. Zijdenbos of the MNI for their mentorship; L. Soualmi of the MNI for generously making image guidance equipment and clinical data available to me; students and members of other institutes with whom I collaborated: D. Gobbi of the Robarts Inst. as well as S. Bouix

and G. Soucy of McGill CIM for help on software implementation; and computer support people at the MNI: P. Neelin and J.F. Malouin.

I would also like to thank my current collaborators at the Advanced Inst. for Industrial Science and Technology (AIST) in Japan, K. Chinzei and Y. Koseki in particular, for their patience and understanding of my less-than-optimal performance in my new position as research fellow at AIST, due to my part-time work on the last few Ph.D. simulations and on the iterative editing of this thesis.

Finally, I would like to thank the members of my family for their love and herculean show of support, both financial and emotional, in the last few years of this research. I will never forget what they did for me, in order for me to finish.

Preface

Image-guided surgery (IGS) describes a class of techniques for localizing anatomical structures on the basis of image data and for determining the optimal surgical path to reach these structures. With the advent of tomographic image volumes produced by computed tomography (CT) and magnetic resonance (MR), it has become possible to elaborate a 3D image model of the patient, whose internal coordinate system (“*image space*”) can be related to that of a *localization device* spanning the physical space occupied by the patient (“*patient space*”). However, the usefulness of image-guided surgery hinges on the accuracy of the transformation between the image and patient spaces. The most important source of error in IGS is intrasurgical brain shift.

In this thesis, I propose a method for characterizing brain shift that is based on non-rigid surface registration, and can be combined with a constitutively realistic finite element approach for volumetric displacement estimation [259, 358, 110]. This registration method integrates in a unified framework all of the stages required to estimate the movement of the cortical surface in the operating room (OR):

- semi-automatic identification (segmentation) of relevant anatomical surfaces within the MRI volume;
- range-sensing of the exposed brain surface in the OR;
- computation of the range-MRI transformation, based on a calibration procedure and sensor base tracking; and
- non-rigid motion tracking, over time, of the range image of the brain.

Chapter 1 reviews literature on image-guided surgery, which is the context for this thesis research, with a particular emphasis on accuracy issues. Chapter 2 surveys the scientific literature which deals with the central problem of this thesis: registration of anatomical surfaces. This chapter is self-contained and presents no new method developed during this thesis research. As such, this chapter can be skipped by the reader whose sole interest is in the original methods presented in this thesis. The rest of this thesis addresses the details of my method for computing the brain surface displacement. Chapter 3 examines how laser range-sensing can provide raw data about the evolving shape of the brain in the operating room. Moreover, chapters 3 and 4 discuss a method for relating this intraoperative surface shape information to preoperative information about the outer brain surface. Part of the registration solution is given in the chapter on range-sensing because there are two registration problems, and it makes sense to discuss the procedure for solving the first one in the context of discussing the acquisition of intraoperative surface shape information. The two registration problems are:

- *Relating range data to MR data.* Establishing the baseline for brain motion in the range referential requires two conditions: finding this baseline, or identifying the outer cortex surface, in MR (this is discussed in chapters 5, 6 and 7), and transposing this baseline to range space. Fulfilling the second condition constitutes the first registration problem of interest.
- *Non-rigid registration of cortical range images.* Registering a time sequence of range images of the brain, to compute a surface displacement map over time. This is the subject of the discussion of chapter 4.

The first registration problem could have been solved by formally registering a skin patch imaged by the range-sensor, with its anatomical equivalent in the MR volume, but as the survey in chapter 2 attests, registering an open featureless skin surface with its MR homologue is a very difficult problem. Registering a near-closed surface encompassing the whole cranium, in particular anchored by facial features, would have been a lot easier [136], but I couldn't assume that the whole head would

be visible in the OR. In practice it is mostly covered by surgical material. A transformation based on a calibration procedure and on the probe-based tracking of the range-sensor base position, which makes use of the image-patient space transformation, is simpler and in all likelihood more robust than a fully autonomous registration procedure, in the absence of discriminant surface shape features, as suggested by the survey in chapter 2.

Finally, the identification of the cortex in the MR volume is laid out over three chapters because the problem of model-based surface identification (or segmentation) is highly complex in itself. In chapter 5, I provide both a literature survey of prior art, and an intuitive discussion of (and motivation for) the so-called surface evolution model which is adopted and refined in my research. Chapter 6 features a discussion of my contributions to making surface evolution models better adapted for identifying the brain surface, in particular for a brain featuring pathologies. Chapter 7 discusses the validation of the various stages of my model-based brain surface identification. Next, I draw conclusions and offer a synthesis and possible continuations of this research.

A literature survey of biomechanical models of the brain is included in appendix A. This review traces the history of the literature from early dynamic models of neurological damage caused by automobile collisions, to recent quasi-static models emphasizing surgical loads. It also emphasizes how physical modelling of the brain relates to accurate image guided neurosurgery and to neurosurgical simulation. In appendix B, the mathematical derivation of the surface evolution model is included.

Original Contributions

The following are believed to be significant original contributions made by this thesis:

1. A framework for characterizing intra-surgical brain deformations, based on non-rigid surface registration.
2. The application of a laser range-sensor to characterize intra-surgical brain shift.
3. The integration of model-based anatomical surface segmentation and registration in a unified framework whereby the registration benefits from a post-processing of the segmentation results, in a manner which is adapted from the surface model computation. In particular, the Closest Point Map obtained by the Fast Marching method makes each iteration of the Iterative Closest Point registration essentially instantaneous.
4. Refinements to surface evolution models for identifying the surface of the brain, particularly one which features imbedded pathologies:
 - A first refinement consists in the use of a classifier to define the spatially varying speed function, which is used to halt the evolving model on the brain boundary. This refinement also entails the recruitment of classifier training points to initialize the surface model, based on small spheres centered at their position.
 - A second refinement is a 2-stage approach which first identifies the white matter (WM) boundary, and uses it to *flexibly* limit the expansion of the surface model through grey matter (GM) voxels deemed close to the former boundary. Existing coupled-surface methods use a constant coupling distance between WM and GM boundaries, which is too restrictive in areas of high GM thickness, such as the cerebellum, and which is inapplicable in the presence of a pathology imbedded in WM.
5. A probe-based calibration and range-sensor tracking procedure for computing the rigid transformation between the internal coordinate systems of the range-

sensor and of the MRI or CT volumetric image data. This work includes the design and machining of calibration tools and the automatic detection of milled pits within range data by template matching, resulting in a fully automatic procedure.

6. A brain-shaped elastic phantom, whose deformations are triggered from within, for duplicating intra-surgical brain shift. This phantom is based on the use of a brain-shaped mold intended for gelatin, but used with Poly-Vinyl Alcohol Cryogel, in conjunction with a movable imbedded disk.
7. Other adaptations for producing usable phantoms:
 - A head mannequin fitted with adhesive surgical fiducials, and converted into a cranial skin phantom for the validation of the range-MR transformation computation procedure.
 - Adaptations of an existing digital anthropomorphic phantom: its brain surface was identified explicitly, and used to quantify various stages of the model-based surface segmentation, as they relate to the brain surface. It was also re-sampled to produce synthetic range images of the brain, to which was applied a known, analytical deformation function, for non-rigid registration validation.
8. A published literature review on anatomical surface registration.
9. A literature review of physical modeling of brain, emphasizing image guidance and surgical simulation applications, to be submitted for publication.

Contributions of Authors

The literature survey appearing in chapter 2 was published by the author in Medical Image Analysis [14]. The co-authors F.P. Ferrie and T.M. Peters contributed as follows. F.P. Ferrie suggested some readings, while T.M. Peters provided supervision and proof-read the final paper.

The literature survey appearing in appendix A is a manuscript in preparation, to be submitted for publication. The co-authors contributed as follows. M. Miga suggested some readings on the physiological basis for poroelastic models. J. Nemes proof-read the manuscript and pointed out inconsistencies in the notation. K. Chinzei provided insight on rheological experiments. T.M. Peters provided supervision and proof-read the final manuscript.

Publications

Various aspects of this thesis research have appeared in the following publications:

1. M.A. Audette et al., An Integrated Range-sensing, Segmentation and Registration Framework for the Characterization of Intra-surgical Brain Deformations in Image-Guided Surgery, submitted to *Comp. Vis. and Image Understand.-Special Issue on Non-rigid Registration*, 2002.
2. M.A. Audette, F.P. Ferrie & T.M. Peters, An Algorithmic Overview of Surface Registration Techniques for Medical Imaging, *Medical Image Analysis*, Vol. 4, No. 3, pp. 201-217, Sept. 2000, selected for re-publication in IMIA Yearbook of Medical Informatics, Schattauer Publ. Co., 2002.
3. M.A. Audette, K. Siddiqi, & T.M. Peters, Level-set Surface Segmentation and Fast Cortical Range Image Tracking for Computing Intrasurgical Deformations, *Medical Image Computing and Computer-Assisted Intervention (MICCAI99)*, Cambridge, England, pp. 788-797, 1999.
4. M.A. Audette & T.M. Peters, Level-set Surface Segmentation and Registration for Computing Intrasurgical Deformations, *SPIE Medical Imaging 1999*, San Diego, Feb. 20-26, 1999.

This thesis work also led to the following oral presentations, the abstracts of which were published:

1. M.A. Audette and T.M. Peters, Level-Set Surface Segmentation and Registration for Computing Intrasurgical Deformations, *Canadian Organization of Medical Physicists-COMP Conference Proceedings*, London, Ont., pp. 134-135, June 1998.
2. M.A. Audette and T.M. Peters, Brain Surface Curvature Computation and Tissue Modelling for Accurate Image-guided Surgery and Surgical Simulation, *Canadian Organization of Medical Physicists-COMP Conference Proceedings*, Charlottetown, P.E.I., pp. 215-216, July 1997.

The following are U.S. patents in my name relating to image guided surgery, while collaborating with ISG Technologies, Inc.:

1. Pat. no. 5951571, Method and apparatus for correlating a body with an image of a body, Sept. 14, 1999.
2. Pat. no. 6096050, Method and apparatus for correlating a body with an image of a body, Aug. 1, 2000.

Some of the work begun in this thesis is still underway, and will be documented in the following publications:

1. M.A. Audette et al., A Survey of Biomechanical Modeling of the Brain for Intra-surgical Displacement Estimation and Medical Simulation, *manuscript in preparation*.
2. M.A. Audette & K. Chinzei, Global Structure-preserving Voxel Classification with Limited Supervision for Highly Realistic, Patient-specific Surgical Simulation, *manuscript in preparation*.

Glossary

The acronyms on the left column appear in this thesis, and their meaning is outlined in the right column.

2D/3D	two/three dimensions
CAD	computer-aided design
CCD	charge-coupled device
CSF	cortico-spinal fluid
CT	computed tomography
FE(M)	finite element (modeling)
FLE	fiducial localization error
FM	Fast Marching
FRE	fiducial registration error
GM	grey matter
ICBM	International Consortium for Brain Mapping
ICP	Iterative Closest Point (registration)
(I)IGS	(interactive) image-guided surgery
MNI	Montreal Neurological Institute
MR(I)	magnetic resonance (imaging)
OR	operating room
RMS	root-mean-square
SNR	signal-to-noise ratio
TRE	target registration error
WDF	WM/GM Distance Function
WM	white matter

Chapter 1

Introduction

1.1 Image-guided Surgery

Image-guided surgery describes a class of techniques for localizing anatomical structures on the basis of image data and for determining the optimal surgical path to reach these structures. IGS is particularly prevalent in neurosurgery, which is justified foremost by the functional importance and fragility of tissue surrounding the target and by the added requirement of limiting the extent of the craniotomy [122], given the protective nature of the skull. Intracranial localization constitutes a major challenge in neurosurgery, and a precise sense of complex three-dimensional relationships is required at all times for a successful intervention [233]. Traditionally, the art of surgical navigation has been taught in a manner where specific anatomical landmarks are recognized and used as beacons to guide the navigation. However, just as modern ground- and satellite-based navigational aids have revolutionized airflight, obviating the need for visual beacons and making it possible for aircraft to routinely land in foul weather and in difficult terrain, IGS has similarly revolutionized the spatial orientation of surgeons.

With the advent of tomographic image volumes produced by computed tomography (CT) and magnetic resonance (MR), it has become possible to elaborate a 3D image model of the patient, whose internal coordinate system (“*image space*”) can be related to that of a *localization device* spanning the physical space occupied by the

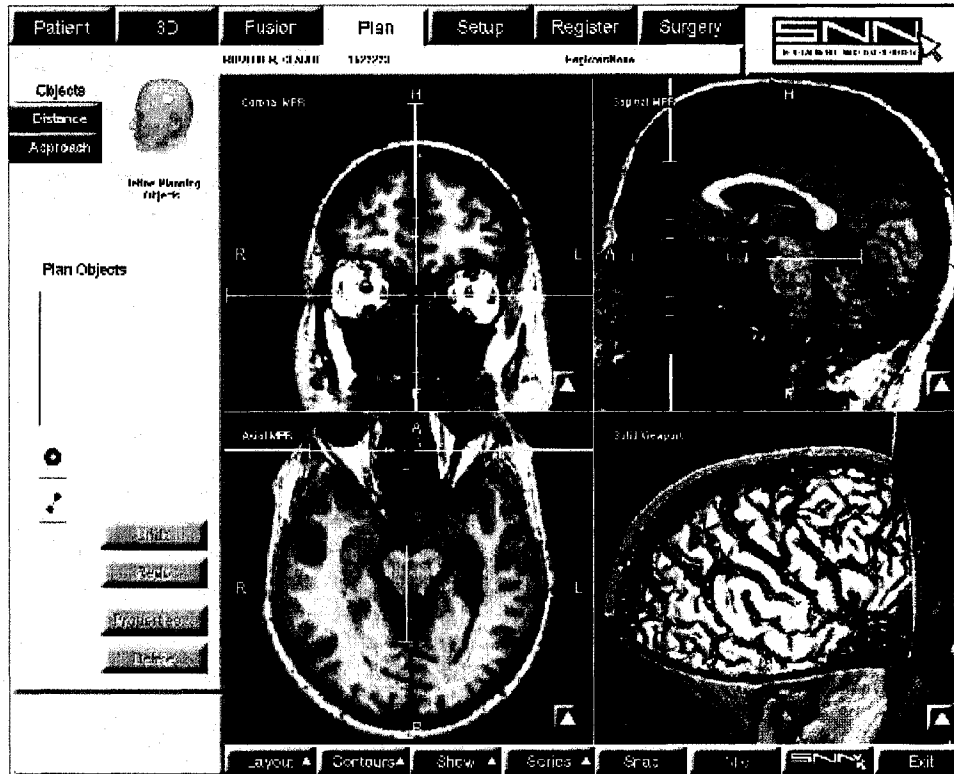


Figure 1.1 : *Image-guided surgery: 3D image model of patient as presented in typical commercial IGS system, namely in tri-planar and 3D view formats, courtesy of Lahbib Soualmi of the MNI.*

patient (“*patient space*”). In other words, anatomical structures can be identified and localized explicitly in the context of surgical planning, inside or outside the operating room (OR), on the basis of morphological information in MR magnetization or CT attenuation properties. The identified structures in the image model can then be related to their physical position in space (within the patient) in the OR, once the transformation between the two spaces has been computed in the course of a registration procedure. Brain IGS applications include the treatment of movement and convulsive disorders and of arteriovenous malformations as well as the biopsy, surgical resection, radiotherapy and brachytherapy of tumours (for surveys see [232, 337, 128]). Other important applications include endoscopic sinus surgery [272], guidance of pedicle screw fixation in spine surgery [286], breast biopsy needle insertion [282], orthopedic treatment of complex fractures [51] and stereotactic liver surgery [150].

The information comprising the image model typically consists of the tomographic volume itself, displayed in tri-planar view format, as illustrated in figure 1.1, and may include anatomical boundary surfaces identified from the tomographic scan and rendered onto a computer screen¹. In the event that the image model includes anatomical surfaces, such as the skin, outer brain boundary, and tumour boundaries, the inner surfaces can be made visible either by exploiting translucency or by making virtual cuts in the outer surfaces. Moreover, as demonstrated in this chapter, it is also possible to co-register other kinds of image data with the reference tomographic volume, such as relevant vascular or functional information, to refine the optimal surgical path.

The position of the point of interest within the 3D image model is conveyed on the basis of:

- the choice of tri-planar view of the underlying tomographic data and the position of the cursor overlaid on each plane, as well as
- the position of a rendered virtual pointer overlaid at the appropriate position and orientation on the displayed anatomical surfaces.

The tri-planar format describes 3 orthogonal planes whose normals are generally along the transverse, anteroposterior and longitudinal (x , y and z) directions, where each plane coincides with an image slice within the volume selected for viewing. For a brain scan, the image planes correspond to *sagittal*, *coronal* and *axial* cross-sections. These cross-sections are typically displayed as separate 2D rasters [363, 46, 365], as in figure 1.1, but may also be projected onto orthogonal planes integrated within a 3D view [374]. To display each 2D raster, one coordinate of the pointer selects the slice of interest, while the other two coordinates appear as a 2D cursor superimposed on the image. Alternately, the tomographic volume can also be displayed as a set of planes along a user-defined orientation, such as those defined about a virtual pointing device, rather than along preset axes. Finally, the explicit patient-image space transformation

¹These anatomical surfaces are generally approximated by a collection of triangles which makes possible their efficient visualization [115, 338], based on synthetic light sources and virtual reflectance properties.

of one tomographic volume can lead to the reconstruction (or resampling) based on the former's sampling orientation, of a second tomographic modality in 3D (e.g.: PET) or in 2D (ultrasound), or to the projection of the tomographic volume onto viewing planes consistent with a second, projective modality (e.g.: DSA). A synopsis of multimodality fusion is provided in section 1.2.

There are two broad categories of localization technique, *frame-based*², shown in figure 1.2, and *frameless*, illustrated in figure 1.3, with the former chronologically preceding and serving as a basis for the latter. The first approach consists of a frame rigidly attached to the patient's skull for the duration of both the pre-operative scan and surgery. The frame features an intrinsic geometrical means for explicitly determining the image-patient space transformation, as well as a structural base to support surgical instruments. In contrast, frameless stereotaxy is generally characterized by a handheld, tracked pointing device, consisting of a *probe* and perhaps a *probe-holder*, whose tracking base is not affixed to the patient³, thereby impeding surgery far less. Moreover, in the frameless case, the landmarks for determining the image-patient transformation are extrinsic to the localization device; i.e.: fiducials glued to the scalp or imbedded in the cranium, or well-localized anatomical features, as opposed to the intrinsic landmarks of the stereotaxic frame.

Finally, and more importantly, frameless localization is far more interactive and easier to use than frame-based, and is also referred to as *interactive image guided surgery* [122] (IIGS) as well as *neuronavigation* [413, 337]. In practice, there is generally no real-time tracking procedure providing a frame-based system with image space feedback about the position of surgical tools or pointing devices on the frame. With the frameless approach, assuming the geometry relating the probe tip to the probe holder is known, and following a registration procedure relating image space to the physical space spanned by the tracked probe, the anatomical position of the probe tip

²Also known as *stereotaxy*, from Greek: *stereo-* three dimensional or solid body, and *taxis-* arrangement. Since the mid-1970's however, the term *stereotactic* surgery is also used to describe surgery on humans, meaning three-dimensional and touching, from Latin [128].

³Although it may be affixed to the patient bed or to the Mayfield clamp which immobilizes the patient head.

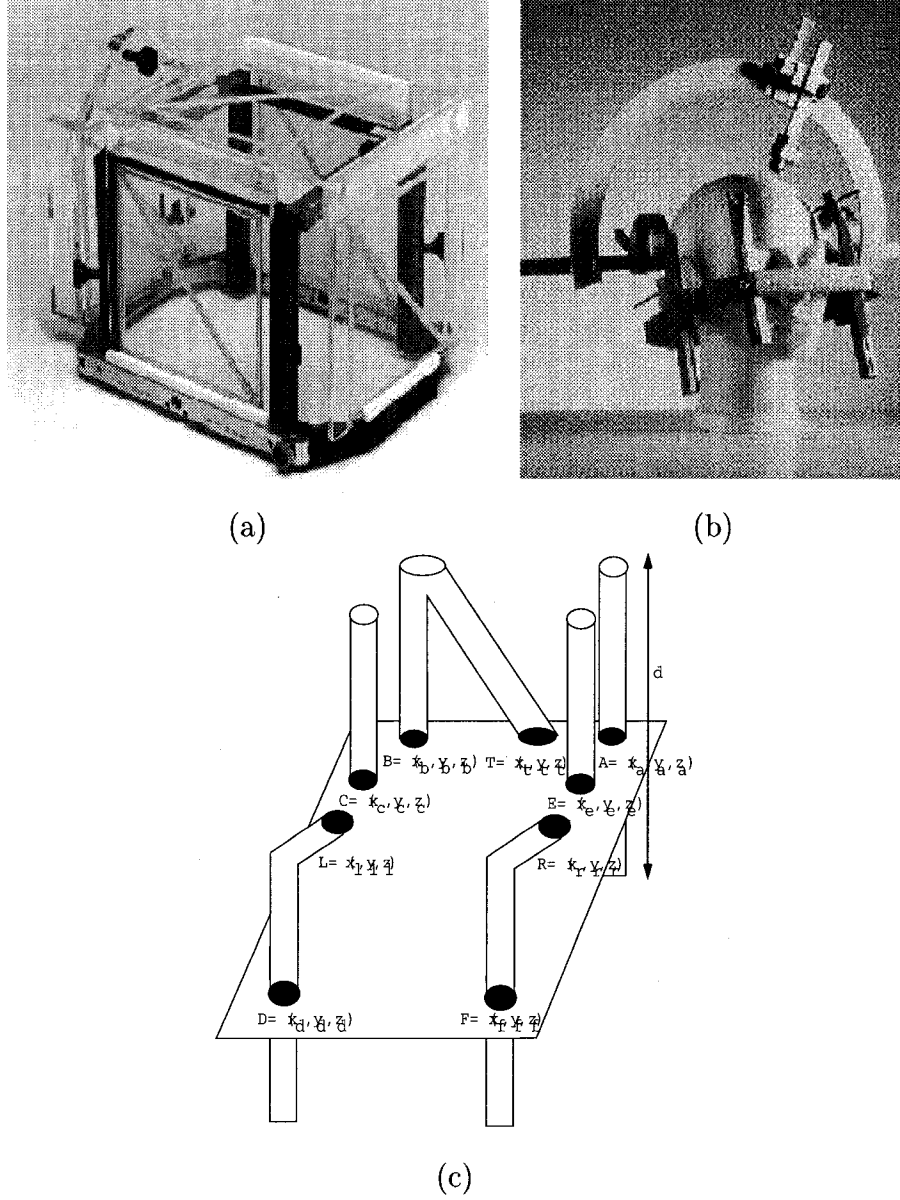


Figure 1.2 : *Frame-based localization: (a) the Olivier-Bertrand-Tipal (OBT) frame, (b) mounting arc for biopsy needles, (c) geometry of frame-based N-shaped localizer (adapted, with permission, from [423]).*

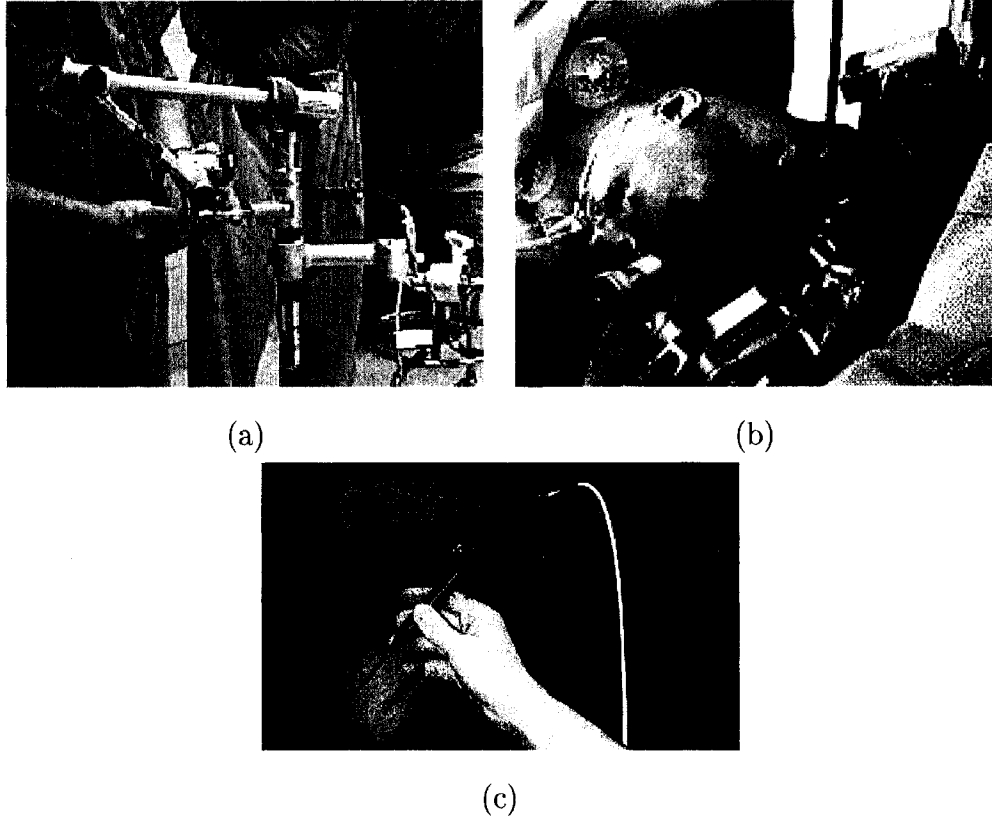


Figure 1.3 : *Frameless localization devices: the FARO arm, during (a) probe calibration and (b) intraoperative use; (c) the Northern Digital Polaris optical tracking system.*

is reflected by that of the 3D image cursor at all times. As a result, the elaboration of the best surgical path is highly intuitive [337]. This intuitiveness may be essential to aggressively resecting a tumour for example, because the tumour margins may have well defined borders in MR or CT, particularly with a contrast agent, but may be visually indistinguishable from healthy tissue during the course of surgery [122].

Nevertheless, the frame-based approach is worth discussing because of its historical and clinical relevance as a predecessor to frameless localization⁴. The modern stereotactic frame is the descendant of, and in many ways similar to, devices introduced roughly a century ago. In 1889, Zernov introduced an “encephalometer”, consisting of an aluminum ring, affixed to the head by ear plugs and plates applied to the upper edges of the orbits, and second ring perpendicular to it [178]. Its purpose

⁴For a survey of IGS emphasizing stereotaxic frames, see [128].

was to predict the *surface topography* of the brain (cranial sutures and cerebral sulci), but not to direct the insertion of stereotactic surgery electrodes. In 1908 Horsley and Clarke invented the first true stereotactic device (used to plan surgery in 3D): a rectilinear frame which integrated image information in the form of an anatomical atlas, for targeting deep brain structures of cats and monkeys [157]. The first stereotactic frame intended for surgery in humans is ascribed to Mussen in 1918, influenced by the Horsley and Clarke design, but he failed to convince a surgeon to use it [312]. It was intended to carry an electrode to any deep structure in order to stimulate it or destroy a tumour present (for a discussion, see also [337]).

At this point, the imaging modality in use was the anatomical atlas based on skull landmarks, which suffered from two drawbacks: first, it didn't account for individual anatomical variability (justifying surgeons' reluctance to use Mussen's frame), and second, it could not compensate for distortions of the anatomy caused by pathologies. By 1945, x-ray studies allowed *intracerebral* landmarks to be seen, leading Spiegel and Wycis to devise a human atlas based on these landmarks, later published as a textbook in 1952 [369]. Spiegel and Wycis also devised a rectilinear stereotaxic apparatus [367], the stereo encephalotome, which made use of intraoperative measurements of cerebral landmarks outlined on a pneumoencephalogram. During this time, comparable work was carried out by Leksell [214] and by Talairach [384, 385], respectively on an arc-based frame and on a brain atlas which also used cerebral landmarks.

The advent of tomographic modalities in the 1970's and 80's led to the modern stereotactic frames, such as that introduced by Brown [50], which integrate a means of relating a point in the image volume to one in the physical space spanned by the frame. These frames feature two or more N-shaped localizers, each of which consists of 2 parallel rods, each perpendicular to the base ring, and a third rod running diagonally from one parallel rod to the other. The localizers are used to determine the position and orientation of an image slice by triangulation. Briefly, figure 1.2(c) provides an illustration of a typical frame geometry, featuring 3 N-bars. In this case the forward relationship mapping a 3D point $\mathbf{x} = (x, y, z)$ to a 2D image point $\mathbf{x}' = (x', y')$ can be ascertained from the image plane positions of the fiducials [423] x'_C, x'_D, x'_E, x'_F ,

assuming an axial image plane:

$$\begin{aligned}
x' &= x'_D + (x'_F - x'_D)s_x + (x'_C - x'_D)s_y + (x'_D + x'_E - x'_C - x'_F)s_xs_y \\
y' &= y'_D + (y'_F - y'_D)s_x + (y'_C - y'_D)s_y + (y'_D + y'_E - y'_C - y'_F)s_xs_y \quad \text{where} \\
s_x &= (x - x_C)/(x_E - x_C) \quad \text{and} \quad s_y = (y - y_D)/(y_E - y_D) .
\end{aligned} \tag{1.1}$$

The inverse relationship is obtained by manipulating the equations 1.1 in order to express x and y in terms of x' , y' , x'_C , x'_D , x'_E , and x'_F (see [423] for full details). Furthermore, the depth z can be computed on the basis of the 3 fiducials coinciding with the intersections of the oblique bars with the image plane. The depth z_T of the fiducial T is given by

$$z_T = d \frac{AT}{AB} . \tag{1.2}$$

The 3 N-bars in the Brown-Roberts-Wells (BRW) [50] frame provide just enough points [423] to determine the so-called *slice transformation matrix* [213] (STM), while the OBT [291] frame, featuring 4 N-bars, enables an overdetermined least-squares solution [213] of the STM: the z coordinate is then evaluated from the slice equation at (x, y) . Some frames that used only two markers required that the frame be mounted such that the longitudinal axis of the frame was collinear with the axis of the scanner. Lastly, the consideration of the fiducials in all slices at once yields a vastly overdetermined system for computing the *volume transformation matrix* (VTM) relating the image model to the patient.

Image-guided surgery made a significant advance in 1986, with the introduction by Kelly et al. [181] of the first interactive stereotactic device, a laser integrated within a *surgical microscope* linked with a computer. This device featured surgical planning by outlining the margins of the tumour on each CT slice and by reformatting the tomographic volume orthogonally to the intended surgical path. The laser focus point served as a pointing device, and the position of both the microscope and the laser focus point were tracked by optical encoders. The surgeon controlled the position of the laser in the surgical field with a joystick. By using the laser in continuous mode (at 65-85 W), the tumour was vapourized “slice by slice”, as the focus point

was tracked over time (see also [182, 183]). A similar surgical microscope has since been commercialized by Zeiss [426]. The drawbacks of the stereotactic device based on the surgical microscope are its complexity and cost. Moreover, an accuracy study conducted on the Zeiss MKM model [176] concluded that measurement of the focus point is sensitive to focusing strategy and to the angle between the focus plane and target surface.

Flexible, low-cost frameless devices soon appeared thereafter, in the form of *articulated arms* [413, 122, 432, 337] and *optical tracking systems* [287, 421, 184]. Other technologies based on the time-of-flight of sound [322] and the estimation of magnetic gradients induced in the workspace [387] have also been explored, but were later shown to be sensitive to echoes and to magnetic interference respectively [337]. Only articulated arms and optical tracking systems are in widespread use, and are further discussed here.

Two early versions of the neuronavigation arm⁵ are attributed to Watanabe [413] and to Galloway [122]. The FARO Surgicom was the earliest commercial frameless device and was integrated with the ISG Viewing Wand [432, 95], the first interactive image guidance system certified for clinical use by the US Food and Drug Administration. Some articulated arms can also be fitted with surgical instruments at the end of the wrist. For example, the Milano [129] arm features fail-safe brakes and is designed to be fitted with a biopsy needle, a tissue retractor and an endoscope. Arm-based stereotaxy computes the probe holder position by sensing each joint angle with a potentiometer or an optical encoder, and by relating the angles and arm segment dimensions to Cartesian coordinates. This approach is comparable to the computation of a robotic end effector position by forward kinematics [82]. The precise relationship between joint angles and cartesian position is ascertained by a calibration procedure performed by the manufacturer. The transformation from the probe holder to probe tip depends on the length of the particular probe, and is determined by another calibration procedure which is typically available in the IGS system and

⁵Schiffbauer [337] has exhaustively compared 7 early designs, along with the recent Oulu Neuronavigator system.

run by the user.

Optical tracking systems, in contrast, are based on cameras seeing and triangulating the position of highly characteristic landmarks. These systems can be categorized as either *active* or *passive*. In the former case, the landmarks are light-emitting diodes (LEDs), usually imbedded in the probe handle, that emit infrared light one at a time, and a precalibrated multiple sensor system determines their position based on this signal. In the latter case, the *sensors* emit infrared light and triangulate from characteristic reflections of this signal from spherical reflecting landmarks. In fact, the characteristic emitters or reflectors can be affixed to the handle of a surgical instrument [421] or to a dynamic reference frame (DRF) [397, 184] attached to the Mayfield clamp, for the purpose of tracking patient movement during brain surgery (not to be confused with intra-surgical brain movement)⁶. Commercial versions of both types of optical systems are made by Northern Digital Inc. [287], and active optical systems are also produced by Image Guided Technologies Inc. [166]. Other related work based on machine vision include Heilbrun’s application of a 2-camera photogrammetric projection system which is pre-calibrated with a localizer featuring 8 fiducials [143], and Kikinis’ registration of laser-based range skin data with the reconstructed skin surface in a MR scan [188], for facilitating the subsequent registration with intrasurgical video.

1.2 Multimodality Fusion

One of the most important recent developments in interactive image-guided surgery is the integration of complimentary modalities with standard tomographic data to alleviate the diagnostic limitations of MR and CT. First, the use of angiographic data such as magnetic resonance angiography (MRA) or digital subtraction angiography (DSA) allows the superposition of explicit vascular structures over morphological data. Second, the consideration of patient-specific eloquent areas of the brain becomes

⁶The DRF is also routinely used to track the piecewise-rigid motion of the vertebrae considered in the context of spine surgery.

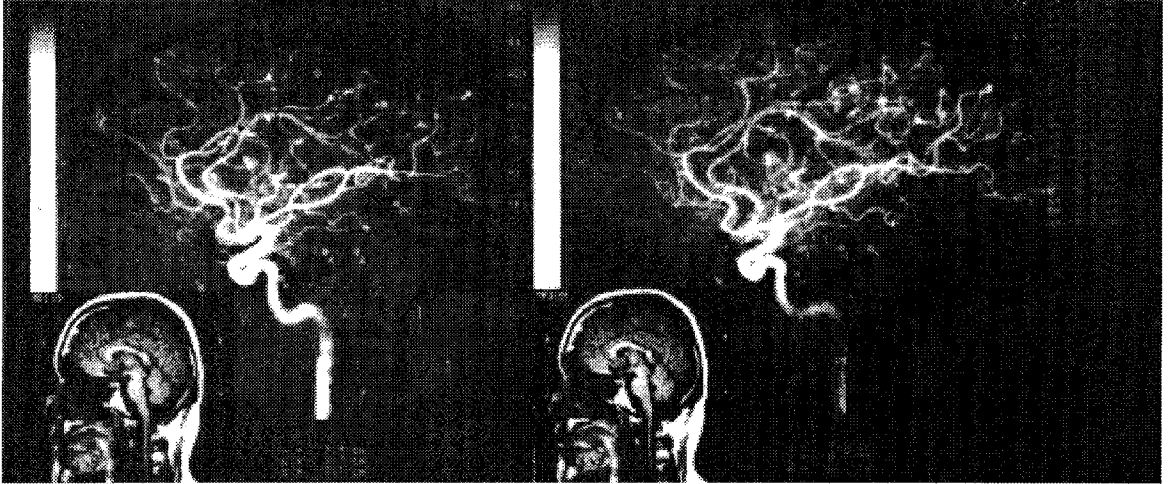


Figure 1.4 : *Integration of MRI with stereoscopic DSA image pair [149].*

important for a surgical intervention assumed to be in proximity of these areas, which may impose the requirement of co-registered functional data such as PET or SPECT, particularly for an aggressive resection. Next, given that conventional MR is limited in the information it gives on the histological nature of pathology, the integration of spatial spectroscopic distributions provides the surgeon with greater discrimination ability, to better gauge where to be most aggressive in resection. Finally, the desire to have the image model remain current during an intervention or to consider data from the viewpoint of a surgical instrument, justifies the integration of intrasurgical modalities.

First of all, the importance of considering vasculature in neurosurgery [307, 294], and in particular in interventions assumed to be close to important veins or arteries, may justify the integration of angiographic data in IGS. Henri et al [149, 309] discussed the relative merits of displaying a MRA volume, co-registered with MRI data and volume- or surface-rendered, versus integrating one or two DSA images with MRI by generating a volume-rendered image of the latter which matches the DSA projective geometry and enables the two modalities to be superimposed, as shown in figure 1.4. The former option exploits MR tagging of flowing blood and is referred to as a MR inflow angiography technique [149]. The latter option is described in detail in [309] and uses 8 frame-based fiducial markers to determine a 2D image to 3D frame

homogeneous transformation matrix:

$$[\alpha u \ \alpha v \ \alpha] = [x \ y \ z \ 1] \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \\ m_{41} & m_{42} & m_{43} \end{bmatrix}, \quad (1.3)$$

where α is a scaling factor. The twelve $\mathbf{M}_{proj} = [m_{ij}]$ parameters which determine each projection are solved for, and once the IGS frame-image volume transformation $\mathbf{T}_{IGS}$ is known, the combination of $\mathbf{M}_{proj}$ and $\mathbf{T}_{IGS}$ is sufficient to specify the projection and overlay of tomographic data onto the DSA viewing plane(s) (see also [217]). Another promising avenue is the extraction of a vascular tree from standard MR images with tube-shaped active surface models [296].

The consideration of functional data stems from a need to take into account the precise spatial distribution of eloquent areas near the site of the pathology of interest. The method of Zhang et al. for PET to MR/CT registration exploits intrinsic frame landmarks [423], based on filling N-bar rods with a solution of germanium-68 for PET, with copper sulfate solution for MR, and using solid aluminum bars for CT. Zhang proceeds from the forward frame-to-image relation of expressions 1.1 and 1.2 to determine the reverse image-to-frame relations for all modalities, which can then be combined to relate any two, such as PET to MR.

Frameless methods for registering anatomical volumes to functional data such as PET, SPECT, and recently, a 3D functional atlas [112], include those based on extrinsic markers as well as those driven by anatomical data, as reviewed by Maurer [243] and by Maintz [236]. These methods exploit landmarks, surfaces, principal axes, atlas-based delineation, correlation and recently, a measurement of similarity from information theory. The landmark approach [242, 310] is an interactive method where anatomically relevant points are expertly identified in each volume of interest, and the resulting homologous point pairs determine a transformation computation, which is closed form for the rigid-body case (see chapter 2). Surface-based methods [303] correspond to the iterative point-based surface registration approaches described in chapter 2, exploiting a measure of proximity to determine homologous

surface point pairs. Principal axis techniques [104, 7] rely on the computation of the voxel-based moments of inertia of the volumes of interest. Methods based on this notion assume that two volumes of similar shape should have similar principal moments of inertia and can be registered by aligning the corresponding principal axes. The atlas-based approach [217] begins by establishing the line between anterior and posterior commissures [55] (AC-PC line) within the sagittal mid-plane of a brain volume, and establishes a local 3D coordinate frame whose origin is the AC point, and whose x , y and z axes coincide respectively with the AC-PC line, with the normal to the sagittal plane, and with the normal to x within the sagittal mid-plane [386]. The axes of this coordinate frame, plus a second vertical line arising from the PC, span the planes which divide the brain into 12 sub-volumes, leading to an anatomically relevant piecewise-affine transformation. The correlation methods [20, 174] rely on a measure of intensity similarity integrated over the domain of interest, under the assumption that the ideal registration should maximize this integral. Finally, the information-theoretic methods [418, 72, 416] refine this integrative approach. The optimal pose between two volumes maximizes the mutual information [301] between two volumes, or rather it iteratively minimizes the dispersion of the joint intensity distribution computed for the two putatively aligned volumes.

As alluded to earlier, one of the foremost applications of IGS is the treatment of brain tumours. An excellent survey of neurosurgical brain tumour treatment appears in Schiffbauer [337], from the standpoint of the histology of tumours, their associated neurological deficits, and the physical properties affecting their imaging. Limitations of anatomical MR include the fact that in general it does not provide a precise histological diagnosis, which would alleviate the need for biopsy, and that not all pathologies are characterized by noticeable changes (without contrast enhancement) in T1 and T2 relaxation (see also [43]). These limitations of conventional MR provide the impetus for integrating magnetic resonance spectroscopy (MRS) information in IGS, as proposed by Preul et al. [315]. MRS imaging is based on the existence of six major chemical resonances associated with metabolites, in the T2-weighted MR spectrum of brain tumours: choline-containing phospholipids (Cho), creatine and

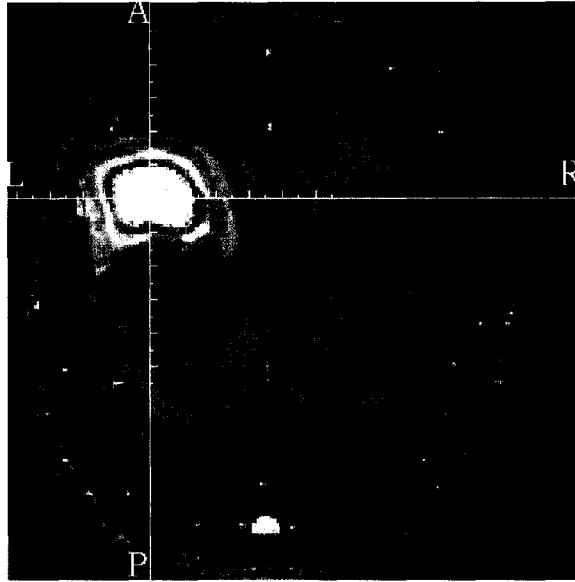


Figure 1.5 : *An illustration of co-registered MRS and MRI data: superimposed choline (colour) and T1-weighted (gray scale) axial images, courtesy of Mark Preul [315].*

phosphocreatine (Cr), N-acetyl groups (NAA), Alanine (Ala), Lactate (La) and finally lipids (Lip). Preul recently demonstrated that a very accurate prediction of the five most common brain tumours is possible, by pattern recognition, from the highly characteristic spectral pattern of each tumour type, and as shown in figure 1.5, has integrated this information in an IGS platform.

Lastly, an increasingly prevalent area of research of multimodal image fusion is the integration of intrasurgical data, in order to provide a surgical instrument-centered view or to update the image model to account for intrasurgical evolution of the anatomy, the latter issue being discussed at length in the next section. First, the integration of a tomographic scanner in the surgical suite begins with Lunsford's dedicated therapeutic CT scanner [227, 197], which was used to confirm the completeness of tumour resection. Interventional MRI (iMRI) was pioneered by GE in the mid-1990's [36, 124], and soon also commercialized by other manufacturers [311, 46]. Samset [333] and Nimsky [283] have recently presented clinical evaluations based on several cases. In the former study, the IGS registration exploited the optical sensors integrated with the iMRI scanner, so that the surgical and image spaces are iden-

tical and a characterization of brain shift accurate to within 1 mm was achievable. Intraoperative ultrasound (US) offers a low-cost alternative for gathering intraoperative tomographic data (typically in 2D, however), and is growing in its clinical acceptance. Comeau [76] and Maurer [249] have devised elegant calibration procedures for relating MR/CT image space to US space, thereby enabling the interactive characterization of brain shift in IGS. Fusion of intraoperative video with a tomographic image model was described by Kikinis [188], and discussed at length by Clarkson et al. [68], the latter within a mutual information maximization framework. Finally, Uecker et al. have demonstrated that the integration of laparoscopic video to IGS [402], in particular to track surgical instruments and to reposition a robotic laparoscope, can provide the surgeon with an instrument-centered viewpoint. Given the growing emphasis on minimal invasiveness, this work also emphasizes the rich potential of IGS to empower the surgeon and to revolutionize neurosurgery.

1.3 Accuracy Issues in Image-guided Surgery

The usefulness of image-guided surgery hinges on the accuracy of the transformation between the image and patient spaces. This accuracy is dependent on a number of conditions, which fall into two groups as follows. Violation of any of the listed assumptions adversely impacts the accuracy of the system.

- *An accurate initial registration:* this assumption requires that
 - the tracking equipment be accurate,
 - the tomographic data be distortion-free,
 - the position of the frame or fiducials used to anchor the registration with respect to the patient be maintained consistent from the scan to the OR,
 - and that these fiducials be accurately localized by the user within the scan and on the patient.
- *A static relationship between patient and image model,* which in turn requires that

- once the patient is registered with the image model, the spatial relationship between the patient and the stereotaxic device remain consistent,
- and that the anatomy being targeted remain static throughout.

In practice, many of these conditions are not met, and the impact of these issues constitutes the object of numerous studies since the integration of tomographic data in stereotaxy. The accuracy achievable with commercial stereotaxic frames [121] and each kind of frameless device [176, 326, 184] based on phantom or calibration data studies, is well-documented. Rohling et al. [326] tested both the FARO arm and the Northern Digital Optotrak optical tracker, for their accuracy in determining the spacing between divots in a precision-machined gold standard. They found that the former estimates distances within 0.5 mm, with a spread of 0.5 mm, and that the Optotrak is generally within the manufacturer’s specifications of 0.15 mm, with a spread equivalent to that value. More recently, Khadem et al. [184] evaluated optical trackers, by attaching a dynamic reference frame on a precision-machined plate, itself mounted on a linear positioner, allowing a fully volumetric study. They defined accuracy in terms of a “jitter” measurement $\mathbf{J} = \{J_x, J_y, J_z\}$, the standard deviation of 100 measured Euclidean distances at a point. After eliminating the worst 5%, Khadem measured active optical trackers at values ranging between 0.028 ± 0.012 mm to 0.059 ± 0.047 mm, and the sole passive model tested at 0.115 ± 0.075 mm. A similar study for the Zeiss MKM surgical microscope [176] found its accuracy to be sensitive to the angle between focus plane and target surface, as well as to focusing strategy. Its absolute error varied between 0.25 ± 0.09 mm and 0.38 ± 0.17 mm.

Accuracy statistics seem to worsen, however, for a registration procedure which involves points in tomographic data and their homologues in physical space, particularly if the positions of the landmarks are not known a priori. Galloway et al. compared the performances of 4 stereotactic frames in a CT study of a plexiglass phantom featuring posts holding radio-opaque targets [121]. At an image resolution of 1mm, the accuracy of the frames, varied from 1.0 ± 0.6 mm to 1.9 ± 1.0 mm, depending on the model. A study conducted by Zinreich et al. with the FARO arm [432]

on a similar phantom assessed the registration error as 1.12 ± 0.61 mm. However, the registration error for plastic skull trials in the same study, where phantom geometry is not known prior to its measurement, was given as 1.74 ± 0.89 mm, with a worse case error of 4.88 mm. The difference is attributable to the uncertainty in identifying fiducials in the image model, which has been called *fiducial localization error* (FLE) by some authors. Research by Maurer and Fitzpatrick et al. [247, 248, 113] has sought to model the relationship between FLE and *target registration error* (TRE) as well as *fiducial registration error* (FRE), which is discussed in further detail in chapter 3. A recent study by Alp et al. [3] produced FRE values for different types of fiducials:

- frame-based, used as a fixation device (Zamorano Dujovny frame): 0.69 mm,
- bone markers (Leibinger OST-REG fiducials [211]): 0.89 mm,
- skin- ball-type (Beekley X-spot [27]), for CT: 1.5 mm, and
- skin- torus-shaped (IZI Medical Products [169]), for MR: 2.62 mm.

The differences in FRE stem from variations in FLE related to the shape of each fiducial and to slippage associated with each type. In related clinical research, Sipos [357] and Doshi [93] reported mean accuracy values of 2.51 mm and 1.8 mm respectively with the ISG Viewing Wand.

An important related issue is the geometric distortion in tomographic volumes, or the accuracy of the representation of the anatomy by the image volume. Michiels et al. [257] classified the sources of geometric distortion for Fourier-based MR reconstruction (see also [289]) into three groups:

- *Machine and environment-dependent parameters*, including magnetic inhomogeneity, imperfect slice or volume selection pulses, nonlinear magnetic field gradients and eddy currents due to gradient switching.
- *Patient-dependent parameters*, in particular magnetic susceptibility variations and chemical shift. Chemical shift describes the shielding effect of a nucleus by its surrounding chemical structure, whereby a proton experiences a local field

$B_l = B_o(1 - \sigma)$. In contrast to the former, magnetic susceptibility χ is a macroscopic property describing the ease by which a material can be magnetized. For a linear isotropic medium the magnetization vector $\mathbf{M}$ is proportional to the magnetic field strength $\mathbf{H}$: $\mathbf{M} = \mu_o\chi\mathbf{H}$, but as susceptibility varies locally, so does the value of the magnetic field, given by $\mathbf{B} = \mu_o\mathbf{H} + \mathbf{M}$, resulting in distortion.

- *Stereotactic frame* (where applicable), in particular due to the choice of materials used. This is an exacerbation of distortion due to chemical shift and magnetic susceptibility variations, related to the rods comprising the frame and to the choice of contrast agent within the rods.

In related work, Sumanaweera [378] exploited measurements along the more accurate phase-encoding directions to quantify the accuracy of the measurements along the frequency encoding directions without prior knowledge of the geometry of the imaged object. Munger [274] adopted an inverse problem framework for characterizing the phase error in echo planar images (EPI). Zylka and Wischmann [434] considered the sources of mechanical (e.g.: gantry tilt angle) misalignment of a CT scanner, and accounted for patient-dependent effects, such as bending of the CT table due to the patient’s weight, in order to model and characterize geometric distortion in CT.

Finally, the most important source of error in IGS is the violation of the assumption of static anatomy, namely intrasurgical brain shift. The amount of intrasurgical brain shift, as documented by numerous authors [292, 154, 245, 246, 324, 283] often measures more than 10 mm and can exceed 25 mm. Olivier et al. [292] were amongst the first to discuss brain shift in IGS, which they ascribed to “the progressive sinking and displacement of the cortical surface as a result of CSF drainage and progressive mesial displacement of the brain due to gravity”. Maurer and Hill et al. characterized intraoperative brain shift using sparse localizer points [154, 246] and later, intraoperative MRI [245]. They suggested possible causes of brain shift, many of which are a consequence of the common practice of reducing intracranial pressure prior to carrying out neurosurgical procedures, but the most important of which is the

loss of cortico-spinal fluid (CSF). Steroids are administered preoperatively to limit inflammation. Intraoperatively, cerebral blood volume is controlled by manipulating ventilation and venous drainage. Moreover, the water content of the brain is controlled through administration of mannitol. Finally and most importantly, with the opening of the dura, CSF is lost, and the brain, which is somewhat buoyant in CSF, progressively sags due to the effect of gravity. Hill [154] has also proposed two simple geometric models for the phenomenon:

- a Uniform Shift Model whereby the volume change of the brain which is essentially *ventricular*, resulting in a deformation uniformly distributed over the brain volume and in a uniformly radial cortical shift, and
- a Fluid Shift Model where the volume change is concentrated near the *craniotomy*, as if the cranial contents were fluid draining through the base.

Roberts et al. [324] recently performed a detailed quantitative analysis of brain shift based on 28 operative cases, including several tumour and seizure focus resections, within the frontal, temporal and parieto-occipital lobes and the posterior fossa. They attempted to characterize statistical relationships between the magnitude of displacement and surgery type, finding for example that the shift is greater for hippocampectomies than the other surgeries considered, and that it is lesser for electrode placements than other surgeries. Most importantly, Roberts also found *intraoperative movement of the brain to be oriented principally along the downward axis regardless of head position, consistent with gravity playing a dominant role in the settling of the brain due to CSF drainage.*

Existing methods for characterizing brain shift in general exploit either a combination of sparse cortical surface displacement data and realistic volumetric finite element models, or a combination of volumetric displacement information from tomographic data (intraoperative ultrasound or MRI) with some form of volumetric interpolation. Several authors have proposed realistic finite element (FE) [261, 266, 110] or mass-spring models [358] models of the brain. Physical models of the brain have been used to compute volumetric brain displacement on the basis of *sparse* cortical surface local-

izer data [246, 259], and recently, from the motion of a surface model coinciding with the brain boundary within an intraoperative MRI volume [110]. Of these methods, Ferrant’s devoted the most attention to meshing issues (the subdivision of the brain volume into simple geometric elements), automatically producing a tetrahedral mesh from the raw MRI volume, where the outer boundary of this mesh coincides with the outer brain surface [110]. Other authors either segmented the volume manually prior to running a commercial meshing algorithm, assumed voxel-sized elements [137] (which is impractical in 3D), or used isotropic elements whose outer surface extends beyond the brain and cranium [99]. Appendix A provides a detailed discussion of physical modelling of the brain. Peters et al. [76, 130] have also demonstrated the computation of volumetric brain displacement from manually identified landmarks in intraoperative ultrasound, using local volumetric interpolation.

1.4 A New Approach to Estimating Brain Shift

In this thesis, I propose a method for characterizing brain shift that is based on non-rigid surface registration, and can be combined with a constitutively realistic finite element approach for volumetric displacement estimation [259, 358, 110]. As illustrated in figure 1.6, this registration method integrates in a unified framework all of the stages required to estimate the movement of the cortical surface in the operating room (OR):

- semi-automatic identification (segmentation) of relevant anatomical surfaces within the MRI volume;
- range-sensing of the exposed brain surface in the OR;
- computation of the range-MRI transformation, based on a calibration procedure and sensor base tracking; and
- non-rigid motion tracking, over time, of the range image of the brain.

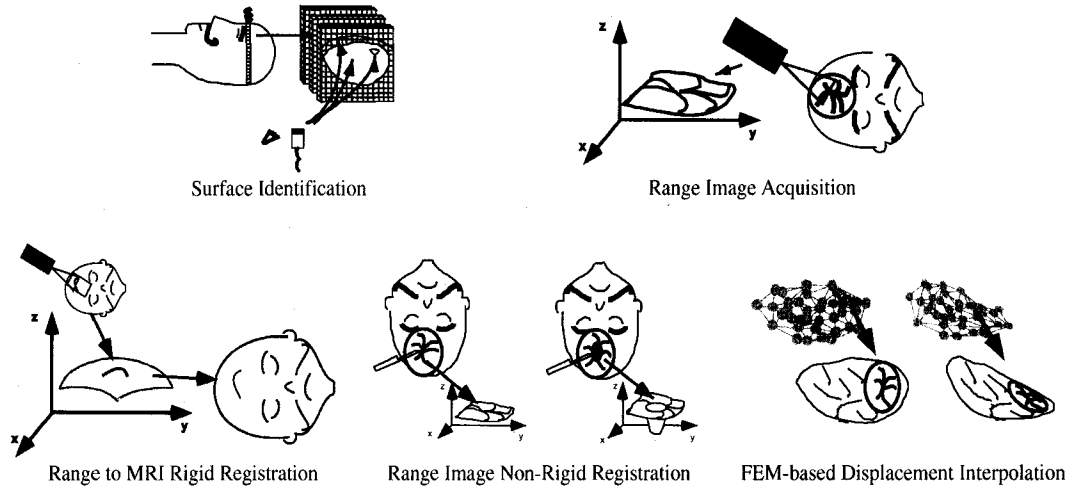


Figure 1.6 : *Brain shift estimation framework, featuring the four stages comprising our non-rigid surface registration framework, which can be combined with a FE-based volumetric displacement estimation stage.*

My approach is complementary to existing finite element models used to estimate intraoperative brain movement, alluded to above, and surveyed in detail in Appendix A.

Current work with FE models tends to incorporate sparse outer cortex displacement information, typically involving probe points taken at locations on a grid imbedded on the brain surface. I propose a method for quickly and easily acquiring and processing *dense* displacement information, using a sensor which could easily be deployed to and retracted from a scanning position, with the appropriate mounting setup. As seen in chapter 2, dense surface displacement information has advantages over sparse data, in characterizing nonrigid motion which is inherently underdetermined. Moreover, this problem is approached with an emphasis on speed of acquisition and of processing in the OR, in order to optimize the *temporal* determination of nonrigid motion. Moreover, existing research involving realistic finite element models is either generic or achieves patient specificity through a laborious manual segmentation stage, whereas the semi-automatic surface identification stage presented here

can be explicitly integrated with meshing software to make the finite element model *patient-specific* with little user interaction. My technique provides a means of semi-automatically identifying all relevant anatomical surfaces within a pathological brain. A final element of justification for this framework lies in its commercial potential, a range-sensor being much cheaper than both ultrasound and intraoperative MR, as well as being very easy to use. Finally, this framework does not preclude the integration of separate resection information, provided by a tracked hand-held probe [356] for example.

Chapter 2

A Survey of Surface Registration Techniques for Medical Imaging

2.1 Introduction

The registration of 3D surfaces is dealt with extensively in machine vision and medical imaging literature¹. Its applications vary from building terrain maps, in the context of providing autonomy to a planetary rover [142], and depth maps of a sea floor for oceanographic studies [177], to the recognition of objects from a computed-aided design (CAD) database [105], and of course, to reconciling various imaging modalities in biomedical imaging [71]. This chapter, based on a published review paper of mine [14], provides a detailed overview of surface registration techniques which have been, or could potentially be, applied to anatomical surfaces. This problem is a subset of the general medical image registration problem, as surveyed recently by Maintz and Viergever [236], who also discussed landmark and volume registration, but I emphasize algorithmic details, with a view to providing some motivation for each technique.

¹This literature survey has been published recently by the author in Medical Image Analysis [14]. This chapter discusses surface models which emphasize registration, while the treatment of model-based surface identification is deferred to chapter 5. The co-authors F.P. Ferrie and T.M. Peters each suggested some readings and proof-read the final paper.

Surfaces provide more information than landmarks, and this descriptiveness may be particularly advantageous for characterizing non-rigid motion. Moreover, we can make a distinction between landmarks automatically extracted from surfaces, which can be seen as feature point-based surface registration and are included in this survey, and manually identified landmarks, which may be tedious to determine and less repeatable than the former. Furthermore, a surface-based approach can still work if the two modalities of interest cover parts of the anatomy which overlap only partially, for example if one modality represents a small subset of the anatomy which appears in the second modality. In other words, we can usually register a subpatch with a larger surface patch, on the basis of local and global surface shape, as will be seen later. Finally, in medical imaging literature, anatomical surfaces are usually explicitly identified within tomographic data such as MRI and CT and are often closed [151, 401]. I expand on this definition by including range images of anatomical structures, such as those obtained by laser-based triangulation, which have a particular relevance to image-guided surgery [354, 188] and which are typically open.

Registration, between modalities A and B , is the estimation of a mapping between coordinate systems Ref_A and Ref_B associated with each modality:

$$\mathbf{x}_B = \mathbf{T}(\mathbf{x}_A), \quad (2.1)$$

where $\mathbf{x}_A = (x_A, y_A, z_A)$ and $\mathbf{x}_B = (x_B, y_B, z_B)$ are points in coordinate systems Ref_A and Ref_B respectively which correspond to the same anatomical point, and where the quality of this mapping can be quantified by a global measurement based on fitting residuals. In an ideal, noise- and distortion-free environment where the same anatomy is imaged by two modalities of like scale, the computation of the transformation from point pairs matched on the basis local information would produce a relation which is also globally consistent. In other words, the resulting transformation would exactly align all pairs of homologous points. In practice, the data contain noise and distortion, and the anatomy itself may distort between images. Therefore the optimal relation (especially if a rigid transformation assumption is maintained) is that which reconciles local homologous point alignment and global consistency in some optimal

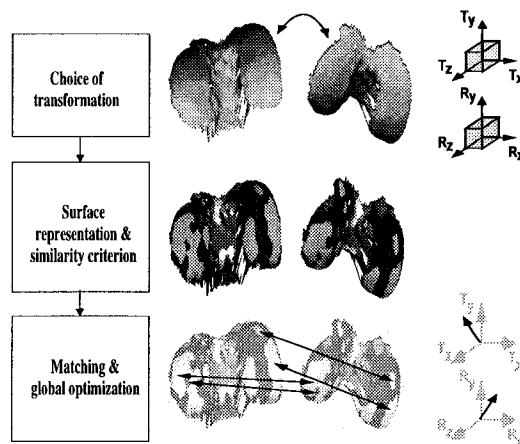


Figure 2.1 : *3D surface registration framework, featuring the choice of transformation between two anatomical surfaces (and of the search space for transformation parameters); the representation of the surface in terms of matching primitives (e.g. regions of consistent curvature); and finally, the matching of surface primitives and computation of the “best” transformation.*

manner.

Surface registration can be roughly partitioned into three stages, as illustrated in figure 2.1: *choice of transformation*, *elaboration of surface representation and similarity criterion*, and *matching and global optimization*. The first issue concerns the assumptions made about the nature of relationships between the two modalities, e.g. whether a rigid-body assumption applies, and if not, what type and how general a relation optimally maps one modality onto the other. The second issue determines what type of information we extract from the 3D surfaces, which typically characterize their local or global shape, and how we organize this information into a representation of the surface which will lead to improved efficiency and robustness in the last stage. The last issue pertains to how we exploit this information to estimate the transformation which best aligns local primitives in a globally consistent manner or which maximizes a measure of the similarity in global shape of two surfaces. Within this framework, this chapter discusses each surface registration issue in detail and reviews the state-of-the-art among existing techniques.

2.2 Choice of Transformation

The first stage is the formalization of the assumptions about the type of relation T that is appropriate for mapping points $\mathbf{x}_A$ on a 3D surface onto points $\mathbf{x}_B$ on another. In most registration problems, T is a transformation between the same anatomy imaged either by different modalities or by one modality at different times. In this context, a *rigid-body transformation* is applicable provided that the deformations sustained by the anatomy are negligible compared with the required accuracy of the transformation. If the deformations between surfaces are significant, and especially if these deformations are caused by factors other than noise and distortion within the modality, then a *nonrigid transformation* must apply. Moreover, one can further classify nonrigid transformations based on whether they are specified by a global or piecewise local fitting.

2.2.1 Rigid-body Transformation

A general rigid-body transformation can be expressed as combination of a rotation and a translation:

$$\mathbf{x}_B = \mathbf{R}_{AB}\mathbf{x}_A + \mathbf{t}_{AB}. \quad (2.2)$$

Consequently, rigid-body registration typically seeks the values of $\mathbf{R}$ and $\mathbf{t}$ which minimize

$$\min_{\mathbf{R}, \mathbf{t}} \sum_{i=1}^N \|\mathbf{x}_{B_i} - (\mathbf{R}\mathbf{x}_{A_i} + \mathbf{t})\|^2, \quad (2.3)$$

given 3D point correspondences $\mathbf{x}_{A_i}$ and $\mathbf{x}_{B_i}$. The problem can be reformulated in such a manner that the computation of $\mathbf{t}$ is decoupled from that of $\mathbf{R}$ by referring the coordinates to the respective centroids of each point set. This leads to the minimization

$$\min_{\mathbf{R}, \mathbf{t}} \sum_{i=1}^N \|\mathbf{x}'_{B_i} - (\mathbf{R}\mathbf{x}'_{A_i})\|^2 \quad \text{where} \quad \begin{aligned} \mathbf{x}'_{A_i} &= \mathbf{x}_{A_i} - \frac{1}{N} \sum_{j=1}^N \mathbf{x}_{A_j} & \text{and} \\ \mathbf{x}'_{B_i} &= \mathbf{x}_{B_i} - \frac{1}{N} \sum_{j=1}^N \mathbf{x}_{B_j} \end{aligned} \quad (2.4)$$

The translation is given by the difference of centroids:

$$\mathbf{t} = \frac{1}{N} \sum_{j=1}^N \mathbf{x}_{B_j} - \mathbf{R} \frac{1}{N} \sum_{j=1}^N \mathbf{x}_{A_j}. \quad (2.5)$$

I review four common representations for rotation, as well as a fifth model which represents both translation and rotation. The relevance of choosing one particular representation is that it may lead to more efficient and/or numerically stable estimation of its parameters than others, or may be better suited to a particular surface representation. It is worth noting that these techniques are applicable not only to the registration of surfaces, but also to any set of explicit point pairs. Furthermore, while they may not have been published in a medical imaging context, they are still applicable to anatomical data.

The *orthonormal matrix* representation consists of a 3×3 matrix, which can be viewed as a mapping from reference frame A to frame B, once the translation between their origins is compensated, where each element R_{ij} is a direction cosine (i.e. the projection of one axis of reference frame A onto one axis of reference frame B [82]). Arun et al. [8] obtained rotation by first computing the singular value decomposition [314] of the matrix

$$\mathbf{H} = \sum_{i=1}^N \mathbf{x}'_{B_i} \mathbf{x}'_{A_i}{}^t \quad (2.6)$$

determined from centroid-referred coordinates:

$$\mathbf{H} = \mathbf{U} \mathbf{D} \mathbf{V}^t, \quad (2.7)$$

where $\mathbf{D}$ is diagonal and $\mathbf{U}$ and $\mathbf{V}$ are orthonormal. The rotation is given by the expression $\mathbf{R} = \mathbf{V} \mathbf{U}^t$.

I briefly address two interesting representations, the *Euler angles* and *axis-angle* models, although I emphasize less the techniques that employ them, because they are iterative rather than closed-form. We can express rotation as the product of three successive rotations (γ, β, α) of a predefined fixed coordinate system about axes $\hat{x}$, $\hat{y}$ and $\hat{z}$, or equivalently, as a succession of rotations about $\hat{z}$, $\hat{y}$ and $\hat{x}$ moving axes [82]. Huang et al. [161] used this Euler angle representation to design a 3D iterative motion

estimation scheme that is a sequence of well-behaved 2D minimizations involving the projections of (partially rotated) points on the $x-z$, $y-z$ and $x-y$ planes. Moreover, a rotation can also be completely specified by a unique vector whose direction is the rotation axis and whose norm is the rotation angle about this axis [19]. Lin et al. [220] adopted the axis-angle representation for a Fourier space approach to rigid-body motion estimation that does not require explicit correspondences.

A widely used representation of rotation is based on *quaternions* [106, 158]. A quaternion can be thought of as a generalization of a complex number, with a real part and three imaginary parts, or as a composite of a 3-vector in $\mathbb{R}^3$ and a scalar in $\mathbb{R}$. Moreover, the rotation quaternion can also be interpreted in terms of the axis-angle representation by the *Euler Symmetric Parameters* [408]: $\mathbf{q} = [\sin \theta/2 \mathbf{n}, \cos \theta/2]^t$. In other words, the orientation of the 3-vector component specifies the axis of rotation, and the norm of the 3-vector and the scalar component are related to the rotation angle about this axis. Horn cast the search for the optimal rotation parameters as a maximization based on quaternion components (in contrast to the minimization of Faugeras and Hebert [106]). His objective function was optimized with respect to rotation by finding the eigenvector corresponding to the largest positive eigenvalue of a matrix $\mathbf{N}$ (see [158]) determined from centroid-referred point coordinates.

The motivation for the *dual quaternion*² rigid transformation estimation technique of Walker et al. [408] is that other rigid transformation estimation techniques first determine optimal orientation and then use this solution to obtain the translation (e.g. [8]), resulting in the accumulation of error in this computation. The dual quaternion technique solves for both relative orientation and position by minimizing a single cost function. The underlying model views the transformation between two coordinate frames as a translation of the original coordinate frame along a direction $\mathbf{n}$ by a distance d , followed by a rotation by an angle θ with respect to a line having $\mathbf{n}$ as its direction and passing through a point $\mathbf{p}$, as illustrated in figure 2.2. Walker reported similar accuracy to Arun’s SVD technique for estimating rotation, but improved

²A *dual number* $\hat{a} = a + \varepsilon b$ is defined as a combination of two real numbers a and b , with a special multiplication rule given for ε given by $\varepsilon^2 = 0$, such that, e.g. $(a + \varepsilon b)(c + \varepsilon d) = ac + \varepsilon(ad + bc)$.

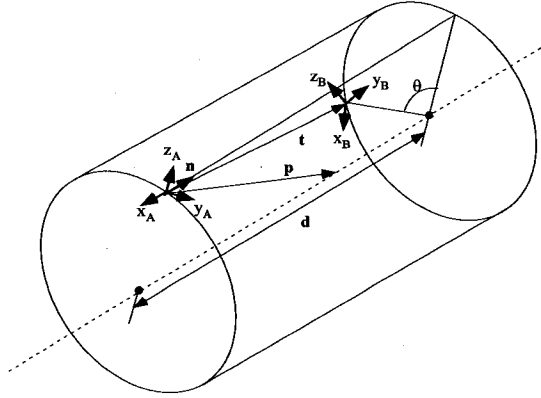


Figure 2.2 : *Illustration of the dual quaternion model for rigid transformation [408] (reproduced with the permission of R.A. Volz, copyright Academic Press 1991).*

accuracy for estimating translation, across identical sets of point correspondences.

2.2.2 Nonrigid Transformations

In the event that surface deformations must be accounted for in the transformation, a nonrigid approach must be adopted. The simplest classes of nonrigid transformations are basically generalizations of the rigid body transformation. A more general formulation is one where a global polynomial function maps surface A to surface B . To make the mapping even more general, a global polynomial can be replaced by piecewise polynomial relations, or *splines*, and such functions can be seen as either *interpolating*, whereby the transformation they express is directly determined by the motion between two sets of primitives (and as such is sensitive to errors in the data), or *approximating*, whereby the agreement between the local motion of any pair of matches is reconciled with global consistency, in some optimal manner.

Generalization of Rigid Body Motion: The Affine Transformation

The generalization of the rigid body transformation relevant to 3D surface matching is the *affine* transformation. The general *affine* class of transformations is characterized by the expression

$$\mathbf{x}_B = \mathbf{A}_{3 \times 3} \mathbf{x}_A + \mathbf{b}_{3 \times 1}, \quad (2.8)$$

where there is no orthogonality constraint on the elements a_{ij} of matrix $\mathbf{A}$ as there is in equation 2.2. Affine transformations do not in general preserve angles or lengths, but parallel lines remain parallel [115]. The affine transform relating two 3D surfaces is solved by Feldmar and Ayache [107], whose method matched “closest” points of compatible shape. This ensured that the solution to the minimization $\min_{\mathbf{A}, \mathbf{b}} \frac{1}{N} \sum_{i=1}^N \|\mathbf{A} \mathbf{x}_{A_i} + \mathbf{b} - \mathbf{x}_{B_i}\|^2$ tended to a stable solution. Moreover, Henri et al. [148] determined the translation, rotation and nonuniform scaling which maps stereotactic frame coordinates to corresponding CT or MR voxel values, prior to superimposing stereoscopic DSA images on the equivalent projections of the volumetric scan.

Global Polynomial Functions

A more general formulation is one where a global polynomial function, typically of order 2 to 5, maps surface A to surface B [208]. Global methods use matched point pairs to generate a single optimal transformation, based on a sufficient number of points to (over)determine the parameters of the transformation, via either approximation or interpolation. Polynomial transformations are typically expressed in either Cartesian or spherical coordinates. Moreover, global polynomial transformations are only useful to account for low-frequency distortions because of their unpredictable behaviour when the degree of the polynomial is high [49]. This component of inter-surface motion is typically computed after an initial rigid alignment.

Approximation is the search for transformation parameters which map matched points or features as closely as possible globally, but not necessarily perfectly individually, whereas interpolation finds the transformation which maps two 3D surfaces

so that matched control points are exactly satisfied. In the former case we assume that some noise or unwanted distortion exists and should not be accounted for in the transformation. For large numbers of control points, this choice makes sense because matches likely include inaccuracies, but taken together they contain sufficient statistical data to make the transformation reliable. Interpolation is more appropriate for a few accurately-matched control points, since it involves an independent parameter for each control point match, which may result in unexpected undulation for a high-order fit.

For a Cartesian formulation, the global polynomial transformation in three dimensions can be stated as follows:

$$x_B = \sum_{ijk} a_{ijk} x_A^i y_A^j z_A^k \quad y_B = \sum_{ijk} b_{ijk} x_A^i y_A^j z_A^k \quad z_B = \sum_{ijk} c_{ijk} x_A^i y_A^j z_A^k, \quad (2.9)$$

where a_{ijk} , b_{ijk} and c_{ijk} are the constant polynomial coefficients to be determined. If interpolation is used, these coefficients express a system of $3N$ unknowns which can be determined by N control points. In a least squares approximation, the sum over all matched feature/point pairs of the squared difference between the left and right-hand side of these equations is minimized, for example by setting the partial derivatives associated with these equations to zero.

Jacq and Roux [170] implemented a trilinear interpolation (i.e. where the summation indices i , j and k each go from 0 to 1, and where the $i+j+k \leq 3$), determined by 8 reference distortion values which spanned the volume to be warped. Subsol et al. [377] registered skulls and cortical surfaces to build a 3D atlas, by first rigidly matching crestlines and then using an iterative closest point algorithm (both techniques are discussed in § 2.3) on crestline points to determine a global 2nd order polynomial to describe the relation between iteratively transformed surface A and surface B. This resulted in a polynomial transformation of order 2^n , based on n iterations.

For closed surfaces that can be modelled as functions on a sphere, some authors prefer to work in spherical coordinates [79, 62]. For example, *spherical harmonic* surfaces are closed surfaces on a sphere that can be decomposed into a set of orthogonal functions. To represent an arbitrary shape, the radius $r(\theta, \phi)$ in the spherical coor-

dinate system (centered on the centroid) can be written as a linear sum of spherical harmonic basis functions:

$$r(\theta, \phi) \approx \sum_{n=1}^N \sum_{m=0}^n [A_{nm} U_{nm}(\phi, \theta) + B_{nm} V_{nm}(\phi, \theta)], \quad (2.10)$$

where A_{nm} and B_{nm} are basis coefficients computed from the data points from a 3D object or surface and N represents the order of the fitting. The basis functions are:

$$\begin{aligned} U_{nm}(\theta, \phi) &= \cos(m\phi) P_{n,m}(\cos(\theta)) \quad \text{and} \\ V_{nm}(\theta, \phi) &= \sin(m\phi) P_{n,m}(\cos(\theta)) \end{aligned} \quad (2.11)$$

where $P_{n,m}(x) = (1-x^2)^{\frac{m}{2}} \frac{d^m}{dx^m} P_n(x)$ and $P_n(\cdot)$ is the Legendre polynomial of degree n [314]. In other words, surfaces S_A and S_B , and the displacement between them, can be seen as functions on a sphere, represented by a set of real coefficients A_{nm} and B_{nm} . Because the shape is approximated as a sum of different harmonics, in theory this representation can reconstruct high-frequency surface detail.

Coppini et al. [79] modelled the epicardial stretch tensor, based on tracked vascular bifurcations, by performing a 3rd-order spherical harmonic fitting over individual displacements $r_A(\theta_i, \phi_i) - r_B(\theta_i, \phi_i)$, after correction for translation and rotation. Chen et al. [62] adopted a similar method, but first characterized global shape with *superquadric* surfaces [23, 24, 21], prior to characterizing local shape variation with spherical harmonics.

Local Nonrigid Transformations: Piecewise Polynomials

Global mapping functions do not always adequately capture deformations of anatomical structures, which are often intrinsically local [208, 39]. Piecewise polynomial functions produce a more general relation. In general, the relative density and reliability of the data determine whether an interpolation or an approximation scheme is used.

One well-documented interpolating scheme is the *thin-plate spline* [96, 39]. The thin-plate spline over a 2D domain can be expressed as $z(x, y) = U(r) = r^2 \log r^2$, where $r = \sqrt{x^2 + y^2}$, and as such U is a fundamental solution of the biharmonic

equation. The interpolant $f(\mathbf{p})$ is optimal in that it has minimum bending energy amongst all functions which pass exactly through points $\mathbf{x}_i = (\mathbf{p}_i, z(\mathbf{p}_i))$; i.e.

$$\int \int_{\mathbb{R}} [(\frac{\partial^2 f}{\partial x^2})^2 + (\frac{\partial^2 f}{\partial x \partial y})^2 + (\frac{\partial^2 f}{\partial y^2})^2] dx dy \quad (2.12)$$

is minimized. This type of interpolation function can then be applied to modelling nonrigid motion. Instead of having $f(\mathbf{p})$ represent a displacement in the z -direction over the $\mathbf{p} = (x, y)$ domain, it can express the x -component of a deformation. Likewise functions $g(\mathbf{p})$ and $h(\mathbf{p})$ can express the y - and z -components of a deformation [39, 102], where now $\mathbf{p} = (x, y, z)$. This method presupposes the use of relatively sparse point-matches in order to determine a set of local polynomials or spline functions exactly, when in fact it may be useful to consider denser displacement information to describe a fundamentally underdetermined relationship. Moreover, the notion of minimum bending is better suited to deformation over a 2D, rather than 3D, domain.

Another useful piecewise polynomial representation is the *B-spline*, particularly in the context of smoothing (regularizing) and least-squares spline approximation [91]. Given a set of displacement data $\mathbf{d}(x_r, y_r, z_r)$, and three sets of knots $\lambda_i, i = 0 \cdots f+1$, $\mu_j, j = 0 \cdots g+1$, and $\nu_k, k = 0 \cdots h+1$, we can compute the least-squares p th order volumetric spline

$$\begin{aligned} \mathbf{u}(x, y, z) &= \sum_{i,j,k} \mathbf{u}_{i,j,k} B_i(x) B_j(y) B_k(z) \quad \text{where, e.g.} \\ B_i(x) &= \sum_{l=i}^{i+p+1} \frac{(\lambda_l - x)^p}{\lambda_l - \lambda_i} && \text{if } x \in [\lambda_i, \lambda_{i+1}] \text{ and} \\ B_i(x) &= 0 && \text{elsewhere,} \end{aligned}$$

and such that $\delta = \sum_r (w_r (\|\mathbf{d}_r - \mathbf{u}(x_r, y_r, z_r)\|))^2$ is minimized. Here, the determination of B-spline coefficients of $\mathbf{u}_{i,j,k}$ is by the least-squares solution of an overdetermined linear system. For a smoothing spline approach, the problem is to find the function $\mathbf{u}_p(x, y, z)$ minimizing a smoothing norm which is a function of the B-spline coefficients, subject to $\delta < S$. Szeliski and Lavallée [382] modelled the nonrigid transformation between two anatomical surfaces as a first order spline in x, y and z , which is constrained by zeroth and first order stabilizers. This approach penalizes large variations of the spline coefficients, while also enforcing agreement with displacement data.

Lastly, a few other ways of characterizing nonrigid motion appear in the literature. Goldgof [131] and Mishra [267] classified the nonrigid motion of surfaces in terms of how it affects their mean and Gaussian curvature properties, namely as *rigid*, *isometric*, *homothetic*, *conformal* and *general nonrigid*³. Moreover, Feldmar and Ayache [108] determined locally affine transformations for individual surface points $(\mathbf{A}_{A,i}, \mathbf{b}_{A,i})$ by a weighted sum of the rigid transformation parameters of locally neighbouring matched surface points $(\mathbf{R}_{A,k}, \mathbf{t}_{A,k})$, where $k \neq i$.

2.3 Surface Representation and Similarity Criterion

The second stage of registration consists of computing a surface representation and defining a matching criterion based on it. In general, the surface representation should be stable over the two modalities or over the time sequence considered. It should afford a similarity criterion which is sufficiently discriminating to associate homologous points unambiguously and efficiently, if the application dictates that algorithmic performance is an issue. There are four approaches to representing a surface for the sake of registration: *feature*, *point*, and *model-based* methods as well as techniques based on *global similarity*. The definition of the similarity criterion is typically closely related to the choice of matching primitive. Furthermore, the criteria for selecting a particular primitive are application-specific, depending for example on whether the anatomy of interest is smooth, such as the cranium, or highly angular, such as a vertebra. Other factors which influence this choice include the size of the transformation to be computed (i.e. whether the two surfaces are separated by an arbitrarily large transformation or roughly aligned), and whether the transformation is rigid or non-rigid.

The feature-based method attempts to express surface morphology as a set of

³Conformal motion is characterized by proportionality of the coefficients of the first fundamental form [92]: $\frac{E_A}{E_B} = \frac{F_A}{F_B} = \frac{G_A}{G_B} = \eta(u, v)$. This function $\eta(u, v)$ becomes a constant over (u, v) for homothetic motion, and identity for isometric motion.

features which are extracted by a preprocessing step. Such features provide a compact description of the surface shape (at the expense of losing information), which is quantifiable by stable, discriminating scalar measurements. The similarity criterion is then an outgrowth of this feature characterization: it consists of a comparison of scalar measurements. The point and model-based methods do not attempt to reduce the surface representation to a more compact description, rather they use all or most of the points. Generally, for the point-based method, the primitive used is often the surface point itself, and the similarity criterion is a distance to be minimized between a pair of surface points. For model-based approaches, often an implicit criterion is used, such as an external force or halting condition driven by two sets of image data, with which an evolving deformable surface model must be reconciled. Finally, a new class of registration methods matches surfaces typically on the basis of their global similarity. While there are currently few anatomical applications, the relevance of these methods stems from the feasibility of precomputing a number of training views of a surface generated from a patient's tomographic data, and from the desirability of computing arbitrarily large transformations for smooth, relatively featureless surfaces.

The feature-based and global approaches are potentially more discriminating than point- or model-based techniques, and can therefore resolve a large motion or transformation. On the other hand, most point and model-based methods may be attractive in the case of small or iteratively estimated motion, because they exploit much of the available information, which is especially useful for estimating locally nonrigid transformations, as these are inherently underdetermined.

2.3.1 Feature-Based Representations

Feature-based matching is largely founded on the use of differential geometry to describe local surface shape [92, 32]. According to Bonnet's fundamental existence and uniqueness theorem, if two surfaces S_A and S_B possess equivalent fundamental forms I and II (or equivalent Gaussian and Mean curvatures K and H), then there exists

an appropriate translation and rotation such that S_A and S_B coincide exactly [32]. In general, feature-based matching is applied to computing rigid transformations.

Features used for surface registration fall into three categories: (sparse) *point features* (distinct from dense point-based schemes, also referred to as free-form surface registration), *curves* and *regions*. Point features are salient, well-localized, sparse loci of important geometric significance, such as extrema of curvature: local peaks, pits, saddle points where the two principal curvatures are most pronounced, or where K is at a local minimum or maximum. The second type of feature corresponds to contiguous lines or curves, consisting typically of differential structures such as ridges or boundaries between regions. Regions, in turn, are areas possessing some homogeneous characteristic, such as consistent curvature sign. Each feature in surface S_A can be matched with its homologue in S_B by first characterizing each feature in either surface by parameters expressing its respective topology, and looking for a compatible vector of parameters in the other surface. If there is more than one suitable candidate for a given feature, the match can be disambiguated by assuming that neighbouring features on the same surface which are matched one-to-one should undergo motion consistent with the ambiguous candidate. False candidates can then be eliminated on the basis of motion inconsistency.

Accuracy issues related to feature extraction include the use of the neighbourhood information around each surface point to stabilize the computation of its differential properties [334, 111] and the sensitivity to noise statistics of a given numerical method for estimating derivatives and surface curvature [114, 1, 329]. Finally, another relevant area of research is the application of recursive infinite impulse response filters [316], particularly near-Gaussian exponential filters [346, 87, 88] to speed up the smoothing stage prior to feature extraction [268, 393].

Point features can be matched on the basis of their intrinsic information, such as surface curvature values, as well as position relative to neighbouring point features. One such feature is the *extremum of principal curvatures*, computed by Thirion [393] within volumetric data by detecting the zero-crossings of two “extremality” functions. The first function is discussed later in this section, and was used to detect

ridge lines, or lines of locally maximal higher principal curvature, while the second function found a local extremum of the lesser principal curvature within each ridge line. The scalar measurements used in the matching included the type of extremal point, which depends on the sign of the principal curvatures, the response of the extremality functions, the principal curvature values, and the distances and orientations of vectors to neighbouring points. An alternate representation, used by Goldgof et al. [132] in terrain matching but nonetheless applicable to anatomical surfaces if they possess sharp corners, is the *extremum of Gaussian curvature* (K), detected on the basis of a threshold on the value of K .

A somewhat less compact way of representing surface shape is as a collection of *curves*. One particularly relevant curvilinear feature in medical images is the *ridge or crest line* (e.g. the trough of a sulcus or the maximum height of a gyrus). As pointed out by Maintz et al. [235], there are many ways to define a ridge, and consequently many ways to detect one. Monga et al. [268, 269] looked for a contiguous set of loci of a surface where the largest principal curvature κ_1 was locally maximal. These correspond to the zero-crossings of the extremality function $e_1 = \vec{\nabla}\kappa_1 \cdot \mathbf{t}_1$, where $\vec{\nabla}\kappa_1$ is the directional derivative of the largest principal curvature, and $\mathbf{t}_1$ is the principal direction corresponding to κ_1 .

Maintz et al. [235] proposed two operators of their own, which stem from the consideration of the gradient $\mathbf{w}$ of a smoothed surface and its right-handed normal $\mathbf{v}$ in the proximity of a ridge. The gradient at any non-ridge position points towards the ridge, but at a ridge position the gradient is aligned with the ridge. A consequence of this geometry is that the directional derivative along $\mathbf{v}$ of a surface is characterized by a highly concave profile at a ridge point in comparison with other points. Maintz then detected a ridge point either as a minimum of the *second derivative of the surface*, or as a maximum of the derivative of the *direction of the gradient*, along $\mathbf{v}$. An alternative to a ridge-based curve is the *distance contour* [318], which could be applied to anatomical data, and which is the set of points of constant distance from a highly salient point.

Guéziec and Ayache [136] presented an elegant technique for characterizing and

matching curves, which in turn is based on the curve matching algorithm of Kishon et al. [194]. Kishon addressed the problem of finding the longest matching subcurve appearing in two curves. He used the local shape signatures *curvature* $\kappa(s)$ and *torsion* $\tau(s)$, of smoothed curves⁴, which are both rotationally and translationally invariant. Each shape signature was represented as a *hash table* (see § 2.4), where entries are associated with pairs of curvature-torsion values (κ, τ) . The improvements introduced by Guézic and Ayache relate to the approximation of the curves, to the hash table and to the statistical analysis of various invariants for matching, particularly in directly determining a 3D transformation from homologous points on matched curves.

Regions constitute an even denser feature-based representation. They are typically characterized either by the homogeneity of their local surface shape or as being circumscribed by some boundary. One advantage is that the notion of neighbourhood between regions is natural [396], and leads to the characterization of a surface as an adjacency graph. Matching is then carried out as an exercise in finding the maximal *clique* of compatible subgraphs (see [319] and the references therein). Here, groups of regions are matched based on neighbourhood topology as well as local region characteristics. One definition of homogeneity used to segment surfaces into regions or surface patches is the *K/H sign combination* [32, 180]. Surfaces can be subdivided into patches according to surface type: elliptic and outwardly bulging ($K > 0/H < 0$), elliptic and inwardly bulging ($K > 0/H > 0$), parabolic and outwardly bulging ($K = 0/H < 0$), parabolic and inwardly bulging ($K = 0/H > 0$), planar ($K = 0/H = 0$) and hyperbolic or saddle shaped ($K < 0/H \neq 0$). In practice, a small nonzero threshold is used to determine sign. Figure 2.3 illustrates a *K/H* surface representation for a range image of a femur epiphysis.

⁴For a curve $\alpha : I = (a, b) \rightarrow \mathbb{R}^3$, parametrized by arclength s , *curvature* $\kappa(s)$ is the scalar $\|\alpha''(s)\|$. It is a measure of how rapidly the curve pulls away from the tangent line $\mathbf{t}(s)$. The direction of $\alpha''(s)$ is given by the unit *normal* $\mathbf{n}(s)$, and the unit vector $\mathbf{b}(s) = \mathbf{n}(s) \times \mathbf{t}(s)$ is called the *binormal* vector. *Torsion* $\tau(s)$ corresponds to $\|\mathbf{b}'(s)\|$, and it is a measure of how quickly the curve pulls away from the *osculating plane* at s spanned by $\mathbf{n}$ and $\mathbf{t}$ [92].

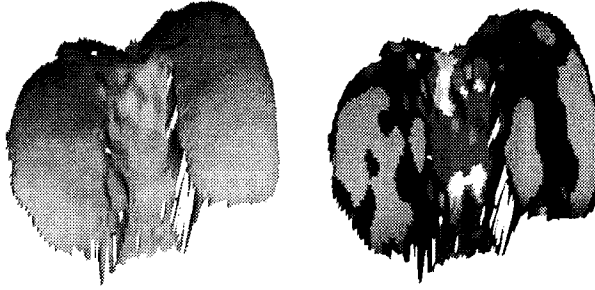


Figure 2.3 : (a) Range image of femur epiphysis; (b) regions based on K/H sign combination. Both are courtesy of Gilbert Soucy of the McGill Centre for Intelligent Machines.

2.3.2 Point-Based Methods

Point-based methods register surfaces on the basis of relatively dense point sets brought into correspondence, where these point sets constitute all, or a significant subset of, the available surface point samples. An alternate name for surface matching based on dense point sets is *free-form* surface matching [33, 424]. The two point sets are generally assumed to be relatively close to being aligned, and are usually registered by iteratively minimizing a global function such as the sum of squared distances between mutually closest points from (possibly transformed) surface S_A and surface S_B , as expressed by equation (2.3). Differences between many of these methods exist strictly at the level of the choice of distance metric and of the methods of optimally finding a match based on this metric (described in detail in § 2.4). Figure 2.4 illustrates this technique.

One common distance metric is the distance from a point $\mathbf{x}_B$ in set $X_B = \{\mathbf{x}_B\} \subseteq S_B$ to the (transformed) point set $X_A = \{\mathbf{x}_A\} \subseteq S_A$:

$$d(\mathbf{x}_B, \mathbf{R}_k \mathbf{x}_{A,min} + \mathbf{t}_k) = \min_{\mathbf{x}_A \in X_A} d(\mathbf{x}_B, \mathbf{R}_k \mathbf{x}_A + \mathbf{t}_k) \quad (2.13)$$

Besl and McKay [33] proposed the *Iterative Closest Point* (ICP) method to determine the closest point pairs according to equation (2.13), then computed the transformation from these pairs with a quaternion technique. The positions of the surface points S_A (“data shape”) were then updated: $X_{A,k} = \mathbf{R}X_{A,k-1} + \mathbf{t}$ and the process iterated until the mean-square distance, or point matching error, stabilized to within some



Figure 2.4 : *Iterative point-based registration of phantom face range data (courtesy of S. Lavallée, R. Szeliski and L. Brunie, copyright MIT Press 1986 [210]).*

tolerance. An accelerated variant of the ICP method (discussed in § 2.4) was also proposed.

Zhang [424] and Lavallée and Szeliski [209] adopted an objective function identical to the Besl's, except for weighting factors to accommodate measurement noise, and in the case of Zhang, to exclude outliers. They also introduced execution efficiencies: Zhang used *k-D trees* to make the search for closest points more efficient, and Lavallée dispensed with this search altogether by precomputing a *distance map*, which unfortunately did not provide explicit point pairs and thus still entailed a search for optimal transformation parameters. Simon et al. [354, 355] applied this technique to intrasurgical registration of a range image of the head with a surface extracted from tomographic data and addressed the issue of sensitivity to the perturbation of individual points. Maurer has complemented this technique with manually identified sparse homologous points pairs, in the context of registering head surfaces [244].

Other point-based techniques, which differ from those discussed above in terms of the distance metric which is minimized, have been proposed by Chen and Medioni [63], by Pelizzari et al. [303], by Soucy and Ferrie [366], and by Rangarajan et al. [321]. Chen used a subset of control points in relatively smooth areas, leading to an iterative technique with very good convergence properties. For each surface normal $\mathbf{n}_{A,i}$ defined at a control point $\mathbf{x}_{A,i}$, its intersection $\mathbf{x}_{B,i}$ with surface S_B is found. Next, defining $s_{B,i}$ the plane tangent to S_B at $\mathbf{x}_{B,i}$, the global transformation which minimizes the sum of squared distances between the set of transformed $\mathbf{x}_{A,i}$ with their

corresponding tangent planes $\mathbf{s}_{B,i}$ on S_B is computed. Pelizzari fitted a “hat” or external surface, consisting of relatively sparse points from the scalp as imaged by the modality of lesser resolution or coverage, to a “head” constituted by the set of 2D contours extracted on a slice-by-slice basis from the higher coverage/resolution modality. The residual minimized by this process is the sum of distances from each hat point to the head surface, along a direction from the former point to the centroid of the head. Soucy proposed an iterative technique which takes local surface shape into account. It matches small surface patches (as small as 3×3 pixels) by minimizing a similarity functional which enforces compatible local shape and piecewise-smooth motion. Rangarajan matched points with a technique (demonstrated on 2D autoradiograph slices but equally applicable to 3D surfaces) that imbeds the search for match pairs and for optimal transformation parameters, as well as the explicit exclusion of outliers, into one elegant minimization, making the method more robust to initial transformation estimation than the ICP technique.

Finally, Feldmar and Ayache [107, 108, 109] did not attempt strictly to find closest points, but to find closest *feature vectors*. This approach is a comparison of 8 parameters, namely the coordinates of each point (x, y, z) , the components of its normal (n_x, n_y, n_z) and the principal curvatures κ_1 and κ_2 of the surface at that point. Feldmar determined a globally affine transformation and then possibly a set of locally affine transformations, by extended Kalman filtering (see § 2.4), based on point-pairs that minimized the following expression:

$$\begin{aligned}
d(\mathbf{v}_{A,i}, \mathbf{v}_{B,j}) = & \alpha_1(x_{A,i} - x_{B,j})^2 + \alpha_2(y_{A,i} - y_{B,j})^2 + \alpha_3(z_{A,i} - z_{B,j})^2 + \\
& \alpha_4(n_{x\ A,i} - n_{x\ B,j})^2 + \alpha_5(n_{y\ A,i} - n_{y\ B,j})^2 + \alpha_6(n_{z\ A,i} - n_{z\ B,j})^2 + \\
& \alpha_7(\kappa_{1\ A,i} - \kappa_{1\ B,j})^2 + \alpha_8(\kappa_{2\ A,i} - \kappa_{2\ B,j})^2,
\end{aligned} \tag{2.14}$$

where α_i are weights determined by minimum and maximum data values. In other words, the closest point to $x_{A,i}$ on S_B is one which best fulfills a compromise between 3D distance, difference in normal orientation, and difference of principal curvatures.

2.3.3 Model-Based Representations

Deformable surface models are reviewed in §5.3.1, in the context of anatomical surface identification. Surface models can also be used to track a surface over a volume sequence, using a model that reconciles the likely shape at any instant and possibly the dynamic behaviour of the surface, with raw image data. With some exceptions [336, 6], such methods generally compute curve and surface motion not by explicit matching, but by implicit consideration of image motion in the form of virtual forces that tend to make the model agree with shifts in strong image gradients. This section discusses model-based methods that emphasize registration, whereas §5.3.1 offer a broader review of active surface models.

Physically-based deformable surface models view a surface, which is tracked over time, as being in equilibrium, or quickly reaching equilibrium, between internal and external virtual forces. Internal forces include inertia, damping and strain, and are typically determined by the current state which is assumed for a 3D object. External forces are typically determined by image data. The relative motion undergone by the various parts of a surface is a result of the interaction of these two sets of forces, according to formal physical principles.

One early physical model applied to registration is Bajscy and Kovacic's multiresolution elastic matching method [20]. This model is expressed as

$$\mu \nabla^2 u_i + (\lambda + \mu) \frac{\partial \theta}{\partial x_i} + F_i = 0 \quad (i = 1, 2, 3), \quad (2.15)$$

where θ is the *dilatation*⁵ at a point on the body, and μ and λ are *Lamé's constants*, which define the elastic properties of the model. The *external forces* $\mathbf{F} = (F_1, F_2, F_3)^t$ bring similar regions of two 3D objects into correspondence by enforcing grey-level value correlation and edge alignment between blurred volumes, in a multiresolution framework.

Research conducted separately by Pentland [306, 100] and by Terzopoulos [389, 250] proposed a finite-element model approach to the numerical solution of deformable

⁵This is the change in volume per unit initial volume under small strains and sum of principal strains in general [239].

surface models, and consequently these surface models are based on the *finite-element equilibrium equation*:

$$\mathbf{M}\ddot{\mathbf{a}} + \mathbf{C}\dot{\mathbf{a}} + \mathbf{K}\mathbf{a} = \mathbf{R}, \quad (2.16)$$

where $\mathbf{M}$, $\mathbf{C}$, and $\mathbf{K}$ are virtual *mass*, *damping* and *stiffness* matrices, $\mathbf{a}$ is the displacement of the FEM nodes, and $\mathbf{R}$ is the sum of external forces, determined by image data.

Improvements relating to surface tracking were proposed by Nastar and Ayache [279] and by Sclaroff and Pentland [336]. If physically based models could be viewed as virtual masses on the surface with springs relating them to their neighbours, this viewpoint could be extended over time to a spring between a boundary point at t_1 and its closest neighbour at t_2 , as pointed out by Nastar. Moreover, salient feature points such as curvature extrema were used to anchor the dynamic behaviour of the model, by attaching particularly stiff springs between feature point pairs. Sclaroff suggested that FEM numerical estimation could benefit from a change of basis from *nodal* to *modal* displacements ϕ , where $\mathbf{a} = \phi \sin[\omega(t - t_0)]$ and ω represents a virtual frequency of vibration, based on object shape, corresponding to each mode (see also the section in Bathe [26] on mode superposition). This change of basis was justified not on the grounds of expected periodic motion, but because it decouples the system of equations (5.5), which was then expressed as the eigenproblem (neglecting damping by taking $\mathbf{C} = \mathbf{0}$):

$$\mathbf{K}\Phi = \mathbf{M}\Phi\Omega^2 \quad \text{where} \quad \Phi = [\phi_1 | \cdots | \phi_p] \quad \text{and} \quad \Omega^2 = \begin{bmatrix} \omega_1^2 & & \\ & \ddots & \\ & & \omega_p^2 \end{bmatrix}. \quad (2.17)$$

The matrix Φ has an interesting interpretation: its vector entries ϕ_i can be ordered according to increasing corresponding eigenvalue ω_i^2 . In this case Φ is an orthogonal, frequency-ordered description of an object's shape and its natural deformations, somewhat like a Fourier series. An important consequence of this representation is that it leads to an explicit matching algorithm, after an initial rigid alignment, where low-order nonrigid modes ϕ_i are used as a coordinate system, and two shapes are nonrigidly registered on the basis of modal feature vectors.

The model of Amini and Duncan [6] uses 3D surface points as its raw data, and views the local surface shape as a bending energy from an idealized thin flat plate $\varepsilon_{be}(u, v) = \kappa_1^2 + \kappa_2^2$. It seeks to match surface points of consistent energy and therefore of consistent principal curvature. Moreover, their model also has a stretching energy term which penalizes *non-conformal* motion (i.e. motion where the proportionality of the various coefficients of the first fundamental form, E , F and G , are not maintained). Consequently, the following energy measure is minimized over two surfaces:

$$\begin{aligned} & \lambda_{be} \{ (\kappa_1 A - \kappa_1 B)^2 + (\kappa_1 A - \kappa_1 B)^2 \} + \\ & \lambda_{st} \left\{ \left(\frac{E_A}{E_B} - \frac{F_A}{F_B} \right)^2 + \left(\frac{F_A}{F_B} - \frac{G_A}{G_B} \right)^2 + \left(\frac{E_A}{E_B} - \frac{G_A}{G_B} \right)^2 \right\}. \end{aligned} \quad (2.18)$$

The tracking model of Cohen et al. [69] is similar to Amini's in that it is characterized by a functional that minimizes curvature differences, between two 2D contours at comparable arclength parameters, while enforcing smooth motion along arclength. The zero arclength value on each contour corresponds to a matched feature point. Cohen has proposed an extension to 3D surface tracking in the same paper.

2.3.4 Techniques Based on Global Shape

The feature, point and model-based algorithms discussed so far can be broadly described as relying on local information to register surfaces. However, there are a few recently-published algorithms which register surfaces on the basis of global surface geometry, that do not rely on a rough prior estimation of the transformation and that may be able to deal with relatively featureless patches. Such approaches are the *spin-map* representation of Johnson and Hebert [173] and the *eigenshape* or *appearance-based* methods for registering 3D surfaces [53].

The spin-map representation is a set of 2D footprints conveying global surface shape in the neighbourhood of selected *oriented points* [173], as illustrated in figure 2.5. A 3D oriented point consists of a point $\mathbf{p}$ and a surface normal $\mathbf{n}$ estimated at that point. In order to describe its neighbouring topology, a 2D basis is formed. To define this basis, Johnson first defines a line L through $\mathbf{p}$ whose direction is along $\mathbf{n}$, and a tangent plane P through $\mathbf{p}$. The two coordinates of the bases are α , the perpendicular distance of a neighbouring point to line L , and β , the perpendicular

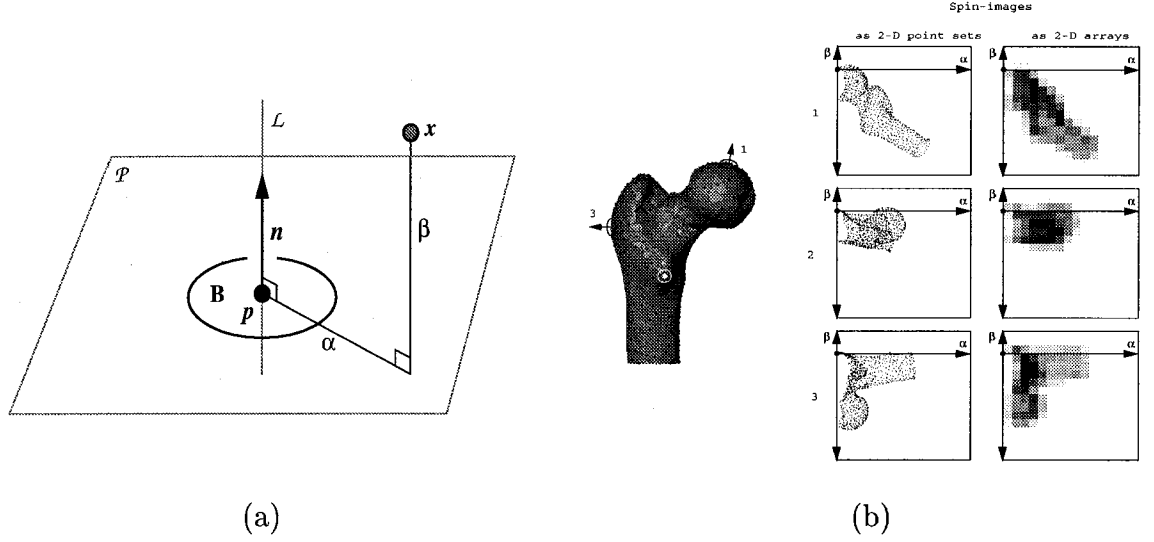


Figure 2.5 : (a) An illustration spin-map geometry; (b) points distributed on the surface of a femur bone, shown with three point bases and their corresponding spin-maps, appearing as point sets and grey-scale images (courtesy of A.E. Johnson).

distance of this point to plane P . A spin-map S_O is the function which maps 3D neighbouring points $\mathbf{x}$ to the 2D coordinates of a basis corresponding to an oriented point O :

$$S_O(\mathbf{x}) \rightarrow (\alpha, \beta) = \left(\sqrt{\|\mathbf{x} - \mathbf{p}\|^2 - (\mathbf{n} \cdot (\mathbf{x} - \mathbf{p}))^2}, (\mathbf{n} \cdot (\mathbf{x} - \mathbf{p})) \right). \quad (2.19)$$

The term spin-map comes from the cylindrical symmetry of the oriented point basis. Given this basis, the 3D shape of the neighbouring surface points can be reduced to a 2D snapshot of global shape, which is a representation invariant to a rotation about its central axis L . 3D surface point correspondences can then be established on the basis of a 2D correlation between their respective spin maps. The applicability of the technique to anatomical surfaces is demonstrated with skull and bone data sets. The use of points matched by spin-map correlation to determine the transformation between two surfaces is comparable to feature-point matching, as addressed in §2.4.

Early work on appearance-based techniques mostly dealt with face recognition in 2D intensity images, and was based on projecting face images onto a feature space that spans the significant variations among known face images [399]. Murase and

Nayar [275] have applied a similar framework to recognize 3D objects, whose models are stored in a CAD database, from 2D images, as well as to estimate the transformation between the two. This in turn was the basis for the eigenshape technique of Campbell and Flynn [53] for object recognition and pose estimation of range images.

In Murase's application, the basic idea is to take several training views of an object, resting on a turntable. The object pose is varied by rotating the turntable by an angle θ in increments of a few degrees. Each $N \times N$ image $I(i, j)$ corresponding to a training view is represented as a N^2 -vector

$$\hat{\mathbf{x}} = [I(1, 1) \cdots I(1, N) \ I(2, 1) \cdots I(2, N) \cdots I(N, N)]^t, \quad (2.20)$$

obtained by concatenating the rows of the image and taking the transpose. An image was then equivalent to a point in a huge N^2 -dimensional space. Fortunately, most of the variation in the images can be accounted for by a subspace spanned by just a few vectors, as in *principal component analysis* (or Karhunen-Loeve expansion) [117]. Each vector corresponding to a distinct training view constitutes a column of an *object image set*

$$\mathbf{X}_{N^2 \times m} = [\hat{\mathbf{x}}_1, \dots, \hat{\mathbf{x}}_m], \quad (2.21)$$

where m is the number of training views. A covariance matrix is then defined from the image set:

$$\mathbf{Q}_{N^2 \times N^2} = \mathbf{X}\mathbf{X}^t. \quad (2.22)$$

The eigenvectors $[\mathbf{e}_1 \cdots \mathbf{e}_k]$ corresponding to the k largest eigenvalues are computed (Murase has listed three algorithms for doing this efficiently), where $\lambda_i \mathbf{e}_i = \mathbf{Q}\mathbf{e}_i$, comprising a *parametric eigenspace*.

Thereafter, the set of training views projected onto the parametric eigenspace constitutes samples of a smoothly varying manifold. If $k = 3$ this manifold is a closed curve whose arclength is a one-to-one function of θ in the 3D space spanned by $[\mathbf{e}_1 \mathbf{e}_2 \mathbf{e}_3]$, which in turn can be fitted with cubic splines. Given a new image, its pose can be estimated by projecting it onto the parametric eigenspace. For example,

if $k = 3$, this can be achieved by finding the position of its projection within the cube spanned by $[\mathbf{e}_1\mathbf{e}_2\mathbf{e}_3]$ and by finding the closest point to this projection on the manifold. The pose can then be estimated by interpolation along arclength between the two closest discrete poses.

Now, for surface registration [53], we are dealing with more than one pose parameter (i.e. not just θ corresponding to turntable rotation), but the framework is the same (and this could be applied to registering an open anatomical surface with a closed surface extracted offline from tomographic data). The steps are (the first three can be implemented offline):

- compute training views of the anatomy at sufficient increments of each pose parameter;
- compute the parametric subspace that captures enough of the shape variation across all poses;
- project back the training views to this subspace to generate a manifold (which may then be smoothed);
- and for a new open surface, project it also onto the parametric subspace and obtain its transformation by interpolating the projections of discrete training poses.

2.4 Matching, Optimization and Transformation Computation

The third stage comprises the search for corresponding points or feature pairs, based on the surface representation of the second stage, and the computation of the optimal transformation as idealized in the first stage. The search for a match can be either a succession of comparisons of discrete candidates, as is frequently the case for feature pairs, or an iterative minimization of an objective function, as usually occurs with point- and model-based schemes. The subsequent transformation computation

depends on the assumptions made in the first stage. It comprises a closed-form computation or an iterative search for the six or more parameters which best align feature pairs. Alternately, a *Kalman filtering* approach presents advantages to tracking surface points over an extended time sequence, as this estimation of transformation parameters is optimal for the noise characteristics of the whole sequence while being recursive over time. As for the techniques based on global shape, points whose spin-maps correlate well are used to (over-)determine a transformation in the same manner as discrete feature points (discussed below). The special case of appearance-based matching was fully described in the previous section.

2.4.1 Discrete Feature Matching and Transformation Computation

Feature matching generally determines a rigid transformation, due the relative sparseness of features, and involves a comparison in terms of shape parameters, such as surface curvature values at extrema, curvature and torsion values of curvilinear features, and shape type, averaged properties (such as Gaussian or mean curvature) and size for regions. In the event a particular matching technique does not imbed a transformation computation, most of the rigid-body motion representations reviewed in § 2.2.1 lead to a closed-form expression for estimating rotation and translation. Early techniques include *pose clustering* (also known as generalized Hough transform), *sequential hypothesis-and-test* (also known as prediction and verification and alignment), and *geometric hashing*. I also review the *eigenvector* technique of Shapiro and Brady [344] for matching point features.

Pose clustering matches similar simple or compound structures and derives a transformation from each such correspondence. This approach involves quantizing the space of relevant transformations and using it as an accumulator in which each match increments a corresponding cell. A globally acceptable transformation is detected as a cluster in the space of all such candidate transformations [376].

Sequential hypothesis-and-test consists of picking a set of feature pairs which

are consistent and which determine a transformation, validating these hypothesized matches based on how other features from the two spaces agree with the putative transformation, then possibly backtracking and proceeding anew with a new set of consistent feature pairs [38, 61].

Geometric hashing involves precomputing local matching information, which is rotationally and translationally invariant, and storing it in the form of a *hash table*. Each table entry is associated with a simple or compound feature from surface A to which a local coordinate system (or basis) can be unambiguously assigned. Subsequently, a feature or set of features in surface B is similarly characterized. Finally, given the transformation between the two local bases, the consistency of the mapped non-basis features is evaluated, where consistent feature pairs validate this transformation by voting for it, or incrementing an accumulator corresponding to this transformation [203, 194].

Shapiro and Brady [344] matched feature points on the basis of consistent same-space distances by an elegant eigenanalysis technique, following the inter-image distance-based matching technique of Scott and Longuet-Higgins [339]. Shapiro suggests accounting for global structure, and proposes that each image or surface be represented by a proximity matrix $\mathbf{H}$ of intra-image (or intra-surface) distances. The eigenvalues of each matrix are then computed, resulting in a modal matrix $\mathbf{V}$ whose columns correspond to the eigenvectors of $\mathbf{H}$. Each row of $\mathbf{V}$ can be thought of as a feature vector $\mathbf{F}_i$ containing the modal coordinates of a feature i . The final stage consists of correlating the two sets of feature vectors $\mathbf{F}_{i,A}$ and $\mathbf{F}_{j,B}$.

In general, techniques involving some sort of voting or accumulation, such as pose clustering and geometric hashing, are likely to be the most robust, particularly to missing information (such as a partial overlap between different modalities, i.e. registering a sub-patch with a larger surface patch).

2.4.2 Closest Point Finding and Global Optimization

Free-form surface matching involves a closest-point finding stage, which may or may not benefit from some form of preprocessing (k-D trees, distance maps), and frequently a global optimization stage which is used either to compute the best current transformation T_k based on the latest closest-point pairs, or to accelerate the convergence of these fundamentally iterative techniques towards a definitive result.

Classical techniques can be used to implement the minimization of equation (2.13), and generally are of the *unconstrained nonlinear optimization* type [225]. These approaches are based on the conditions that, for a minimum, the objective function must have a null gradient and a positive semidefinite Hessian matrix. Starting at some initial point, one determines, according to a fixed rule, a direction of movement in the domain, then one moves in that direction to a relative minimum of the objective function along that line. At that point, a new direction is determined and the process is repeated, until some termination condition is met. These techniques include the *method of Steepest Descent*, *multivariate Newton's*, *Conjugate-Gradient* and *Quasi-Newton methods*, and related methods [225].

These methods are also used, once point-matches are determined, to iteratively compute optimal transformation parameters. Lavallée and Szeliski employed a Levenberg-Marquardt technique, which is a hybrid of the Steepest Descent and multivariate Newton's Method [225], to iteratively compute both rigid [209] and non-rigid [382] transformations. Besl and McKay's accelerated ICP technique involved a series of line searches in a 7-parameter space spanned by rotation quaternion and translation [33]. The unaccelerated stage found closest point pairs and used a quaternion technique to find the current iteration's least-squares registration. Each such iteration is equivalent to a small step in 7-space, and the accelerated stage fitted an interpolant in the direction of this step (from the last 7-vector), whereupon the convergence could be improved by a line search for the minimum. Special switching logic determined whether a linear or parabolic interpolant was used.

The *k-D tree* [424] is a sequence of bisections in a space of k dimensions (here

$k=3$). Each successive cut plane passes through a point chosen such that it divides the remaining points into clouds of roughly equal numbers of points, producing a left and a right son. Each son is split into two grandsons by choosing the appropriate plane parallel to $x-z$, followed by a plane parallel to $x-y$, and so on, while alternating the cut plane orientation, until sufficient divisions occur such that no resulting rectangular parallelepiped contains a data point. Such a cut is termed a “leaf”. Each node ν of the 3-D tree is characterized by a point $\mathbf{x}_{A,i}(\nu)$ through which it passes, and a parameter $t(\nu)$ indicating the orientation of the cut plane. The search for the closest point via a 3-D tree calls a recursive procedure that begins at the root of the 3-D tree assigned to X_A , and exploits the tree structure, branching off into one son or the other depending on the signs of the distance components, to zero-in on the best candidate.

A *distance map* can be precomputed off-line, to determine for each voxel in a volume containing a surface S_B , the closest point on that surface and the distance to it. Thereafter, for a surface S_A falling within this mapped volume, each of its points inherits the closest point precomputed for the voxel on which it falls. In the case of an *octree spline* implementation [209], the representation is carried out by a classic octree subdivision [115]. For each corner vertex of all terminal cubes, the signed distance to its closest point on the surface is computed. The determination of the distance to the closest point on S_B is achieved by interpolation. Alternate techniques for computing the distance map of a surface are by *fast marching level sets* [341, 191], where distance is the arrival time of a moving front starting from the initial surface, and by various local mask-based distance transforms [40, 41, 299].

Alternately, each match pair $\hat{\mathbf{x}}_k \equiv (\mathbf{x}_{A,k}, \mathbf{x}_{B,k})$ can be viewed as a measure $\hat{\mathbf{x}}_k$ of a match $\mathbf{x}_k$ resulting from the application of the true transformation relating the two surfaces, corrupted by a random error $\mathbf{v}_k$. This leads to an *extended Kalman filtering* (EKF) [175, 364] formalism for recursively estimating the optimal (with respect to noise statistics) transformation parameters. The application of EKF to tracking surface points over time is an issue of expressing a relationship between positions or feature vectors of the matched pairs and the transformation parameters as a *mea-*

surement equation that can be linearized, and for which the linear Kalman filter can recursively compute the best transformation estimate. This is the technique used by Feldmar and Ayache [109] for estimating the global affine transformation relating matched feature vector pairs. Because it is recursive, it presents advantages for tracking surface points over time. Only the new measurements and the statistics and transformation parameters of the previous iteration need to be considered, in computing the parameters which are optimal for the noise statistics of the whole sequence.

2.4.3 Model-Based Motion Estimation

The method of estimating motion from surface models depends on the nature of the model. As discussed in § 2.3.3, many physically-based models segment the surface explicitly, but surface motion is often just the difference between consecutive shapes at a given surface coordinate (u, v) , particularly when the shape at time t_k is used to initialize the shape estimation at t_{k+1} [391, 250]. The application of surface evolution models to motion estimation is still in its infancy, but as I will show in subsequent chapters, this framework can be exploited not only to extract a surface reliably, but also to alleviate an explicit search for matches in a subsequent point-based registration scheme.

Finite element modelling is the most prevalent technique for computing physically based models [100, 389]. Here, the displacement FEM represents a surface as a mesh of 2D simple polygonal elements whose global behaviour (where a surface reaches equilibrium between internal and external forces) can be characterized by an element-by-element analysis of the dynamics involved. The global mass, damping, and stiffness matrices are summations of the corresponding element matrices, as are non-concentrated load matrices. In practice, the integration of expression (5.5) is simplified by neglecting either mass [389], leading to the *explicit first-order Euler integration*⁶ $\mathbf{a}^{(t+\delta t)} = \mathbf{a}^{(t)} + \delta t (\mathbf{C}^{(t)})^{-1} (\mathbf{R}^{(t)} - \mathbf{K}\mathbf{a}^{(t)})$, or damping, leading to the change

⁶Based on virtual time steps which need not be related to the time lapse separating two surfaces.

of basis to modal coordinates and the expression (2.17), which can also be integrated numerically.

The Amini surface tracking technique finds closest points in a manner similar to free-form surface registration techniques [6]. For each point of surface A , the point on surface B that minimizes expression (2.18) is selected. This ensures that the match is most similar in local surface shape and that its relative movement is nearest to being conformal. The result is a 3D flow field, which may be noisy, defined over the first surface. Thereafter, a *vector smoothing* technique [159] is applied to components of each flow vector, and flow estimates are propagated over other regions of the surface.

Other formulations have also been proposed for computing or tracking deformable model points. The change of basis of expression (2.17) allows Sclaroff and Pentland [336] to adapt Shapiro's *modal matching* technique to explicitly match FEM surface nodes. Many authors [305, 256, 390] apply a Kalman filter formalism to estimate 3D motion from a physical model. Cohen [70] has suggested a *finite-differences* based technique for finding the energy-minimizing contour of his balloon model. For alternative approaches to estimating deformable contours and surfaces, see [390] and [251].

2.5 Conclusion

This chapter has presented a survey of surface registration techniques, particularly those that apply to anatomical surfaces, with an emphasis on their mathematical or algorithmic foundations. I chose to represent the process of registration as the succession of three stages: choice of transformation representation, choice of surface representation and similarity criterion, and matching per se and global optimization. According to § 2.2, transformations can be categorized as rigid and non-rigid. Several different representations for rigid-body transformations were surveyed, and I demonstrated that non-rigid transformations can be further categorized into global and local polynomial representations, and into formulations of Cartesian or spherical coordinates.

We saw in § 2.3 that a feature, point, global shape or model-based approach can be employed to represent a surface for the purpose of registration, and that each has certain advantages. Feature-based and global shape-based techniques can determine an arbitrarily large rigid transformation. A point-based technique is best suited to bringing two surfaces into very close alignment, given a good starting point for the final transformation, particularly if we wish to quantify non-rigid motion. On the other hand, a model-based approach is appropriate when the surface within a volume must be identified prior to registration. Furthermore, I discussed three categories of features: sparse points, curves and regions. I demonstrated that dense point-based techniques vary according to their definition of proximity between points and according to their means of estimating the closest point to a surface. Finally, I pointed out that models for registering surfaces imbedded in volumetric data are typically based on physical equations. I identified some techniques which compute surface motion implicitly, from the evolution over time of a 3D model, and emphasized methods which explicitly establish correspondences between surface points. A more complete literature review on surface models, emphasizing anatomical surface identification, appears in § 5.3.1.

Next, in § 2.4 I offered a survey of the main numerical estimation schemes or algorithms for matching features (and points correlated on the basis of spin-maps), points and model coordinates (matching on the basis of appearance was concluded in the discussion on representation). Features are typically matched by discrete comparisons, in a manner which may imbed the computation of the resulting transformation in the matching process. Since the transformation specified by these matches is usually rigid-body and the robustness of each match is typically very good (due to the discriminating ability of features), the global registration parameters can then be obtained by a closed-form expression. In contrast, matching and registration for point-based techniques are imbedded into a highly iterative process, with the matching characterized either by a classical optimization technique or by a more efficient procedure requiring offline preprocessing, such as a k-D tree or distance map. The registration may also require an optimization or Kalman filtering procedure for finding the

best-fitting transformation parameters in the global sense. Lastly, for physical model-based approaches, surface motion computation frequently amounts to identifying the surface over consecutive volumetric scans. This identification is usually implemented with the finite-element model, where a surface at t_1 initializes that at t_2 . In contrast, surface evolution models are less prevalent in registration, and typically estimated by finite differences.

In conclusion, the design choices, namely how to model the transformation, whether to use a feature, point, global shape or model-based technique (as well as how to preprocess the data) and how to optimally compute the transformation, are very context-specific. There are instances, e.g. when dealing with rigid tissue, when a rigid-body transformation will suffice, and others, such as tracking soft tissue deformations, where a spline-based non-rigid transformation model is necessary. In between these two poles, there is a whole spectrum of situations where a rigid model may be too restrictive, but an approach based on splines may also be inapplicable, by being underdetermined or inappropriate for a given noise distribution, for example. Likewise, the choice of surface representation is influenced by the size of the transformation (e.g. whether surfaces are already roughly aligned or not), by the smoothness, noise distribution, and mutual overlap of the data, as well as by time constraints. Moreover, establishing correspondences between two sets of primitives is basically an outgrowth of the formalization of the transformation and the choice of primitives, and consequently is also highly related to the situation at hand.

Chapter 3

Laser Range-sensing and Calibration-based Range-MR Transformation

3.1 Introduction

In order to characterize brain shift, one must first find a simple-to-use, robust means of measuring the 3D shape of the brain exposed in the OR. In order to address this objective, I chose to measure the external surface shape, rather than use a tomographic method, because of the ease of use and potential commercial applicability of a surface tracking and FEM approach. Moreover, tomographic studies have indicated that much of the brain movement is concentrated near the brain surface exposed by the craniotomy [245, 77], so that a good characterization of the volumetric displacement function is possible with a surface tracking and FEM approach. I note however that the accuracy of the volumetric estimation of this method would degrade if there were motion in deep structures of the brain which bore little relation to the motion of the exposed brain surface.

Laser-based range-sensing is the industrial standard for quickly and accurately acquiring dense, quantitative information on the 3D shape of a visible surface (as

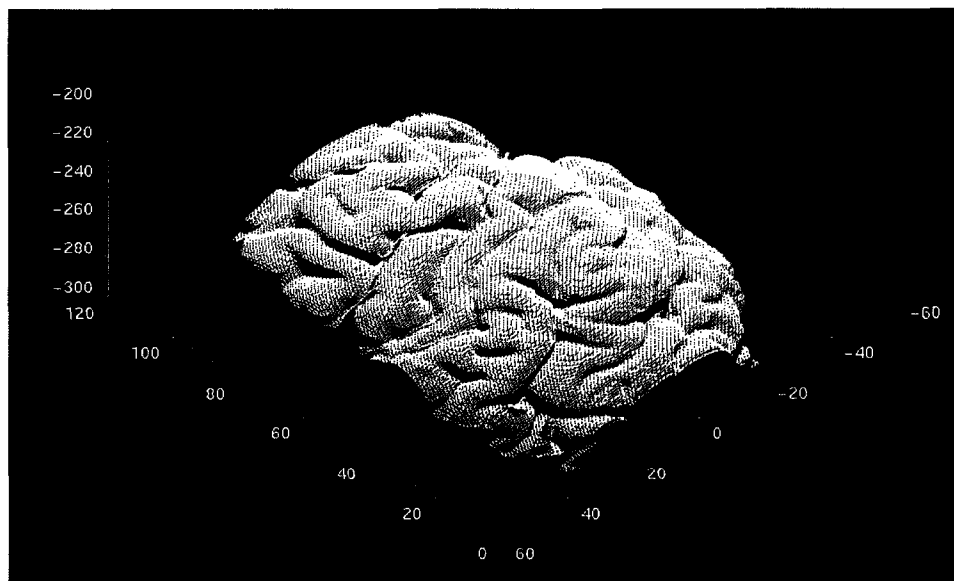


Figure 3.1 : *Range image of the left hemisphere of a human brain.*

opposed to a tomographic modality producing volumetric data), used for example in robotic navigation, industrial quality control, and welding automation [103]. Figure 3.1 illustrates the dense shape characterization achievable with this technology. It is worth noting that other 3D surface shape estimation methods could have been used, such as shape-from-stereo [90, 12, 11], or shape-from-focus [280]. In particular, our brain shift estimation framework would transpose well to a binocular surgical microscope, by replacing the range sensor with a stereo algorithm running on the binocular image pair. The underlying correspondence problem could be simplified and accelerated by projecting an infrared pattern, unseen by the surgeon, on the exposed brain surface, and using the disparities of identifiable points seen in each view to determine stereoscopic displacements [284]. Nevertheless, I use a commercial laser range-sensor to capture 3D brain surface shape, because of the ease of use, widespread availability, proven robustness, and sub-millimeter accuracy of this technology. Section 3.2 of this chapter provides technical details of the sensor.

As alluded to in the preface, there are two registration problems which are relevant to our method for characterizing brain shift:

- *relating range data to MR data*, in order to establish the baseline for brain

motion in the range referential, and

- *tracking the non-rigid cortical motion* implicit in the time sequence of range images.

The first problem could perhaps have been solved by computing the MR-range transformation on the basis of a fully autonomous skin patch registration. However, as indicated in chapter 2, registering a small range image of smooth cranial skin with its MR homologue is a very difficult problem. A feature-based method, normally used for arbitrarily large transformations, would likely not be robust here because of the lack of anatomically relevant, reliably identified high-curvature points on the top and back of the head. Furthermore, a model-based technique is essentially inapplicable here because it is designed for two closed surfaces (and in any event presumes a good initial alignment). Finally, a technique based on global shape description which did not have the benefit of an initial alignment would also be unreliable, because this situation tends towards matching a patch from a near-spherical ellipsoid with its homologue, so that global shape descriptions of such a surface patch lack discriminating power.

As discussed in § 3.3, the technique which seems most useful is a calibration-based procedure. My approach is inspired by Comeau’s ultrasound-to-MR registration technique [77], and, as illustrated on figure 3.2, it features:

- A *calibration* stage which relates the internal range coordinate system to external plates (featuring milled divots) bolted to the range-sensor, which I call the *sensor base*¹.
- A *sensor base tracking* stage which determines the range-probe transformation in the OR. The divots on the external plates, referred to range coordinates by the previous stage, can be matched with their homologous probe-space positions in the OR in order to determine the range-probe space rigid transformation.

¹A consistent sensor base position presumes that the sensor is at its *home position*, or fully retracted ($y = 0$) on the linear positioner, when not scanning.

- *The IGS patient registration* is finally multiplied with the range-probe transformation resulting from sensor base tracking to relate the range and MR coordinate systems.

If the calibration, sensor-base tracking and image-probe registration procedures were all error-free, the resulting range-MR transformation would be sufficient to establish an accurate baseline from the identified brain surface in the MR scan. Because no procedure is error-free, and any errors are cumulative, one may opt to refine the final result with a rigid-body Iterative Closest Point (ICP) registration. The latter part of this chapter provides a discussion of the procedures for relating range space to MR space: calibration, sensor-base tracking, and ICP refinement, as well as details of a validation study of these procedures.

3.2 Laser Range-sensing

The three-dimensional coordinates of the exposed cortical surface are computed by a commercial range-sensor made by Vitana Corp. [405] (Ottawa, Canada), based on the Canadian National Research Council's Biris design [35]. This sensor uses both laser-based *triangulation* and *defocusing* techniques to estimate range. The sensor is mounted on a commercial linear positioner, produced by Intelligent Actuator [167]. The range-sensing configuration is featured in figure 3.3.

Laser-based triangulation involves projecting a thin ray of laser light onto an object surface, capturing the reflected light by a charge-coupled device (CCD), and inferring 3D position from the pixel position of the image point and from the geometric relationship between the position of the laser source and the CCD. The relationship summarizing the triangulation is illustrated in figure 3.4(a) and expressed as [35, 405]:

$$z_p = \frac{1}{l^{-1} + [(l - f)/(f \cdot D \cdot l)] \cdot p}. \quad (3.1)$$

In equation (3.1), z_p is the range value of the anatomical point obtained by triangulation, D is the *triangulation base* separating the laser source from the CCD optical center, l is the *focusing distance* defined from the lens to the image plane, f is the

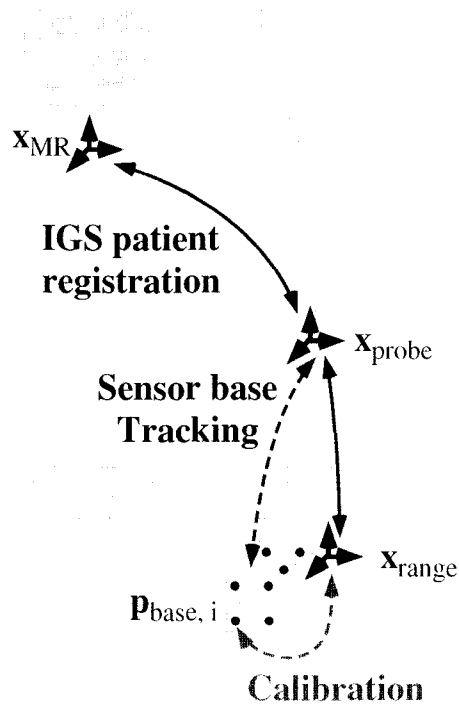


Figure 3.2 : *Range to MR transformation geometry, illustrating the range, probe and MR referentials and the sequence of procedures for relating them to each other.*

focal length of the lens, and p is the pixel position of the laser spot image on the CCD. Intuitively, the further an object is from the sensor, the more the ray reflected by the object tends to align with the optical axis of the lens; and the more the ray refracted by the lens tends to the optical center, the smaller the pixel offset p . The x -component of position is obtained by a similar triangulation. These relationships are characterized by a factory calibration procedure (not to be confused with the calibration procedure for computing the range-MR space transformation) after which each laser spot image detected at p on the CCD is interpreted as the (x, z) position of the anatomical point illuminated by the incident laser ray.

The Biris sensor allows the laser light to pass through a mask featuring two apertures at a predetermined distance d apart, rather than one, as shown on figure 3.4(b).

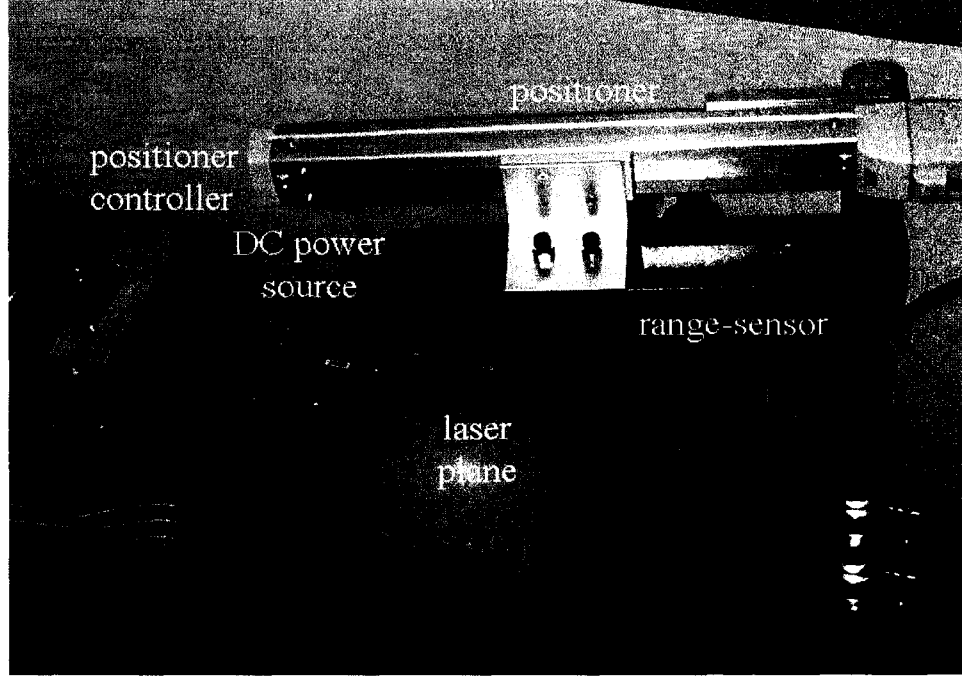


Figure 3.3 : *Illustration of range-sensor and positioner configuration in the laboratory, with labelled components.*

The sensor sees two laser spot images at p_1 and p_2 , instead of one. Triangulation in this sensor is implemented by considering the *average position* of the two spots $p = \frac{p_1 + p_2}{2}$ in expression (3.1). Defocusing allows the range value z_b to be determined independently from the *distance* $b = p_2 - p_1$ between the two laser spot images, according to the expression [35]:

$$z_b = \frac{1}{l^{-1} + [(l - f)(f \cdot d \cdot l)] \cdot b}, \quad (3.2)$$

where l and f are as in expression (3.1). Intuitively, defocusing relates to a) the positions of the two laser spot images produced by the rays reflected by an object, in relation to the position of the object, and b) to the position of the image p' which would occur through one central aperture located on the optical axis. The further an object is from the sensor, the smaller the effect of the offset between each aperture and the optical axis, and the closer the two images p_1 and p_2 are to the putative central-aperture image p' . Consequently, the separation between p_1 and p_2 decreases as z_b increases.

The range measurements obtained by triangulation and by defocusing are com-

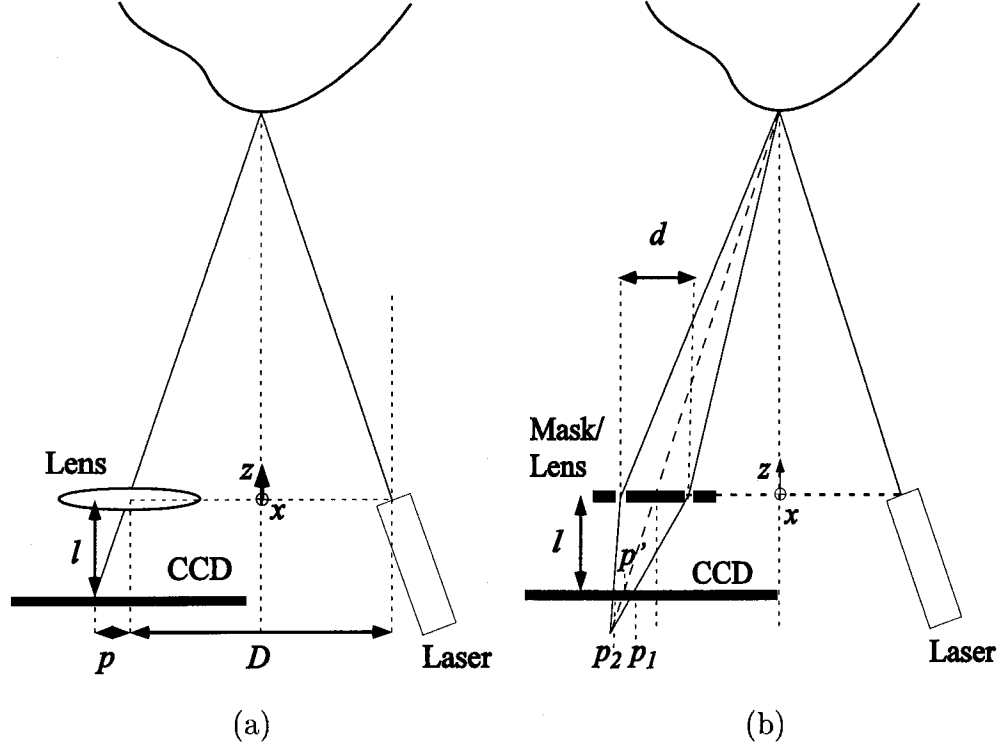


Figure 3.4 : *Biris laser range-sensing geometry: (a) triangulation; (b) defocusing*

combined by assigning weights α and $1 - \alpha$ to the two expressions for z given in expressions (3.1) and (3.2):

$$z = \frac{1}{\alpha \cdot (K_b b + l^{-1}) + (1 - \alpha) \cdot (K_p p + l^{-1})}, \quad (3.3)$$

with most (80%) of the weight given to former, because laser peak position varies much more than peak separation as a function of range [35]. The combination of the two techniques provides constraints for detecting laser intensity peaks unambiguously in any column of the CCD [405].

Ordinarily, in order to get range information at regular samples of a surface, the laser would have to be swept along two axes which are normal to each other and span a plane normal to the anatomical surface of interest. However, this sensor dispenses with one sweep axis by optically spreading the ray into a *laser plane* (this is visible in figure 3.3). The intersection of this plane with the imaged object instantaneously produces a 2D range-map of a “slice” of the surface, which appears as a curve of 256 points, or *laser profile*, on the CCD². Viewed through a two-aperture mask, the

²Laser profiles can be acquired at up to 60 Hz: collecting a 256×256 range image, consisting of

intersection of the laser plane with the surface of interest is seen as two laser profiles in the CCD. The triangulated result is obtained from the average position of the two laser spot images within each column, and the defocused result is given by the separation between the two spots in the each column. The final *range profile* coincides with the weighted sum of expression (3.3) evaluated across all 256 CCD columns. A 3D surface is then obtained by linearly sweeping the sensor in a direction normal to the laser plane, across the anatomy of interest, while taking range profiles at regular intervals. The final 3D surface is composed of 256 such range profiles. The coordinate y is given by the linear positioner, and it is orthogonal to the $x - z$ plane spanned by the range profile.

Finally, the sensor has a depth of field of 135 mm (from 108 mm to 243 mm), which for a 256-pixel CCD corresponds roughly to a pixel resolution of 0.5 mm. The field-of-view of this sensor at a typical range (150-200 mm) is around 100-120 mm (depending on the proximity of the sensor to the surface), which gives it a 0.4-0.5 mm lateral pixel resolution. The positioner, which linearly sweeps the sensor across the skin or cortical surface, possesses a repeatability of 0.02 mm and is capable of 150 mm travel at speeds up to 330 mm s^{-1} . During this linear travel, the program acquires 256 laser profiles at a constant rate, when the positioner is moving at a constant speed. The regular sampling nature of the CCD (in $i-k$ pixel coordinates) and the constant-rate data acquisition along the y -axis result in the collection of uniformly spaced samples. This fact is exploited later in the registration process. Both the sensor and the positioner are supplied with Windows-based software interfaces and dedicated control hardware. Both software interfaces support a standalone program which calls vendor-supplied libraries.

256 profiles spanning the surface of interest, requires just over four seconds.

3.3 Relating Range to MR Space by Probe-based Calibration, Sensor Base Tracking and IGS Patient Registration

The baseline for cortical movement is implemented by computing the rigid-body transformation between range and MRI spaces and applying this transformation to the identified cortex surface in the scan. As discussed in the introduction of this chapter, I might have computed this transformation on the basis of an autonomous registration procedure between a range image of the skin and its anatomical equivalent taken from the skin surface identified in the MR volume. In the literature, computing an arbitrarily large transformation between two surfaces is generally approached with features based on the differential geometry of the surfaces, or global descriptions of surface shape, but neither local features [393, 136] nor global shape signatures [172, 53] would be effective on a small patch taken from the top or back of the head. However, the technique presented here uses the IGS tracked probe to establish the position of the *sensor base* in probe space and ultimately in MRI space. All that is needed to complete the picture is a *calibration* procedure to relate the sensor base to the sensor internal reference.

An important point to be made here is that a very accurate range-MR transformation is achievable from the automatic detection of implanted fiducials [409]. However, because of the practice at the MNI of not using implantable fiducials for IGS registration, this approach was not used. The automatic detection of implantable fiducials in both in MR and range data, afforded by a well thought-out fiducial design, would lead to their unambiguous, accurate localization, which in turn would lead to a direct computation of the range-MR transformation on the basis of explicit homologous pairs. For example, the cylindrical design published by Wang et al. [409], which can be filled with a contrast agent appropriate to MR or CT, could be modified so that the flat, circular top surface, visible to the range sensor, would reflect laser light in an intense, non-specular manner, allowing automatic localization on the basis of the

known geometry of the fiducial.

The geometry of the calibration, sensor base tracking and IGS registration is seen in greater detail in figure 3.5³. The objective of the overall procedure is to find a way to relate range coordinates to MRI coordinates. The goal of calibration, which can take place outside of the OR and is illustrated in figure 3.5(a) and (b), is to relate the inner range-sensor reference to the milled divots on its outer side-plates, or sensor base. The goal of sensor-base tracking, as shown in figure 3.5(c), is to establish the position of these sensor base points in probe coordinates at time t_o in the OR, and use this information along with the calibration results to relate the inner range-sensor reference to probe space. The probe-image transformation, produced by the IGS patient registration, can then be used to relate range to MR coordinates, as featured in figure 3.5(d).

To simplify the formalism adopted for this section, I use the letters R , P and M to designate *range*, *probe* and *MR* coordinates. Furthermore, t_c and t_o correspond to time instants coinciding with *calibration* and the intervention in the *OR*. Next, the symbols $\mathbf{p}_{B,i}$ ($i = 1 \dots 9$) and $\mathbf{p}_{T,j}$ ($j = 1 \dots 15$) refer to points coinciding with milled divots on the sensor *base* and calibration *tool*, respectively. The superscripts of these points indicate the coordinate frame and time instant considered, where relevant. Finally, a transformation from a coordinate frame A to frame B is represented by $\mathbf{T}_A^B$.

The calibration procedure (refer to figures 3.5 (a) and (b), 3.6 and 3.7) consists of the following steps, assuming that the sensor is brought to *home position*, $y = 0$, after each range image acquisition (as currently programmed):

1. *Linear positioner calibration.* The relation which maps (i, k) pixel coordinates to (x, z) in the range-sensor is established by Vitana Corp. and assumed accurate. However, the accuracy of the y -coordinate relies on an assumption of constant speed during range image acquisition, as well as a consistent acceleration pattern

³The detection or computation of a set of points in a given coordinate system is shown as a dashed line, whereas a transformation computation between two coordinate systems appears as a full line

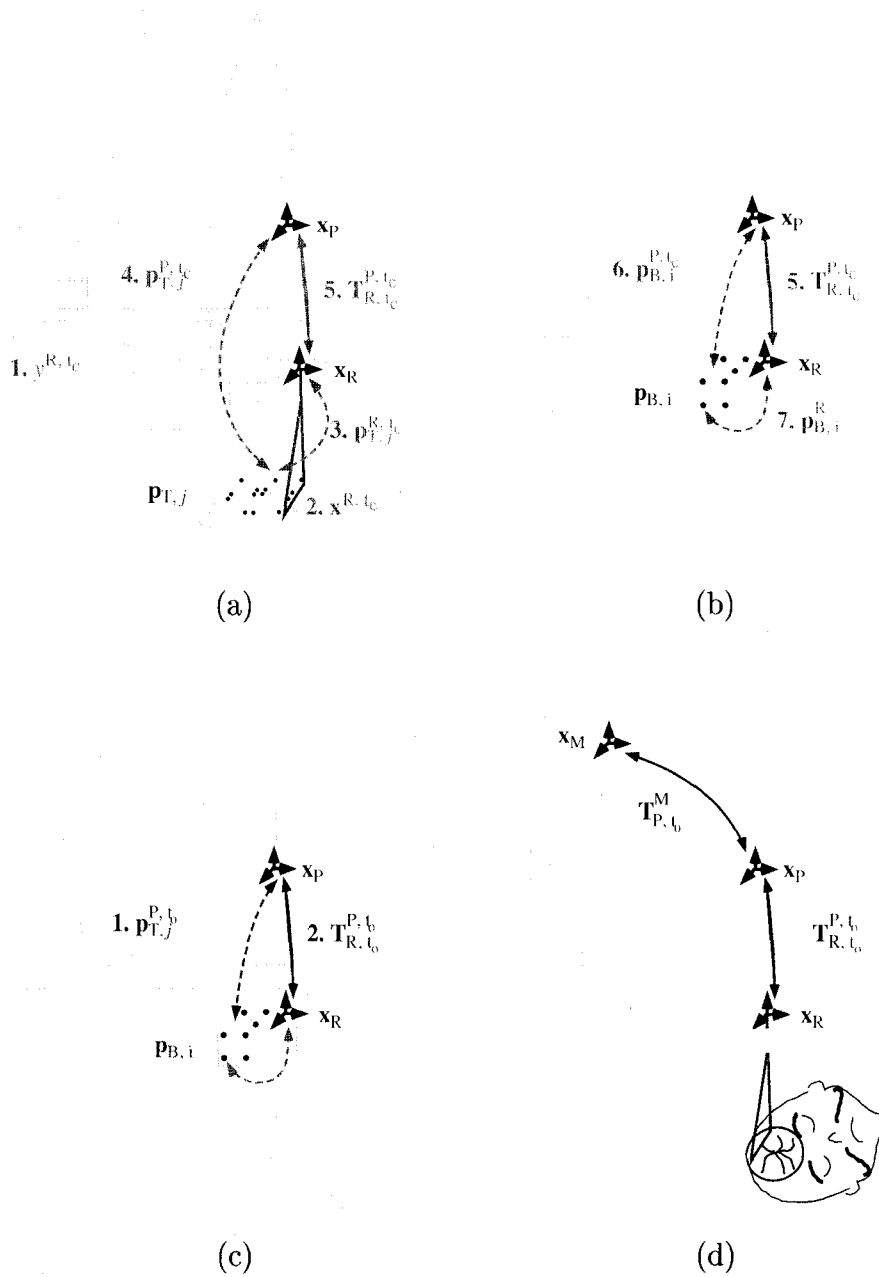


Figure 3.5 : *Range to MRI transformation computation. At $t = t_c$, the calibration procedure first relates range to probe space (a), then determines the sensor base points in range space (b). (c) Sensor base tracking determines sensor base points in probe space at $t = t_o$ and finds the range-probe transformation in the OR. (d) The final step uses this and the probe-MR transformation to map range to MR.*

from one range image acquisition to the next. Our experience has shown that both assumptions were undermined by the manner in which the positioner is integrated with the range-sensor setup, i.e.: communicating by serial port. This positioning error was demonstrated and corrected with a calibration experiment with the positioner calibration plate shown in figure 3.6(a), which features a row of machined divots at a $10\text{mm} \pm 0.01\text{mm}$ spacing.

In particular, the time delay between the trigger signal and the acquisition of the first laser profile, initiated by a return signal from the positioner controller, varied sufficiently to introduce inaccuracies in the y -position of up to 4mm. This issue was alleviated by operating the positioner at a much slower speed than originally intended, typically 1mm/sec rather than 24mm/sec, which reduced the effect of a variable delay to less than 1mm. In many range-sensors, communication between a positioner and sensor would not occur over a serial port, but rather on an internal computer bus, thereby obviating this time delay issue and allowing faster acquisition speeds.

2. *Range-sensing of a “wedding cake” (i.e.: non-coplanar) calibration tool.* The wedding cake tool features an array of hemispherical milled holes of 2mm radius, corresponding to the size of the probe tip, as shown in figure 3.6 (b)⁴, resulting in a range image such as seen in figure 3.7(a). Furthermore, the range-probe transformation can benefit from increasing the volume spanned by range-probe point pairs, by using the same plate at two different positions, as shown in figure 3.7(d), while maintaining the sensor in place throughout. This refinement is made possible by the 135 mm depth of field of the range-sensor, or in other words its ability to sense objects placed anywhere from 107 to 242 mm from the outer casing of its optical assembly, or at a distance from 187 to 322 mm from its optical origin.

3. *Automatic identification of hemispherical pits $\mathbf{p}_{T,j}^{R,t_c}$ (see figure 3.7) within the*

⁴Note one missing pit, to break the symmetry and to disambiguate the orientation of the tool within the range image.



Figure 3.6 : *Precision machined tools for calibration: (a) positioner calibration plate and (b) “wedding cake” sensor calibration tool.*

range image of sensor calibration tool, assuming a hemispherical template of 2mm radius, by *template matching* [328], i.e.: finding points whose local neighbourhood correlates well with this template. Everywhere in the range image a correlation value is computed, and subsequently thresholded, in order to retain only those points whose neighbourhood strongly agrees with the template (using a histogram of these values to select the threshold). The correlation computation is made tolerant of false reflections which can occur within the pits and which would otherwise corrupt the results, by considering also the intensity profile coinciding with the range image, as shown in figure 3.7(b): points of low intensity are discounted from the neighbourhood correlation computation. A *connected component analysis* [328] follows, in order to establish the contiguous regions of high agreement, and within each region, I find the point of *maximal correlation value*⁵, which is labelled a pit coordinate $\mathbf{p}_{T,j}^{R,tc}$. Typical results are

⁵The software eliminates false positives (such as occurring near discontinuities within the cali-

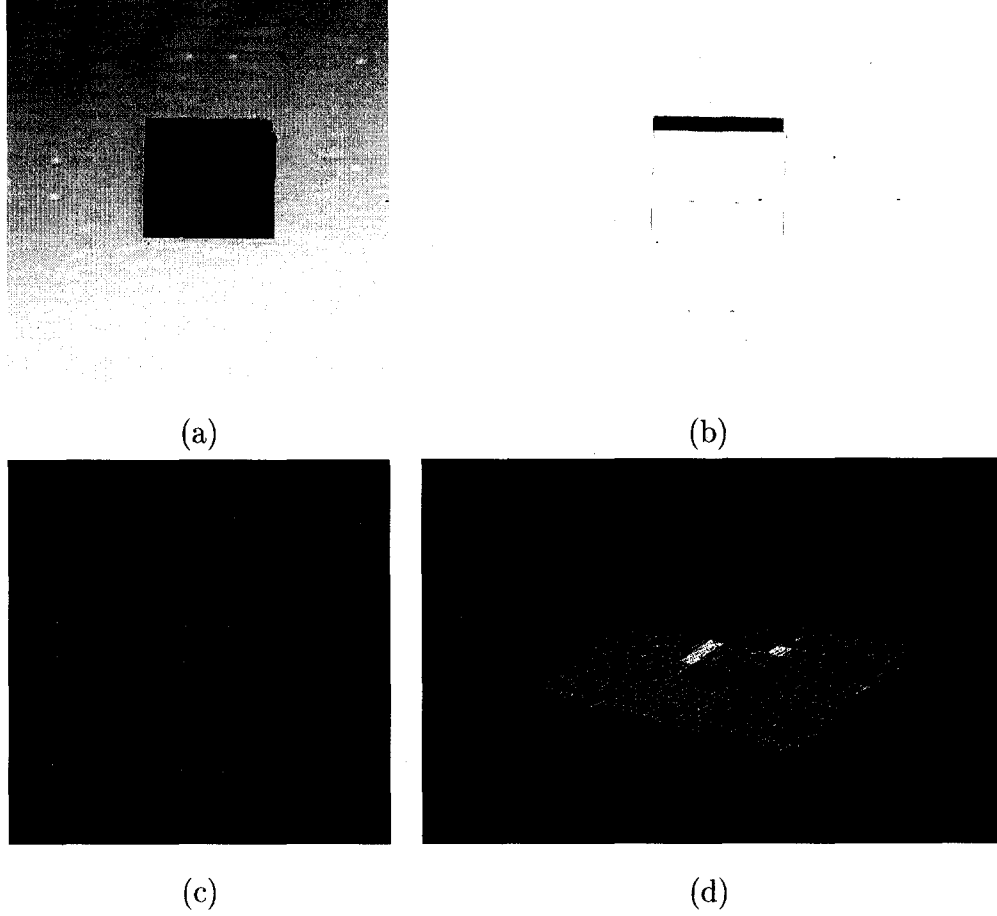


Figure 3.7 : *Illustration of calibration procedure: (a) range image of calibration plate, displayed as a grey-level; (b) intensity image overlapping range image, used to refine pit detection; (c) detected pits in range image; (d) range images comprising the 2-level calibration procedure.*

shown in figure 3.7(c).

4. *Probe-space identification of pits* in the tool, $\mathbf{p}_{T,j}^{P,tc}$, by insertion of the probe tip.
In the case of the 2-level procedure illustrated in figure 3.7(d), both range and probe points are identified for the calibration tool at the first level, prior to moving the tool to the second level.
5. *Registration of range pits* $\mathbf{p}_{T,j}^{R,tc}$ *with probe-space coordinates of the pits* $\mathbf{p}_{T,j}^{P,tc}$:
the transformation from probe reference to internal range reference (for the
bration tool) on the basis of known distances to neighbouring pits, and selects between two or more
strong correlates in close proximity to each other, based on the highest correlation value.

calibration procedure) $\mathbf{T}_{P,t_c}^{R,t_c}$ is determined from the point pairs by Arun's SVD technique [8].

6. *Probe-space identification of sensor base points*, i.e.: the 9 milled divots in outer side-plates bolted to the sensor: $\mathbf{p}_{B,i}^{P,t_c}$.
7. *Apply probe-range transformation to sensor base points*:

$$\mathbf{p}_{B,i}^R = \mathbf{T}_{P,t_c}^{R,t_c} \mathbf{p}_{B,i}^{P,t_c}. \quad (3.4)$$

Note that the position of the sensor base points $\mathbf{p}_{B,i}^R$ in range space shouldn't depend on time, since the side-plates are bolted to the sensor. Consequently, this calibration procedure need not be repeated often.

The sensor base tracking procedure in the OR at t_o is as follows:

1. *Identify the location of pits in sensor side-plates in probe space*, to characterize the new sensor base position: $\mathbf{p}_{B,i}^{P,t_o}$.
2. *Match probe-space pits with the points $\mathbf{p}_{B,i}^R$* to determine the probe reference-range transformation $\mathbf{T}_{R,t_o}^{P,t_o}$ for the configuration at time t_o .

Finally, the range data comprising surface S_B are then related to MRI-space anatomical surfaces by considering the probe-MRI transformation computed in the IGS patient registration, along with the probe-range transformation:

$$\mathbf{x}_{S_B,t_o}^M = \mathbf{T}_{P,t_o}^M \mathbf{T}_{R,t_o}^{P,t_o} \mathbf{x}_{S_B,t_o}^R = \mathbf{T}_{R,t_o}^M \mathbf{x}_{S_B,t_o}^R. \quad (3.5)$$

In general, the range-probe transformation $\mathbf{T}_{R,t_o}^{P,t_o}$ can change over time (particularly if the sensor is retracted out of the way of the surgeon in between range scans), necessitating further sensor base tracking, whereas the probe-MR transformation $\mathbf{T}_{P,t_o}^M$ will tend to be stable due to the use of a Mayfield clamp which generally keeps the patient's head in place, although some slippage may occur and can be ascertained with a *tracker* [397] affixed to the patient's head.

Moreover, since none of these procedures is error-free, and these errors are cumulative, the final result can be further refined with a rigid-body Iterative Closest Point

(ICP) registration. The goal ultimately is to establish a baseline for cortical movement by referring the MRI-based pre-operative brain surface to range coordinates via $\mathbf{T}_{R,t_o}^M$. In addition, corresponding skin patches in range and MRI spaces are related by the same transformation. The final transformation between skin patches, as determined by an ICP procedure, can be exploited to refine $\mathbf{T}_{R,t_o}^M$. Such a refinement should in general reduce the component of the local registration error which is normal to the skin surface. The details of the ICP registration implementation are presented in chapter 4.

3.4 Validation of the Range-MR Transformation Procedure

3.4.1 Materials and Methods

Validation of all procedures involved in relating range to MR is based the skin phantom [28] shown in figure 3.8(a). This phantom features 12 adhesive torus-shaped fiducials [169] (each one consisting of a sealed disk, containing contrast gel, with an central opening for probing), which are visible under both CT and MR. I opted for a CT validation, because the phantom features metal wire in its core. The CT voxels are 0.89 mm in the x and y directions, and 1 mm in z. Finally, the localizing device used in this study is a FARO arm [432, 95], which was discussed in chapter 1. The overall range-CT transformation makes use of the prior IGS probe-CT transformation based on 7 pairs of fiducials $\{\mathbf{x}_{F,k}^C \leftrightarrow \mathbf{x}_{F,k}^P\}$ visible in probe space, as shown in figure 3.8(b). The IGS transformation results are illustrated in figure 3.8 (c) and (d) for left and right side studies respectively. Note that for both left and right side studies, one fiducial is taken from the opposite, less exposed, side in order to better anchor the transformation (so that the homologous pairs span a volume, rather than a near-planar section).

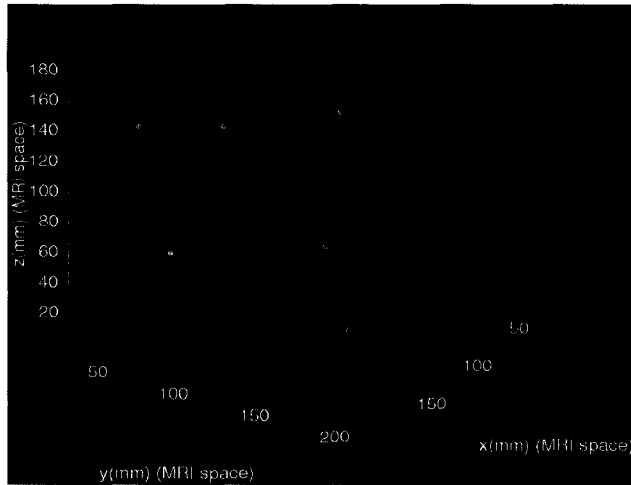
The decision to use a least-squares or robust [330] transformation computation depends on the assumption made about the localization error distribution amongst



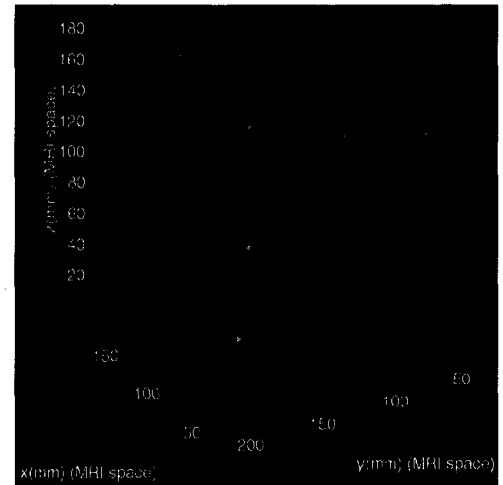
(a)



(b)



(c)



(d)

Figure 3.8 : (a) *Poupoune*, the exotic skin phantom, prior to validation study, and (b) in validation configuration. Illustration of least-squares IGS image-probe transformation and the 7 homologous point pairs used for left (c) and right (d) side studies: CT space (green) and probe space fiducials transformed to CT space (light blue), overlaid on outer skin phantom voxel.

homologous pairs. If the error is assumed to be normally distributed and without severe outliers, then robust regression can cause the transformation to degrade at the anatomical points coinciding with the worst residuals by throwing out useful information, which is highly undesirable in this application (a least squares estimator is optimal under Gaussian noise). If severe outliers are likely, then the use of least-median-of-squares or any other robust scheme simply discards the erroneous homologous pair from the final transformation computation. One would expect a careful manual IGS registration not to suffer from spectacular outliers, but the robust approach could conceivably be refined by a test on the residual magnitude to discriminate between a typical registration error and a truly outlying homologous pair.

The validation study consists of a comparison of

- *computed CT fiducial positions*: manually identified positions of the fiducials in the range image, transformed to CT space with the range-CT (or range-MR) procedure presented in the previous section, against
- *manually identified fiducial positions in CT*.

Typical validation results are illustrated in figure 3.9 and are presented succinctly in table 3.1. While an automatic localization of the fiducials in both spaces may have been desirable (from a repeatability standpoint), to do so reliably was difficult to achieve. In the both the left and right skin phantom range images, at least one of the fiducials present is self-occluding (the wall of the torus obscures part of its central opening, from the range-sensor field of view). In fact, even the manual identification of all three fiducials in each range image was difficult and subject to some interpretation (which would have been difficult to replicate automatically). Moreover, to my knowledge, there exists no published automatic localization technique for the IZI Med fiducial in tomographic data [431]⁶.

⁶One published technique for fiducial localization in CT and MR makes use of a simpler fiducial, designed with automatic localization in mind [409].

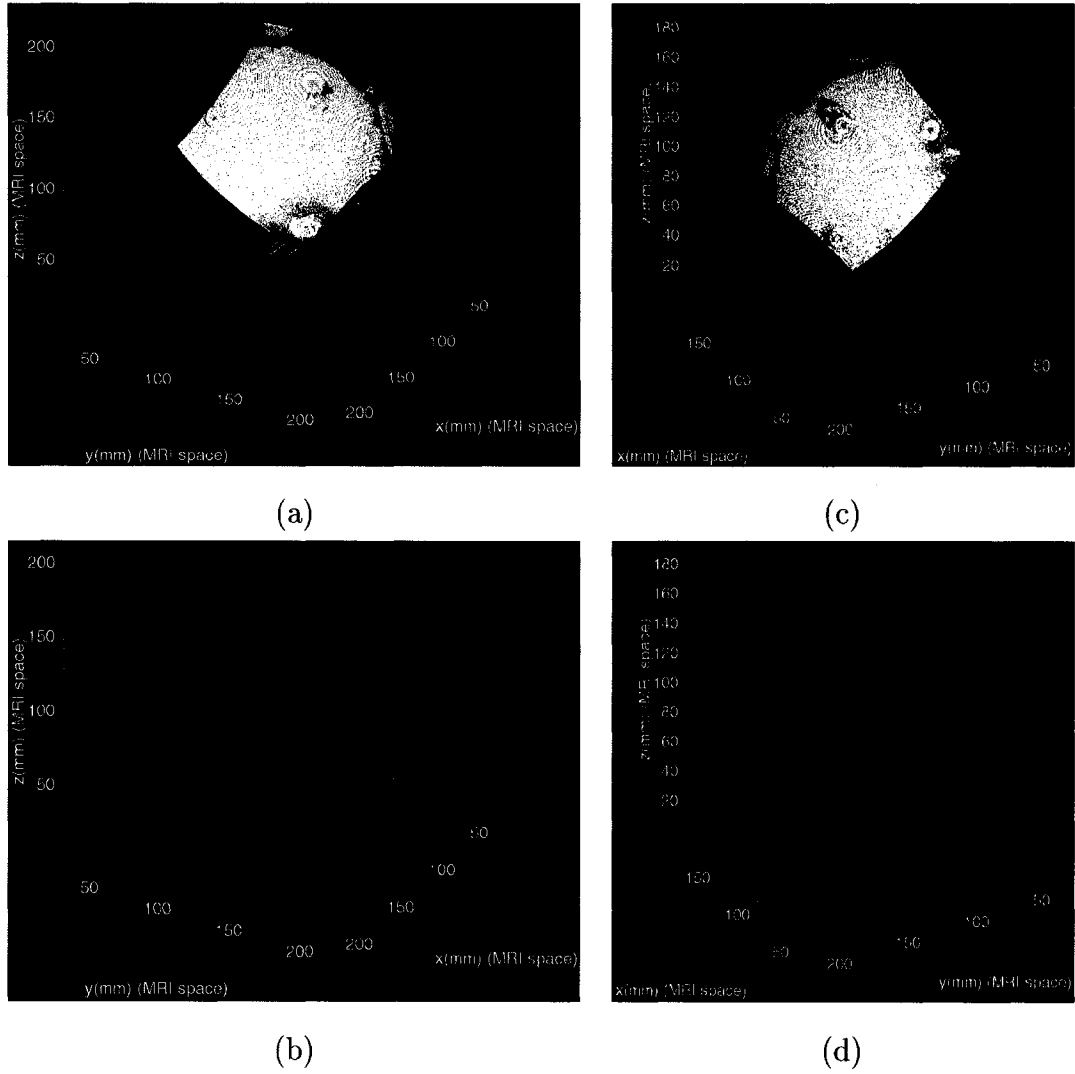


Figure 3.9 : *Range-CT transformation procedure results (trial 1): (a) left side range surface data transformed to CT space, overlaid on CT skin phantom voxels; (b) original 7 CT space (green) and transformed range space (pink) fiducials of left side study; (c) transformed right side range surface data overlaid on CT skin voxels, as in (a); (d) CT space and transformed range space fiducials of right side study, as in (b).*

3.4.2 Validation Results

Table 3.1 sums up the average values for disparities between identified and computed fiducial positions in CT space for the left and right side studies. Also featured in table 3.1 is a description of the magnitude of the transformation between the two sensor base positions at t_c and at t_o (during validation), from side-plate probe measurements. This information is compiled to establish whether the disparity is sensitive to the magnitude of the translation or of the rotation between the two sensor positions, measured in probe space. The transformation is expressed as two values: the magnitude of the calibration-OR translation vector $\|\mathbf{t}_c^o\|$ in mm and the angle of rotation θ_c^o in degrees, about the rotation axis⁷.

3.4.3 Discussion

These results are encouraging and they are consistent with the localization technology used for implementing the IGS registration and with the inherent uncertainty in assessing the position of the torus-shaped fiducials (and therefore the uncertainty in the validation). Nevertheless, the goal of this procedure should be to achieve a characterization of the range-MR transformation to the nearest millimeter. The studies cited in § 1.3 suggest that substantially better accuracy is achievable with optical localization and more compact fiducials, possibly on the order of 1mm. In particular, the Rohling [326] study notes that the FARO estimates distances to within 0.5mm, with a spread of 0.5mm, whereas the Optotrak is generally within 0.15mm, with a spread equivalent to that value. In my procedure, the repeated use of the probe causes probe-based localizing errors to be cumulative: this cumulation would be far less significant with optical tracker accuracy.

These results can be interpreted in the context of the formal relationship between the *fiducial localization error* (FLE) and the *target registration error* (TRE) explored

⁷Recall that a rotation can be expressed in terms Euler Symmetric Parameters, i.e.: as an axis of rotation (k_x, k_y, k_z) and an angle of rotation θ about this axis. The conversion from this representation to an orthonormal matrix proceeds through quaternions, is closed-form, and appears in [82] (pp.55-56).

Left side study						
		Trial 1	Trial 2	Trial 3	Trial 4	Trial 5
$\ \mathbf{t}_c^o\ $ (mm) / θ_c^o (deg)		293.4 / 11.0	139.1 / 11.1	201.6 / 12.9	265.5 / 53.8	167.0 / 14.4
least squares	ϵ_{rms} (mm)	2.48	2.48	2.98	2.57	1.99
	d_{surf} (mm)	0.97	0.88	0.99	0.79	0.61
robust	ϵ_{rms} (mm)	2.01	2.41	4.31	2.69	1.74
	d_{surf} (mm)	0.91	0.91	1.53	1.01	0.67
least squares, ICP	ϵ_{rms} (mm)	2.44	2.47	2.71	2.57	2.09
	d_{surf} (mm)	0.46	0.47	0.47	0.47	0.46
robust, ICP	ϵ_{rms} (mm)	2.11	2.16	3.90	2.51	2.01
	d_{surf} (mm)	0.46	0.46	0.49	0.46	0.45

Right side study						
		266.7 / 1.5	141.9 / 1.0	231.1 / 2.7	243.5 / 43.8	191.2 / 24.4
$\ \mathbf{t}_c^o\ $ (mm) / θ_c^o (deg)		266.7 / 1.5	141.9 / 1.0	231.1 / 2.7	243.5 / 43.8	191.2 / 24.4
least squares	ϵ_{rms} (mm)	2.14	2.48	3.39	2.43	2.19
	d_{surf} (mm)	0.97	0.52	0.89	0.91	0.63
robust	ϵ_{rms} (mm)	2.11	3.08	5.40	2.71	3.08
	d_{surf} (mm)	0.66	0.92	1.65	0.86	0.72
least squares, ICP	ϵ_{rms} (mm)	2.12	2.12	2.92	2.33	2.13
	d_{surf} (mm)	0.41	0.41	0.45	0.42	0.41
robust, ICP	ϵ_{rms} (mm)	2.20	2.49	4.27	2.42	2.78
	d_{surf} (mm)	0.41	0.40	0.45	0.41	0.40

Table 3.1 : Trial results for calibration and sensor base tracking procedure, for range images of the left and right sides of the head. Rigid body transformation by least-squares and by robust regression, with and without ICP refinement based on the skin identified in CT and range spaces. The value ϵ_{rms} represents the RMS error (in mm) between CT fiducial positions and their homologues in the range image transformed to CT space by $\mathbf{T}_{range}^{CT}$. The value d_{surf} is the RMS distance (in mm) between the transformed range points and the skin surface in CT space, as read from a precomputed distance map. The symbols $\|\mathbf{t}_c^o\|$ and θ_c^o indicate the magnitude of the translation and rotation between the sensor base positions at t_c and t_o .

by Maurer, Fitzpatrick et al. [349, 205, 247, 248, 113]. Langron and Collins [205], based on prior work of Sibson [349], showed that fiducial registration errors (FRE) add in quadrature. Maurer [247] showed through simulations that the relationship between TRE and FLE is linear: $TRE = k \cdot FLE_{eff}$, where in turn $FLE_{eff} = \sqrt{[(FLE_1^2 + FLE_2^2)]/2}$, where FLE_1 and FLE_2 represent the errors in the two spaces. Furthermore, the slope TRE/FLE is related to the number of fiducials and is sensitive to the fiducial configuration: TRE/FLE varies between “very accurate” (0.8) and “poor” (> 2.5). These simulations were complemented with phantom experiments and clinical trials, with both sets of results somewhat worse than predicted by simulations, which was abscribed to other factors such as geometrical image distortion and patient motion. Fitzpatrick et al. [113] showed that the *fiducial registration error* (FRE) is related to FLE by the following relation: $\langle FRE^2 \rangle = (1 - 2/N) \langle FLE^2 \rangle$ and is chi-square distributed, where N is the number of fiducials. For a large N , FRE tends to FLE, and its chi-square distribution tends to a normal distribution.

The relevance of the FLE-FRE-TRE analysis to the present study is that a lower bound on the final TRE of the calibration and sensor base tracking procedure can be estimated on the basis of the FLE or TRE associated with each step. Assuming that $TRE \approx TLE$ for the sake of simplicity, we can analyze the cumulative error as follows.

1. The error relating the range reference to the sensor base (outer plates) depends on the localization error of the pits in the calibration tool, in range and Faro probe space. These are separate processes which can be simplified as normally distributed, 0-mean, with standard deviation proportional to the FLE, and multiplicative, so that the variance of the cumulation of the two localizations simplifies to the sum of the two variances, and FLE_{tot} is given by the RMS sum of the two localization errors:

$$FLE_{tot} = \sqrt{(FLE_1^2 + FLE_2^2)} \quad (3.6)$$

This formula is consistent with the idea of adding error sources in quadrature.

2. Assuming a probe error FLE_P of 0.5 mm, consistent with the Rohling study [326],

and a range $FLE_R \approx \sqrt{(0.5^2 + 0.5^2 + 0.5^2)} = 0.86$ mm, where a value of 0.5 mm is also presumed for the y -axis (due to limitations of the linear positioner calibration procedure), I assess the accuracy of the calibration plate localization as $TRE_{R-P,t_c} = \sqrt{(0.5^2 + 0.86^2)} \approx 1.0$ mm.

3. The accuracy with which the sensor base divots can be located with respect to the range reference, at calibration, is a function of the accuracy TRE_{R-P,t_c} with which range space is related to probe space, and once again of the accuracy FLE_P with which the FARO probe can locate each divot. Adding them in quadrature once again: $FLE_{B_P,t_c} = \sqrt{(1.0^2 + 0.5^2)} \approx 1.12$ mm.
4. The accuracy of the sensor-base tracking procedure, from t_c to t_o , is then the TRE incorporating the previous FLE_{B_P,t_c} and the usual FARO-related $FLE_P = 0.5$ mm for localizing the sensor base divots at t_o :

$$TRE_{R-P,t_o} \approx FLE_{tot,B_P} \approx \sqrt{(1.12^2 + 0.5^2)} \approx 1.22 \text{ mm.} \quad (3.7)$$

5. If one factors in the registration error of the probe-MR transformation, which is in the 1.5 mm range, then a reasonable lower bound for the final range-MR registration error is

$$TRE_{R-M,t_o} = \sqrt{(TRE_{R-P,t_o}^2 + TRE_{P-M}^2)} \approx \sqrt{(1.22^2 + 1.5^2)} \approx 1.93 \text{ mm.} \quad (3.8)$$

This value is not incompatible with the values observed in table 3.1, if one also factors in the uncertainty involved in determining the position of the fiducials in the range and CT spaces in the context of the validation itself. Worth taking into consideration are the Zinreich [432] and Alp [3] studies emphasizing limitations in assessing fiducial position in a tomographic volume, particularly for the torus-shaped fiducial. In retrospect, compact ball-shaped fiducials would probably have been more accurately identified in both the CT and range data, alleviating the self-occlusion problem.

It is worth re-emphasizing that the calibration and sensor-base tracking procedure is made necessary by the clinical constraint that implantable fiducials, which could

be automatically detected, were not used at the MNI during this research. If they had been available, and if their design were optimized for automatic detection in MR and range data, then the need for this procedure would have been obviated, and far better accuracy would be achievable because the transformation between range and MR would no longer be limited by the accuracy of the probe-MR transformation, where manual localization error becomes an issue.

The worst disparity in both left and right studies was associated with the same inaccurate calibration procedure, namely trial 3. In turn, there doesn't seem to be a correlation between the value of the registration error and the magnitude of the transformation, as reported here: the magnitudes of the translation $\|t_c^o\|$ and rotation θ_c^o of trial 3 do not stand out. The inaccuracy of trial 3 can probably be ascribed to less than optimal use of probe data: in particular, in all trials but this one the probe coordinates that were recorded were consciously taken in the middle of the range of values seen while manually pivoting about a divot point. The probe data would have benefitted from a spatial averaging strategy, if a reasonably stable position were ascertained.

Robust statistics yielded comparable results for most of the left side trials, but significantly worse results for trial 3 for the left side study, and most of the right side trials. The homologous pairs which determine the various transformations comprising the final range-MR relation appear to preclude a severe outlying pair, since the systematic removal of information in the robust method [330] tends to make the registration degrade. This finding also underscores the importance of the geometric configuration of homologous point-pairs [247]. In short, unless a test is feasible to separate a severe outlier from "typical" registration error, least-squares generally yield more predictable results and should be adopted.

Finally, the respective skin surfaces were brought to within 1mm of each other with the calibration-based procedure, and this distance could be reduced further with a subsequent ICP refinement. However, the improvement in the fiducial registration, as measured here, was marginal. This result suggests two possible explanations: that *much of the mis-registration occurred along the plane normal to the skin surface*, for

which there is no obvious explanation in the procedure itself, or that there may also be some *limitations in the validation due to the uncertainty in localizing the fiducial positions* in range and CT data. Conceivably, a large part of this uncertainty may lie in assessing the position of the fiducial on the skin within the range image: a pixel offset $(\delta i, \delta j)$ in fiducial localization would project mostly onto the tangent plane of the skin rather than normal to it.

Future directions for this research include a repetition of the validation with compact fiducials, optical tracking, a more refined approach to gathering probe data, and an adaptation of the sensor for intra-surgical use. In the latter case, the sensor and positioner would have to be mounted on an articulated arm (or robot) which could repeatably lock into two positions: *scanning* and *retracted* (out of the surgeon's way). If these two positions were repeatable to a fraction of a millimeter, subsequent sensor base tracking procedures would be obviated. Otherwise, a simple repetition the sensor base tracking and range-probe transformation computation would update the range-MR transformation and allow the surgeon to proceed.

3.5 Conclusion

This chapter presented a method for easily and cheaply determining the 3D shape of the exposed brain surface, using a laser range-sensor. Alternately, an active shape-from-stereo technique could have been substituted, particularly in the context of a binocular surgical microscope. I also discussed a procedure for relating 3D range data to pre-operative MR data by a rigid transformation, based on 3 stages: a calibration stage to assess the relationship between sensor base and range coordinates; a sensor base tracking stage to determine the sensor base position in probe space while in the OR, and in the process relate probe to range coordinates, and a final transformation between range and MR coordinates on the basis of the IGS patient registration results. A final refinement with ICP-based skin registration was also proposed. The validation made use of a skin phantom featuring adhesive torus-shaped fiducials. While the validation involved manual identification of the fiducials in both range and MR data, the

transformation procedure itself was fully automatic and reproducible. These studies suggested an average disparity obtained with a FARO arm and torus-shaped fiducials that is in the 2-3 mm range, although this value can degrade if the probe-based localization is jittery. A substantial improvement in the results is achievable with optical tracking, with more compact fiducials and perhaps with an improved probe sampling strategy. Better results still could be obtained if implanted fiducials were used, provided that they were optimized for automatic localization in MR and range data, thereby obviating the need for a probe-based calibration procedure. Future directions for research include a re-enactment of this validation with the suggested improvements and the mechanical adaptation of the sensor for use in the OR.

Chapter 4

Non-rigid Surface Displacement Estimation

4.1 Motivation and Overview of Non-rigid Cortical Motion Estimation

This chapter presents a method for estimating brain shift in the OR, from one or more range images of the brain as in figure 3.1. Brain shift estimation requires that a baseline for this movement be first assessed. This is achieved by applying the rigid range-MRI transformation, computed as presented in chapter 3, to the brain surface identified in the MRI volume. Various components of a method for semi-automatically identifying the cortical surface in MRI are described in chapters 5, 6 and 7.

Brain shift at the exposed cortical surface, measured with respect to the rigidly transformed pre-operative brain surface identified in MRI, can be characterized as non-rigid surface motion captured by the initial range image, as well as those displacements captured in subsequent range images. This surface motion is estimated with a non-rigid Iterative Closest Point (ICP) [33, 209] registration technique. As discussed in § 2.3.2, ICP techniques start from a rough initial alignment between two surfaces S_A and S_B , and iteratively use the set of *closest point pairs* (as opposed to features of similar shape as characterized by differential geometry) between the two surfaces

to determine a new incremental transformation. The incremental transformation $\mathbf{T}_k$ corresponding to the k th iteration is then applied to one of the surfaces (which may already be transformed), namely $\mathbf{T}_{k-1}S_B$, gradually improving its alignment with S_A . In most applications, the transformation $\mathbf{T}_k$ consists of an incremental rotation $\mathbf{R}_k$ and translation $\mathbf{t}_k$ which minimize the least-squares k th iteration registration error:

$$\min_{\mathbf{R}_k, \mathbf{t}_k} \sum_{i=1}^N \|\mathbf{x}_{A_i} - (\mathbf{R}_k \mathbf{x}_{B_i} + \mathbf{t}_k)\|^2. \quad (4.1)$$

The iterative process is stopped when this error is less than some preset tolerance.

In contrast, the use of this technique to characterize brain shift is based on $\mathbf{T}_k$ representing a local piecewise polynomial function: a *smoothing surface spline*, as suggested by the discussion in § 2.2.2. It is worth noting that thin-plate splines [96, 39] as well as volumetric B-splines [91] were also considered. However, thin-plate spline interpolation is better suited to *sparse* homologous pairs, rather than dense range and MR data: a small error in dense homologous pair data, in conjunction with exact interpolation, leads to wild oscillations in the resulting interpolant. On the other hand, volumetric approximating B-splines were used effectively by Szeliski and Lavallée for volumetric anatomical warping [382], in conjunction with an innovative ICP registration technique. In the context of this application however, each range image describes the outwardly visible cortical surface, as opposed to fully three-dimensional tomographic information. Therefore, the displacement implicit in each range image can effectively and quickly be characterized by a piecewise function over a 2D domain: i.e., a surface spline. Here I employ a smoothing spline, to make this characterization tolerant of small misregistrations amongst dense closest point homologous pairs. The implementation details of this approach are presented in the following section.

Finally, although it is conceivable that an active surface model framework could be used here for characterizing brain shift, at the cost of replacing the initial MRI-based image function by one driven by the range image sequence, I prefer an approach that restricts the surface tracking to the *domain of the range image*. By deferring the computation of the motion of the surface outside the range domain (i.e.: not visible through the craniotomy) to an ulterior FEM stage, one can estimate this motion in a

manner which accounts for *realistic* material properties, in contrast to the arbitrary elastic properties characteristic of most surface models.

In summary, the choice of a surface spline-based ICP technique is justified by its suitability to non-rigid motion estimation, in contrast to approaches based on features and global shape. Moreover, this technique is better suited to tracking an open range image (which does not require prior identification within a tomographic volume) than an active surface model [14]. The main drawback of such an iterative technique is the requirement of a rough initial alignment, but this is provided by the range-MR transformation procedure described in chapter 3. A method employing a combination of dense surface points and sparse features, either curve or point-based [244], could also have been employed if comparable sparse features could be identified reliably in range and MR spaces. A fully automatic feature extraction from range data would likely be complicated by the intra-surgical presence of cortico-spinal fluid and blood at the exposed brain surface, which could occlude the most reliable, sulci-based, features, while most gyri may be too flat to produce discriminating sparse features.

The rest of this chapter is organized as follows. Section 4.2 presents the implementation details of the proposed iterative surface registration procedure. This procedure features some efficiency innovations, in the form of a pre-registration processing stage of the semi-automatically identified brain surface in the MRI volume, described in § 4.2.1, and an extremely fast surface spline implementation, presented in § 4.2.2. Taken together, these refinements make each iteration essentially *instantaneous*. Special considerations for the non-rigid ICP technique of range and MR-based surfaces are further emphasized in § 4.2.3. Section 4.3 presents validation studies featuring a new and original elastically deformable brain-shaped phantom, as shown in figure 4.1, which is based on Polyvinyl Alcohol Cryogel (PVA-C) [65]. Further validation studies make use of an analytical deformation function applied to the brain surface extracted from the anthropomorphic digital head phantom of Collins et al. [74]. This phantom is also used in tandem with a MRI simulator in the context of semi-automatic brain surface identification technique proposed later, and is described in detail in § 7.1.

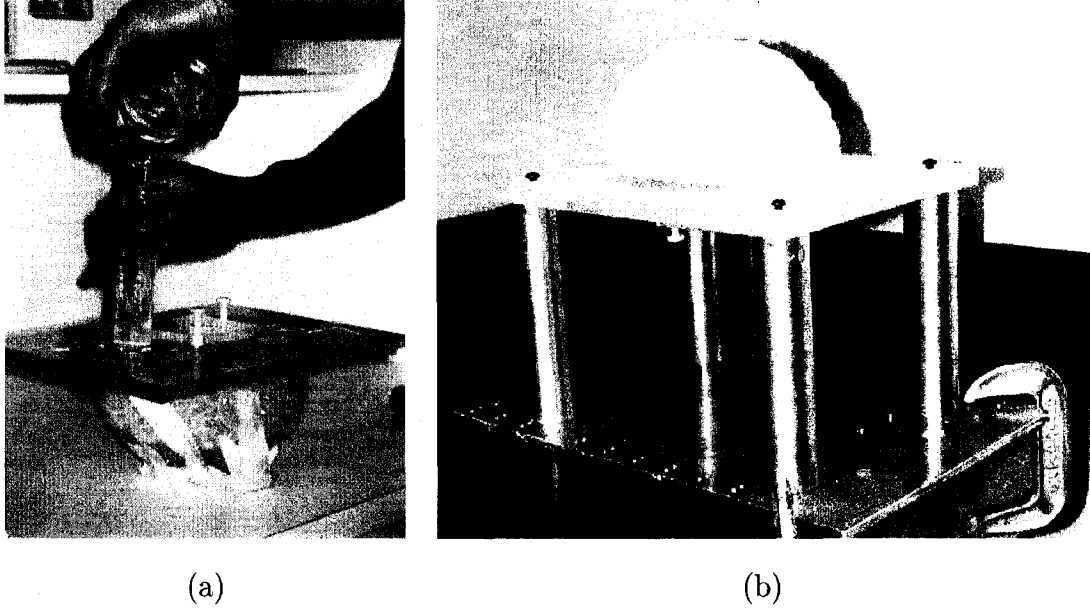


Figure 4.1 : *Elastic PVA-C brain phantom: (a) liquid stage during construction
(b) final result after freezing and thawing.*

4.2 ICP Registration Technique: Implementation Details

4.2.1 ICP Homologous Pairs by Look-up: the Closest Point Map

The iterative registration technique proposed here bears comparison to that of Lavallée and Szeliski [209, 382] (see also § 2.3.2), in that the volume spanning one of the surfaces is preprocessed to accelerate its registration with the second surface. Lavallée and Szeliski represent the volume containing the first surface by an octree structure, then determine the closest surface point to each corner of each cube in this representation, compute the resulting distance to it, and finally interpolate the resulting sparse distance map with trilinear splines. When a second surface falls within the mapped volume, each point on it inherits a distance value to its closest homolog of the first surface. The closest homolog $\mathbf{x}_{A,i}$ on surface S_A of point $\mathbf{x}_{B,j}$ on the surface

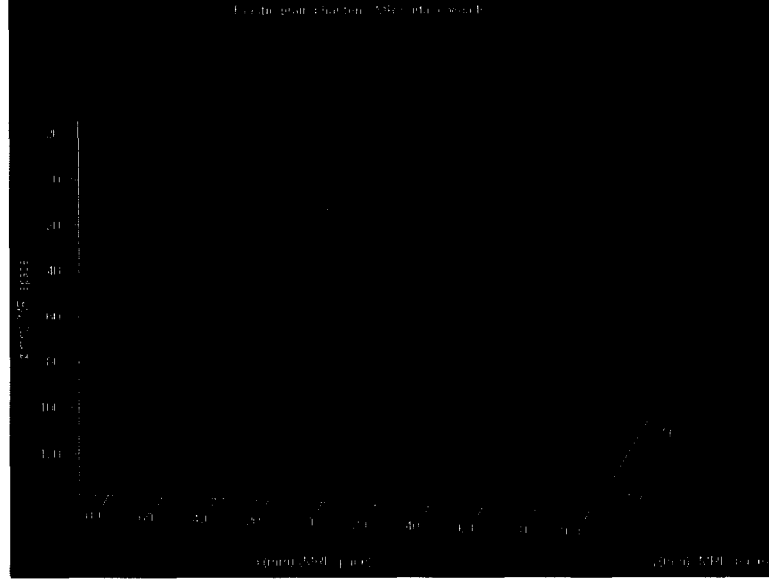


Figure 4.2 : *Elastic brain phantom surface from MRI scan (rotated for better visibility). The outline of the plate used to implement deformation can be discerned within the boundary of the phantom.*

S_B , possibly transformed by $\mathbf{T}_k$, is by definition

$$\left\{ \mathbf{x}_{A,i} \equiv \inf_{\forall \mathbf{x}_{A,m} \in S_A} \|\mathbf{x}_{A,m} - \mathbf{T}_k(\mathbf{x}_{B,j})\| \right\}. \quad (4.2)$$

The resulting optimal transformation between two surfaces is that which best accounts for these distances, or which best registers the second surface to the first, in a weighted least-squares sense. Because Lavallée’s preprocessing produces only a distance map, but does not provide explicit information about mutually homologous pairs, the subsequent registration must still resort to a search method, such as the Levenberg-Marquardt approach [225], to determine the optimal transformation.

In contrast, the method presented here incorporates a processing stage applied to the identified brain surface in the MRI volume, based on a propagative distance-to-surface algorithm called the *Fast Marching method* (see sections 6.4.1 and 7.4), and which is modified to produce a *Closest Point Map* as well. The Fast Marching method is used to compute distances from a given surface, such as shown in figure 4.3 (a) for tri-planar images of the elastic brain phantom. The Closest Point Map is an adaptation of this method, whereby the *label of the closest surface point*, from

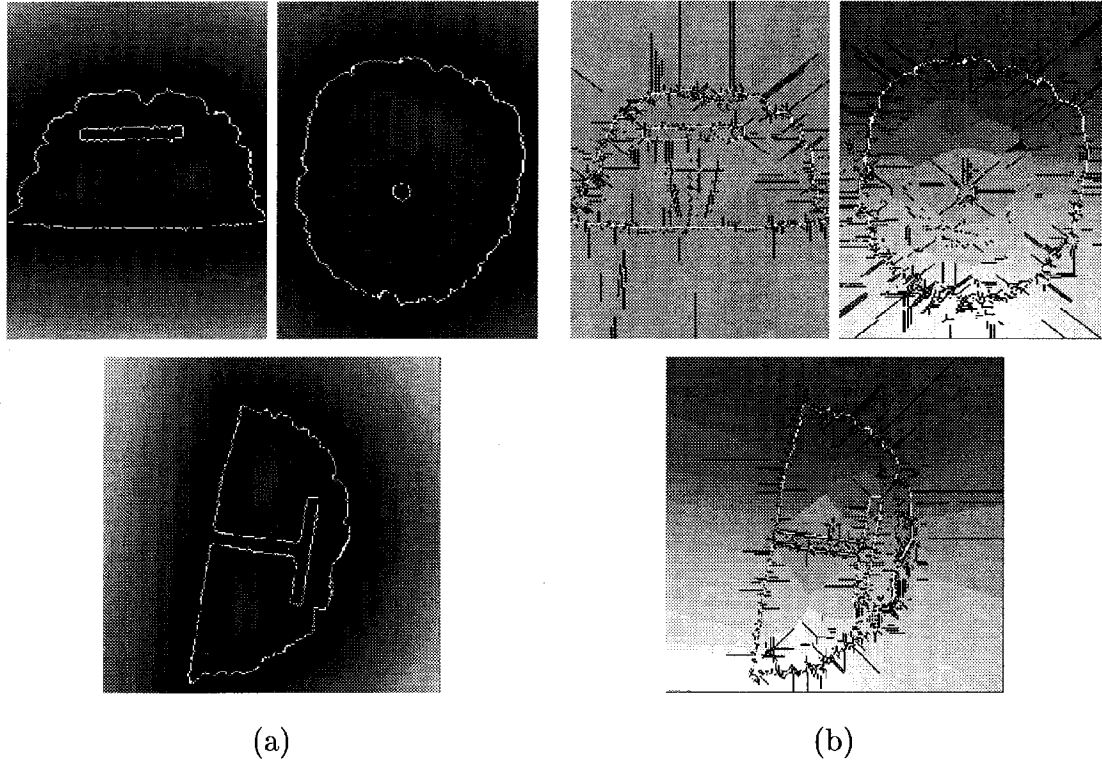


Figure 4.3 : (a) Distance and (b) Closest Point Maps computed from the surface of the elastic phantom (white), and shown in tri-planar views, mapped to a grey level. In (b) ambiguous voxels, possessing more than one closest surface point, are shown in black (see discussion in § 7.4).

which the distance is computed, is also stored. For every voxel in MRI space, this map associates the label of the brain (or phantom) surface point closest to it ¹. This concept is shown in figure 4.3 (b) for tri-planar images of the elastic brain phantom, with labels mapped to a grey-level and with ambiguous voxels (more than one closest point) identified in black.

The flowchart of the registration algorithm, shown in figure 4.4, illustrates how the range image of the brain, transformed to MR space (based on the procedure discussed in the previous chapter), is non-rigidly registered to the MR brain surface on the basis of the Closest Point Map. For a range image of the brain brought into rough alignment with the brain surface in the MR volume, the closest surface

¹The external examiner of this thesis has indicated that a similar registration approach was developed independently by another group [123].

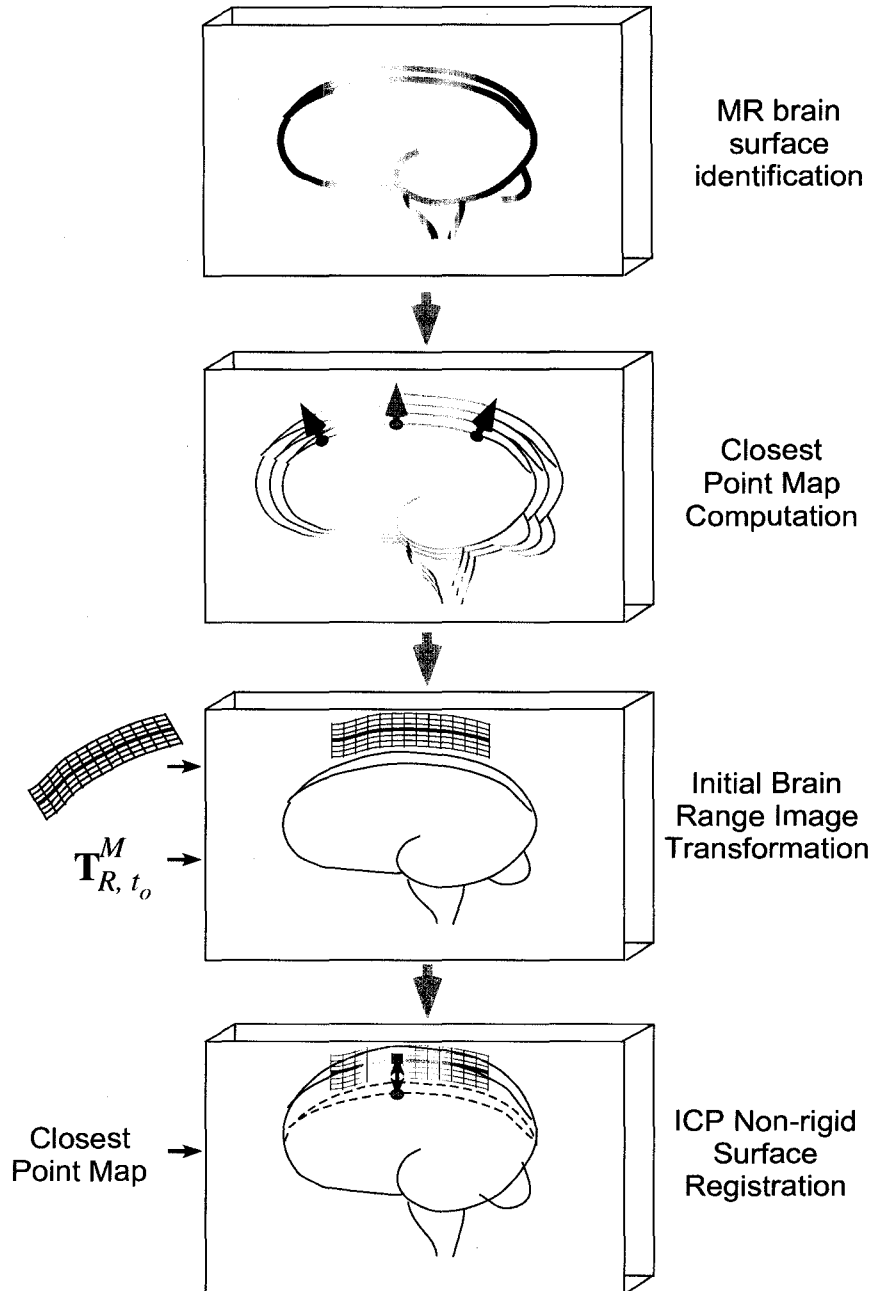


Figure 4.4 : *Symbolic description of non-rigid surface registration. Top: the labels of the identified brain surface voxels are symbolized by colour. Row 2: the brain surface labels are propagated to all non-surface voxels. Row 3: the range data, of initially arbitrary orientation, are transformed to MR space on the basis of the sensor-base tracking procedure. Bottom: each range point inherits a Closest Point Map label from the MR voxel with which it coincides, thereby specifying its MR homologous point.*

point to it *is simply determined by look-up*, thus constituting a homologous pair. In other words, given partially registered range data, each point of which falls near a voxel in MR space, the closest MR brain/phantom surface point is read as the Closest Point Map label associated with that voxel. These homologous pairs are used either to determine an incremental rigid transformation that globally improves the registration of the two surfaces, or to produce a densely estimated displacement function which is refined to yield a smoothed surface displacement function over the range domain.

The characterization of non-rigid motion becomes much more efficient if its starting point is a dense set of homologous pairs, as opposed to a distance map which imposes a costly minimization at each ICP iteration. Each rigid transformation iteration can exploit a closed-form computation method[8, 158, 106, 408], any of which requires explicit point-pairs, not distances between them. The Closest Point Map makes such rigid-body transformation computations instantaneous, since the calculation simply involves the determination of homologous point-pairs by table look up. Each non-rigid iteration can make use of a dense vector displacement function obtained by look-up, and as will be seen in the next section, a highly efficient numerical scheme for smoothing this function with surface splines.

The justification for emphasizing computational efficiency here is two-pronged: *clinical acceptability* and the inherent *temporal under-determination* in estimating non-rigid motion. Clinical acceptability hinges on allowing the surgeon to easily deploy the system, get the surface displacement quickly, use this surface information with efficient FEM schemes [9, 25] to estimate volumetric displacement, and proceed with the surgery. The under-determined nature of the inverse problem of estimating non-rigid motion, from a time sequence of homologous point-pairs, implies that it makes sense to enable the acquisition of a large number of displacement maps over time if needed, just as it does to use dense point displacement information to characterize non-rigid motion between any two successive time points.

Assuming that the baseline cortical surface (and brain volume) is updated over time, based on successive non-rigid registrations with intra-operative range data, the same non-rigid registration method may also be applied to the updated cortical

surface (in MRI space) at t_n , at the cost of recomputing the Closest Point Map by the Fast Marching method. This computation can be performed in a matter of a few minutes (typically less than 5 minutes on a 866 MHz Pentium III machine), while the current brain shift information is rendered on the screen, and while FEM-based volumetric motion estimation proceeds from it and is provided to the surgeon. This scenario assumes that a surgeon desiring brain shift information wants it as quickly as possible (which justifies preprocessing in the form of a Closest Point Map), but once it is provided, would not need it to be updated until some time later; i.e.: it assumes that the frequency of use would not prohibit such processing. Alternately, the registration can proceed from the original preoperative cortical surface at t_0 to the points of the range image at t_{n+1} . However, while iteratively registering each new range image at t_{n+1} with the cortical surface at t_n (for which a new Closest Point Map is computed) is more expensive computationally, it makes better use of the temporal resolution available to estimate non-rigid motion.

4.2.2 Exploiting Regular Range Data Spacing: Surface Splines by Recursive Filtering

The regular spacing of the range data is exploited by using extremely efficient *recursive smoothing splines* [403, 404] to smooth the displacement function. These surface splines draw on the work of Unser et al [403, 404], which expresses the fitting of interpolating, smoothing or approximating splines as a sequence of *filtering* stages, all of which can be implemented as recursive filters, provided that the data samples exist at regular positions. As mentioned earlier, this is the case for the range surface points. Recall from the previous chapter that the range-sensor incorporates a uniformly spaced CCD array, and that its horizontal pixel coordinates i is also uniform in sampling. Recall further that image acquisition points j along the positioner axis are also characterized by uniform spacing, as ensured by the linear positioner calibration procedure outlined in § 3.3.

For this work, I adopt a smoothing spline approach, whose first stage is a convo-

lution with a smoothing kernel²

$$S_{\lambda}^n(Z) = 1 / \left(B_1^n(Z) + \lambda(-Z + 2 - Z^{-1})^{\frac{n+1}{2}} \right), \quad (4.3)$$

where $B_1^n(Z)$ is the *indirect B-spline transform* given by Unser [403], n is the order of the spline fitting and λ is the regularization parameter of the smoothing. This stage produces B-spline coefficients and is followed by a convolution with the indirect transform $B_1^n(Z)$ to yield the smoothed output. The filters can simply be cascaded to implement smoothing in 2D. A first-order fit is chosen, i.e.: $n = 1$, to limit overshoots due to a higher-order continuity constraint, such as near the boundary of the craniotomy, outside of which the displacement should be null after rigid registration. The shape of the cranial opening does not exactly coincide with a rectangular range image domain, with the result that part of the range image coincides with areas covered by the skull, where the displacement cannot be estimated and is therefore undefined. If a surface of order 2 or 3 had been chosen, the smoothed displacement function might appear even more continuous, but the transition from the area of *measured* cortical displacement, to areas covered by the skull where displacement is *undefined* or *presumed null*, would lead to inappropriate continuity-preserving oscillations.

Given the smoothing kernel expression 4.3, setting $n = 1$, and multiplying by indirect transform $B_1^1(Z) = 1$ [403, 404], the recursive spline filter is characterized by the transfer function:

$$H(Z) = B_1^1(Z)S_{\lambda}^1(Z) = 1 / (1 + \lambda(-Z + 2 - Z^{-1})) . \quad (4.4)$$

This filter can be broken down into factors in Z and Z^{-1} , and implemented as separate anti-causal and causal filters respectively [316]. Furthermore, this is a first-order recursive IIR filter, so that the transfer function in Z -space translates into local operations in 1D involving the *current and immediately neighbouring* pixels: i , $i + 1$ (anti-causal) and $i - 1$ (causal), independently of the value of the smoothing parameter λ . The same reasoning applies to smoothing along the j -pixel axis.

²Note: Z here relates to the Z -transform, not to be confused with the depth axis of the range-sensor.

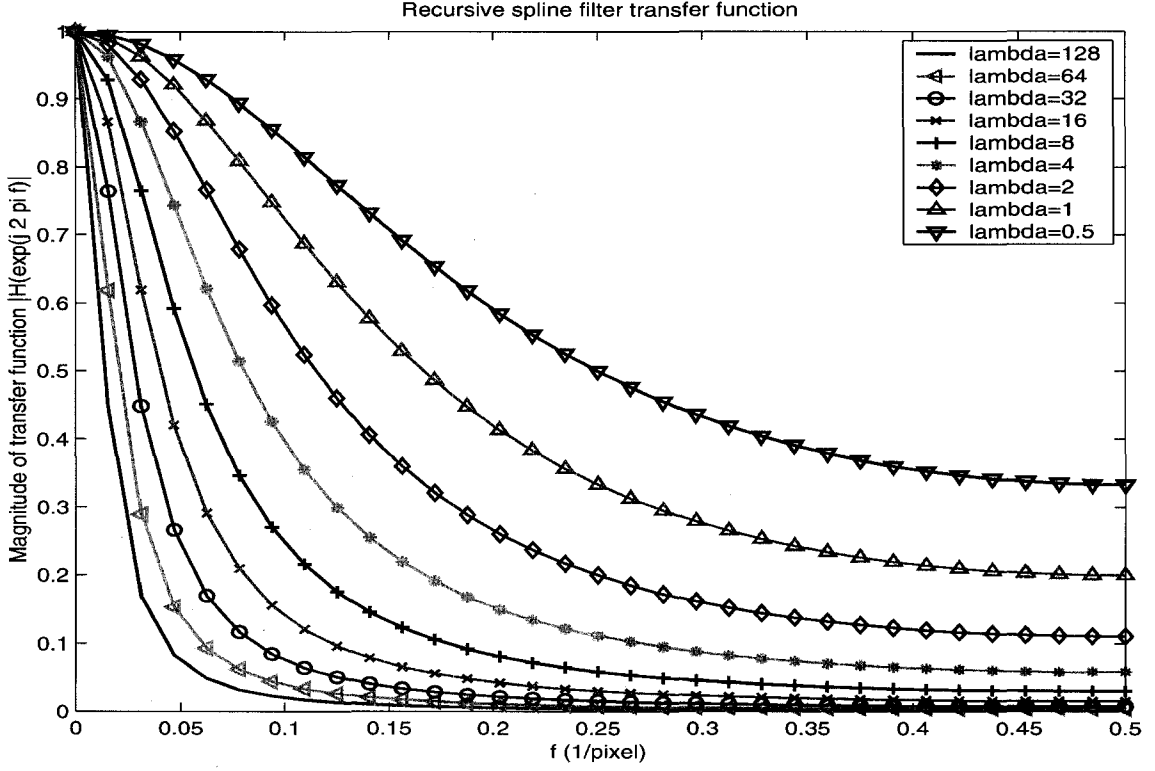


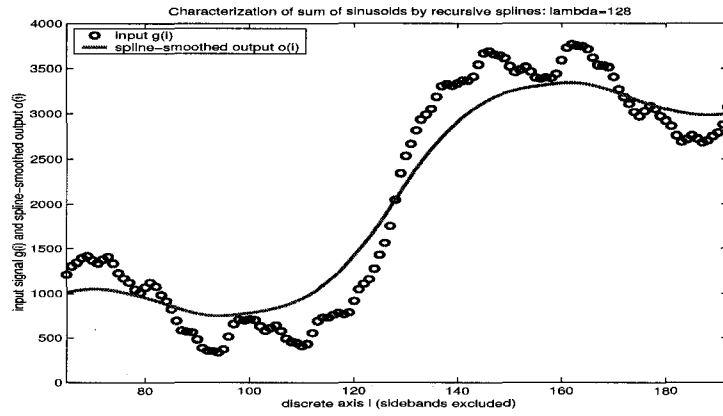
Figure 4.5 : *Frequency of response of the recursive approximating spline filter.*

The frequency response of the transfer function in expression 4.4 can be obtained by substituting $e^{j2\pi f}$ (where here $j = \sqrt{-1}$) for Z , and plotting the magnitude as a function of λ , as in figure 4.5. A further illustration of the smoothing filter response, as scale decreases from $\lambda = 128$ to $\lambda = 1$, is given in figure 4.6, with an input signal consisting of the sum of a ramp and of 7 sinusoids:

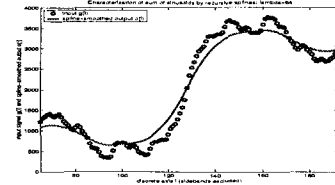
$$g(i) = 1024 \sin(4\pi i/N) + 512 \sin(8\pi i/N) + 256 \sin(16\pi i/N) + 128 \sin(32\pi i/N) + 64 \sin(64\pi i/N) + 32 \sin(128\pi i/N) + 16 \sin(256\pi i/N) + 4096i/N \quad i = 1, N, \quad (4.5)$$

where N is the number of samples, in this case 256. This example reflects the progressive characterization of the displacement function which is achievable as the scale factor is decreased over successive ICP registration iterations.

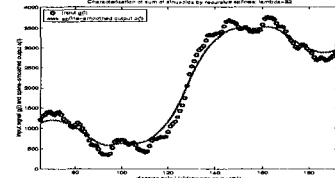
Note that the sidebands coinciding with $i < 64$ and $i > 192$ in figure 4.6 are not considered, to eliminate border effects. The inputs used in the recursion relations over these sidebands are the values at borders $i = 64$ and $i = 192$, propagated outwards. This practice is also adopted when processing the range image with cascaded spline filters. In this case, the displacement function defined over the 256-by-256 domain



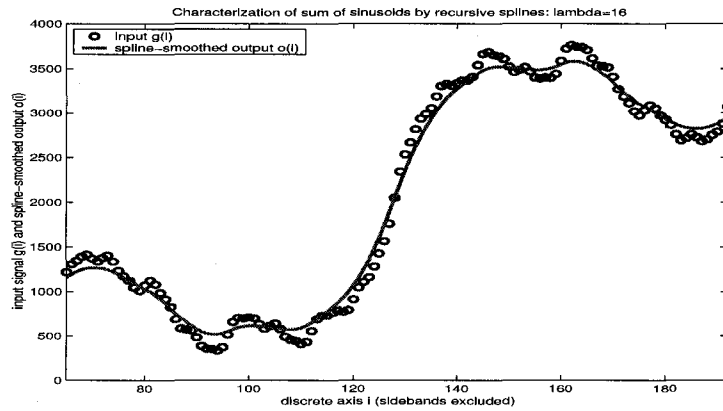
(a) $\lambda = 128$



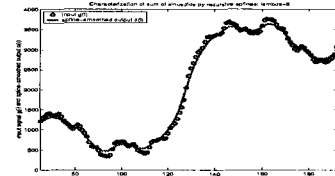
(b) $\lambda = 64$



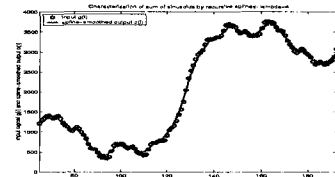
(c) $\lambda = 32$



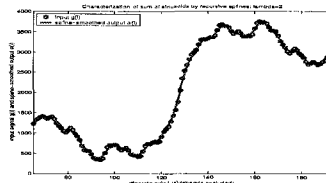
(d) $\lambda = 16$



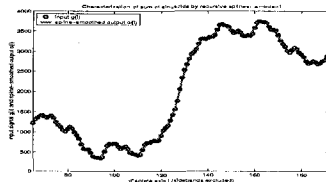
(e) $\lambda = 8$



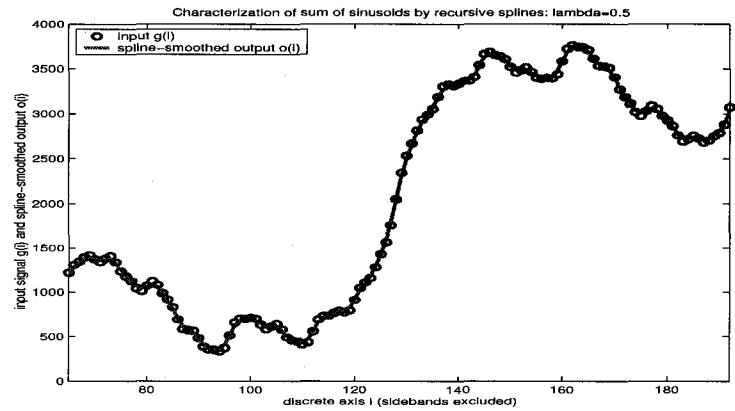
(f) $\lambda = 4$



(g) $\lambda = 2$



(h) $\lambda = 1$



(i) $\lambda = 0.5$

Figure 4.6 : Illustration of recursive spline filter response, given an input consisting of a sum of a ramp and 7 harmonic sinusoids, as λ decreases from 128 to 0.5.

of the range image is actually imbedded in a function defined over a larger, 512-by-512 domain, with border values also propagated on the outer bands. The function defined over the larger domain is then input to the 2D smoothing filter. It is likely that similar logic would also have been necessary with other spline implementations, including thin plate splines, to deal with border effects.

Finally, in order to make the non-rigid motion estimation well-behaved, the registration features the following stages:

- A *rigid registration* stage: a series of instantaneous rigid ICP iterations, each of whose transformations is obtained by Arun’s closed-form, SVD technique [8], where homologous pairs are obtained by look-up from the Closest Point Map. This stage is halted when the average distance between the two surfaces is no longer improved significantly by another iteration (or if two consecutive iterations increase the average distance), as determined by a distance threshold (typically 0.01mm).
- A *non-rigid registration* stage: a series of non-rigid ICP iterations where the smoothing parameter λ is initially set very high, and is gradually lowered as the process iterates, progressively resolving finer-level motion detail. Just as for the rigid ICP stage, each iteration of the non-rigid surface motion estimation is essentially instantaneous, and the reduction of λ is triggered by a lack of improvement at the current level of resolution.

4.2.3 Special Considerations for Range-MR Non-rigid Surface Registration

When performing the non-rigid registration stage, each iteration may require some additional consideration, as illustrated in figures 4.7 to 4.10, in comparison to the rigid-body stage:

1. *Missing range data*: as illustrated in figure 4.7, a low signal can occur at some pixels in the range-sensor CCD, caused by an occlusion or artefact related to the

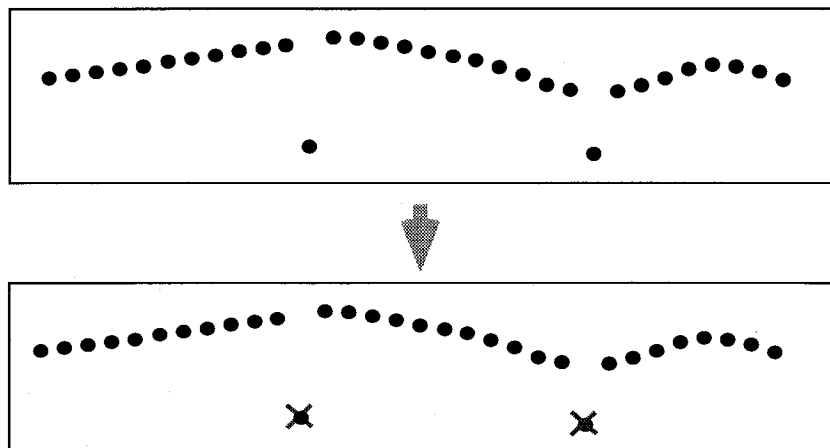


Figure 4.7 : *Range data featuring unreliable, low-signal pixels which are identified and eliminated from further consideration.*

reflectance of the imaged surface. Such a low signal value can be detected from the intensity profile of the range image. With this sensor, low reflected laser intensity is associated in practice a null z -value, leading to simple detection. Moreover, the intensity profile on the CCD is as accessible as the range profile, so that it is also possible to place a threshold on intensity. If the range image were characterized by noise impulses which were more difficult to detect, a formal impulse noise removal technique [64] could have been used. It would then have been simple to replace this void with an interpolated range value on the basis of neighbouring range values. However, given that the range data are considerably denser than MR cortical surface data, and that interpolated points can be viewed as more speculative than valid range points, the unreliable range pixels are simply identified and eliminated from further consideration.

2. *Difference in data density:* if the MR and range data were of equal density, the homologous pairs would directly determine a spline-based displacement function. However, because the range data are somewhat denser than the MR-based cortical surface points (65536 vs. 5782 points on the brain phantom), the range-MR homologous pairs are *many-to-one*. Consider the displacement implied by each homologous pair $\mathbf{d}_i$ in such a situation, as shown in figure 4.8: each such

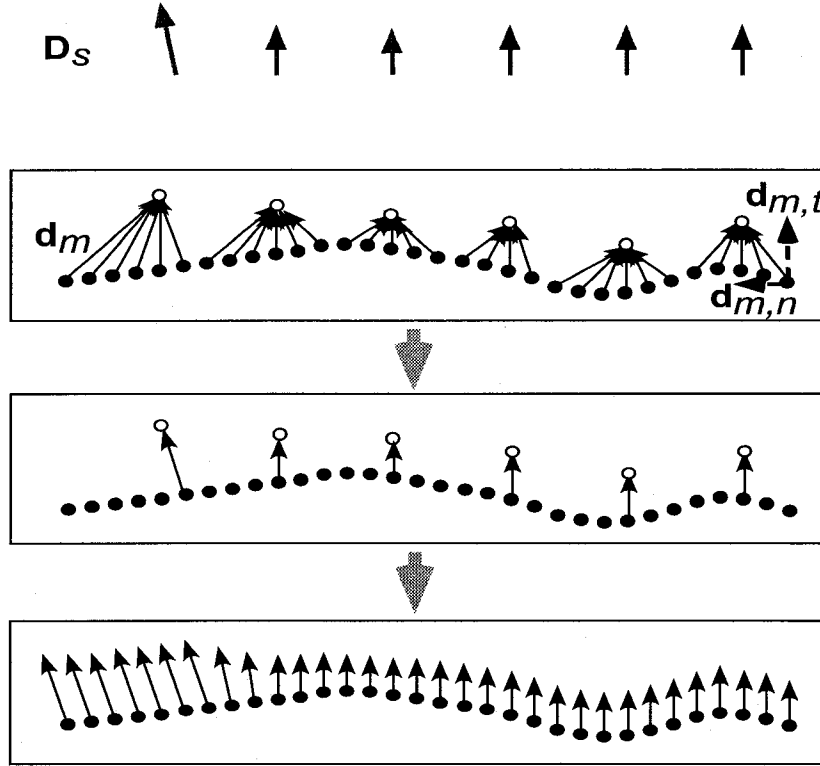


Figure 4.8 : *Dealing with many-to-one range-MR matches arising from unequal data density (MR points in white, range points in black). Top: true surface displacement $\mathbf{D}_S$; second row: displacement field implicit in many-to-one homologous pairs; third row: disambiguation by pruning all but mutually closest homologous pairs; and bottom: distance-weighted propagation.*

displacement can be idealized as the vector sum of a component $\mathbf{d}_{m,t}$ along the true displacement $\mathbf{D}_S(\mathbf{x}_A)$ between two surfaces, and an error component $\mathbf{d}_{m,n}$ normal to it.

In a rigid registration, the error components $\mathbf{d}_{m,n}$ would tend to cancel each other out because of spatial symmetry, provided that the spatial distribution of the denser points were uniform. Conversely, a bias could be introduced if the density of range data were monotonically increasing or decreasing along one or both axes. The main consequence of this geometry is that the rigid-body ICP convergence in this case would be slower than in the case of uniformly

dense data, because of the under-estimation of the component $\mathbf{d}_{m,t}$ along the true displacement. However, in the non-rigid case, consideration of all of the many-to-one homologous pairs, even under uniform spacing, would lead to a large erroneous displacement component $\mathbf{d}_{m,n}$, because of the strong influence of each homologous pair on the local displacement function. In other words, the consequence of the many-to-one homologous pairs caused by the difference in density is not just a slower convergence as in the rigid case, but spurious oscillations due to a large displacement component $\mathbf{d}_{m,n}$ normal to the true surface displacement $\mathbf{D}_S$.

This issue suggests a pruning of the many-to-one homologous pairs obtained by look-up, resulting in a set of *mutually closest pairs*. In order to define a displacement map everywhere in the range domain from the now-pruned range-MR displacement map, a distance-weighted propagation is exploited. This pruning and propagation operation remains far less expensive than a formal search-based ICP registration, and each iteration remains virtually instantaneous.

3. *Multiple and Spurious MR Closest Point candidates*: as will be discussed in detail in §7.4, it is geometrically feasible for a voxel to have more than one closest surface points, given a Closest Point Map computed from a surface consisting of a collection of voxels. In other words, even after the refinement in step 2, a many-to-one (MR-to-range) situation is possible, because of an artefact in the Closest Point Map. The possibility of more than one MR candidates, or even of a spurious closest homologous pair, must be addressed for each non-rigid iteration. The impact of any ambiguous or spurious match in each rigid iteration is negligible, given the weight of the many unambiguous homologous pairs determining 6 independent transformation parameters. But in the non-rigid case, each homologous pair has substantially more weight in determining, and adversely affecting, the local displacement function.

Simple disambiguation and validation of each pair are possible however, based on the assumption of *continuous displacement*, a constraint widely used in

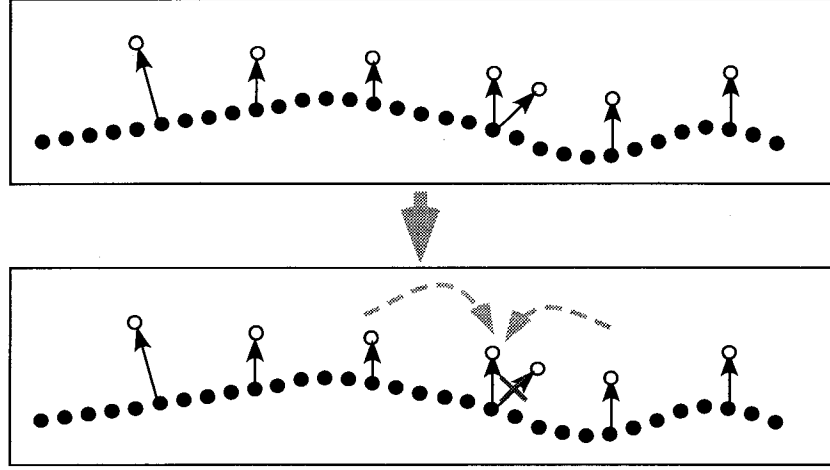


Figure 4.9 : *Multiple or spurious MR closest point: disambiguation from continuous displacement assumption, by neighbourhood estimate.*

machine vision to deal with ambiguity, for example in binocular correspondence [240, 12]. This assumption holds particularly true for a set of points arising from the same smooth anatomical surface. In simple terms, one expects the displacement implicit in a homologous pair $\{\mathbf{x}_{R,i} \leftrightarrow \mathbf{x}_{M,j}\}_l$ to be fairly consistent with the displacement of neighbouring homologous pairs, so that each displacement $\mathbf{d}_m = \mathbf{x}_{R,i} - \mathbf{x}_{M,j}$ can be compared with an estimate $\hat{\mathbf{d}}_{N(m)}$ computed from neighbouring unambiguous pairs (i.e.: one range point, one Closest Point in the MR cortical surface), as illustrated in figure 4.9. This test is a comparison between two displacement vectors in 3D.

If the objective is to choose the best among several candidates for a match, one can simply select the one which implies the displacement closest to the neighbourhood estimate. However, if the goal is generally to eliminate spurious matches, then every homologous pair must be validated on the basis of a comparison with the neighbourhood estimate. The measure of disagreement is the magnitude of the vector difference between the two displacements:

$$\epsilon_{N,m} = \|\mathbf{d}_m - \hat{\mathbf{d}}_{N(m)}\|. \quad (4.6)$$

The threshold for enforcing continuity can be determined from statistical measurements on the displacement difference $\epsilon_{N,m}$ between each homologous pair

and its neighbourhood. For example, one can tolerate a displacement difference, between a homologous pair and its neighbourhood, whose magnitude is less than the average difference plus some multiple n of the standard deviation of this difference; i.e.: we exclude a homologous pair $\{\mathbf{x}_{R,i} \leftrightarrow \mathbf{x}_{M,j}\}_l$ if

$$\epsilon_{N,m} > \bar{\epsilon}_N + n \times \sigma_{\epsilon_N} . \quad (4.7)$$

The choice of this multiple n is a trade-off between a desire for high density of displacement information, and a low tolerance for erroneous registration anywhere on the brain surface. This neighbourhood validation is based on a local average, as opposed to a more expensive median filter, and as shown in § 4.3 provides satisfactory results.

4. *Computing the inverse MR-range transformation:* the final result of the procedure described so far is a *forward range-MR transformation* $\mathbf{T}_R^M$, composed of a final multiplicative rigid transformation $\mathbf{T}_{R,r}^M$ which is the product of successive ICP transformation matrices³, followed by an additive spline-based non-rigid displacement function $\mathbf{T}_{R,n}^M$:

$$\begin{aligned} \mathbf{x}'_{R,i} &= \mathbf{T}_R^M(\mathbf{x}_{R,i}) \equiv \mathbf{T}_{R,n}^M(\mathbf{T}_{R,r}^M \mathbf{x}_{R,i}) \text{ where} \\ \mathbf{T}_{R,r}^M &= \mathbf{T}_{R,r,K}^M \mathbf{T}_{R,r,K-1}^M \cdots \mathbf{T}_{R,r,2}^M \mathbf{T}_{R,r,1}^M \text{ and} \\ \mathbf{T}_{R,n}^M(\mathbf{x}) &\equiv \mathbf{T}_{R,n,L}^M(\mathbf{T}_{R,n,L-1}^M \cdots (\mathbf{T}_{R,n,K+2}^M(\mathbf{T}_{R,n,K+1}^M(\mathbf{x})))) . \end{aligned} \quad (4.8)$$

The indices K and L indicate the final iterations of rigid (multiplicative) and non-rigid (additive) ICP registration. In turn, the following recursive expression further defines the l -th non-rigid transformation

$$\mathbf{T}_{R,n,l}^M(\mathbf{x}) \equiv \tilde{\mathbf{d}}_l(\mathbf{x} + \tilde{\mathbf{d}}_{l-1}) + \mathbf{T}_{R,n,l-1} , \quad (4.9)$$

where $\tilde{\mathbf{d}}_l$ represents the smoothed displacement at iteration l , evaluated at putatively transformed range point. In other words, the method described so far maps to MR space the range surface data $\mathbf{x}_{R,i}$, where $\mathbf{x}_{R,i} \in S_R$, in a manner which conforms to the pre-operative MR cortical surface data $\mathbf{x}_{M,j}$, where

³Homogeneous coordinates reduce the application of successive rigid body transformations to a sequence of matrix multiplications [82].

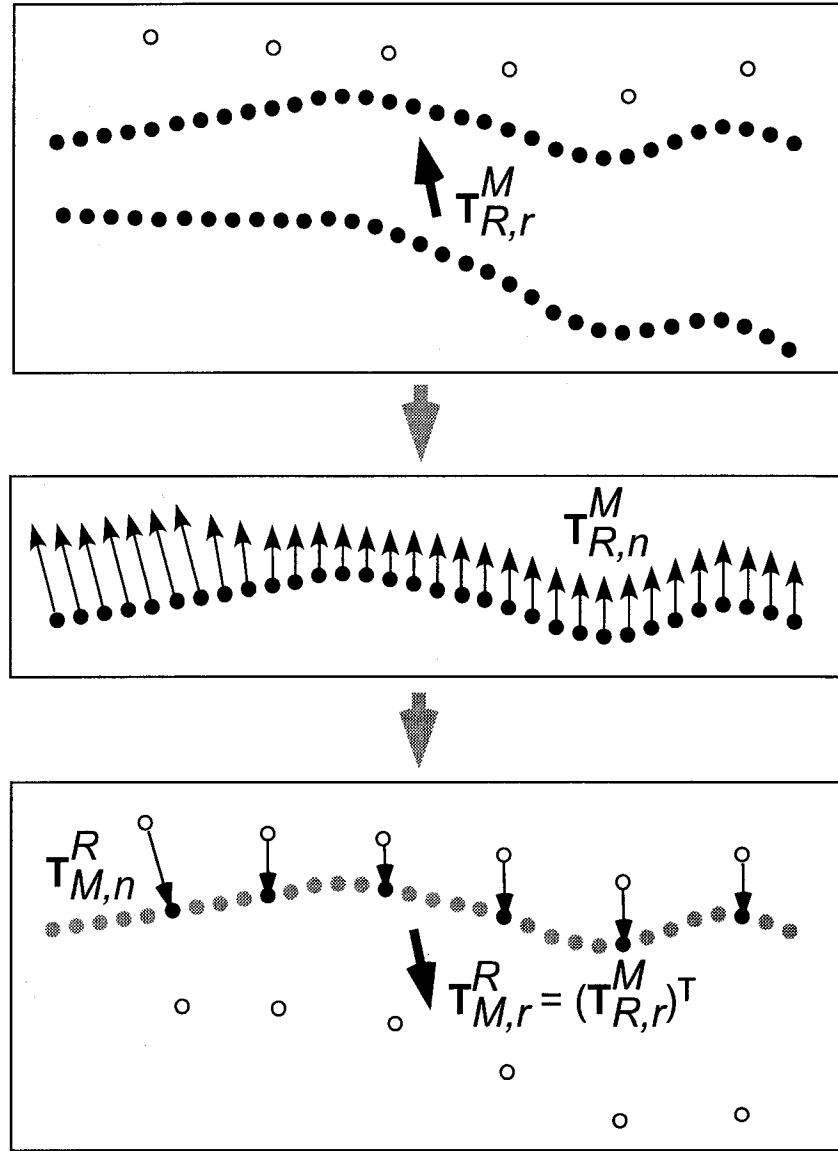


Figure 4.10 : Generating the MR-range transformation $\mathbf{T}_M^R$ from $\mathbf{T}_R^M$. Top: overall rigid range-MR transformation $\mathbf{T}_{R,r}^M$; middle row: overall non-rigid range-MR displacement $\mathbf{T}_{R,n}^M$. Bottom: the non-rigid and rigid components of the transformation which maps the MR brain surface to the range image are as follows: $\mathbf{T}_{M,n}^R$ is the negative of $\mathbf{T}_{R,n}^M$, and $\mathbf{T}_{M,r}^R$ is the transpose of $\mathbf{T}_{R,r}^M$.

$\mathbf{x}_{M,j} \in S_M$. This mapping is based on a sequence of additive spline-smoothed displacement functions, whose magnitude progressively diminishes over successive iterations, following a sequence of incremental multiplicative rigid-body transformations, also diminishing in magnitude over successive iterations. In fact, rather than construct the final transformation by adding or multiplying incremental transformations, it is preferable to consider the final set of homologous pairs, whose labels are retained and which by steps 1-3 are presumed one-to-one, and explicitly computing the final rigid transformation or smoothed distances from the final matched MR surface points to their range homologues. The latter method reduces the accumulation of registration errors over the sequence of iterations.

What is needed to characterize brain shift, however, is the *inverse transformation*: we must find the transformation $\mathbf{T}_M^R$ that non-rigidly maps the MR data $\mathbf{x}_{M,j}$, where $\mathbf{x}_{M,j} \in S_M$, to range space, in a manner that conforms to the shifted cortical range data. The inverse transformation of $\mathbf{T}_R^M$, as illustrated in figure 4.10, is the subtraction of the overall range-MR non-rigid displacement function followed by the application of the inverse, or transpose, of the overall range-MR rigid-body transformation:

$$\begin{aligned} \mathbf{x}'_{M,j} = \mathbf{T}_M^R \mathbf{x}_{M,j} &= \mathbf{T}_{M,r}^R \mathbf{T}_{M,n}^R (\mathbf{x}_{M,j}) \text{ where} \\ \mathbf{T}_{M,n}^R(\mathbf{x}) = -\mathbf{T}_{R,n}^M(\mathbf{x}) \quad \text{and} \quad \mathbf{T}_{M,r}^R &= \mathbf{T}_{R,r}^M{}^{-1} = \mathbf{T}_{R,r}^M{}^T. \end{aligned} \tag{4.10}$$

This approach to the computation of the inverse transformation pre-supposes the smoothness of the forward non-rigid mapping, such that the inverse is well-defined, and the absence of significant occlusions.

4.3 Non-rigid Registration Validation

4.3.1 Brain Shift Simulation by Realistic Elastic Brain Phantom

For the purpose of reproducing non-rigid cortical movement outside the OR, I have implemented a brain-shaped phantom with elastic material properties. I use a jello mold in the shape of the brain (currently available at the Barebones website [22]), into which we pour *PolyVinyl Alcohol Cryogel* (PVA-C), as illustrated in figure 4.1(a). The latter is a relatively *viscous liquid*, which upon freezing and thawing, becomes an *elastic solid* [65] as shown in figure 4.1(b). Furthermore, PVA-C can sustain several freeze-thaw cycles to acquire more rigidity, and the PVA concentration can also be manipulated to that effect, producing a Young's modulus E in the 0.1 to 0.7 MPa spectrum [379]. This is comparable to values published in the literature for gray and white matter material properties⁴.

This phantom rests on a support plate of plexiglass (usable within an MR scanner), and features a moving assembly consisting of a small disk and rod, also of plexiglass, imbedded within the elastic material. The other end of the rod (outside the elastic material) is attached to an external plate which rests on the opposite side of the support plate. The position of the moving assembly can be modified by means of some set screws pushing the external plate away from the support plate, thus drawing the embedded disk toward the support plate and triggering a compression of the elastic material.

Finally, because both the liquid and solid forms of PVA-C are translucent, and as this characteristic has an adverse impact on the range-sensor (leading to multiple reflections of each laser plane), the surface is spray-painted to make it opaque. Translucency is obviously not a characteristic shared by the human brain, and fig-

⁴Zhou [425] lists bulk moduli K of 0.219 GPa and 0.349 GPa, for GM and WM, as well as a Poisson's ratio ν of 0.4996 for each. These values translate to values of E of 0.525 MPa and 0.837 MPa respectively, based on the formula $K = E/[3(1 - 2\nu)]$ for isotropically elastic material [239]. The values found elsewhere in the literature are closer to the lower end of the PVA-C spectrum [261].

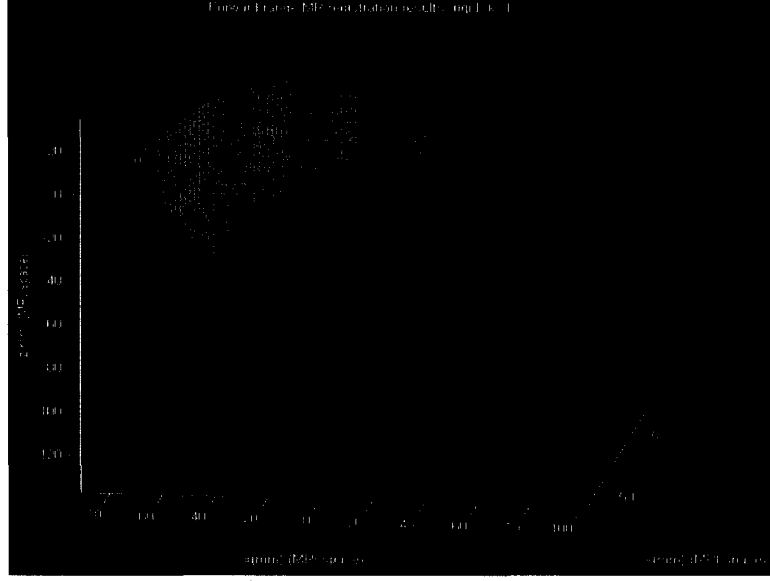


Figure 4.11 : *Elastic brain phantom surface from MRI scan with roughly registered range image overlaid, decimated by a factor of 2 (both rotated for improved visibility).*

ure 3.1 illustrates that our sensor can image brain tissue without any problems or pre-treatment.

In order to assess non-rigid surface tracking, the set screws under the phantom support plate are tightened, triggering a deformation of up to 15mm at the top, and the phantom surface is then imaged by the range-sensor. A rough manual alignment is provided, based on finding the points on the MR surface corresponding to the corners of the range domain and adding a vertical offset, as illustrated on figure 4.11.

The gradual improvement of the registration of the range and MR surfaces of the elastic phantom, over successive iterations and progressively more local characterization, can be illustrated as follows.

- *Distance statistics:* average and standard deviation values computed from the distance between homologous transformed range and MR point pairs. These statistics appear in figure 4.12 for the registration algorithm running with the usual parameter values, in particular with the smoothness parameter $\epsilon_{N,m} = \bar{\epsilon}_N + 3 \times \sigma_{\epsilon_N}$. In fact, this distance statistic overestimates the surface registration error, because the ray between any two mutually closest points will have a

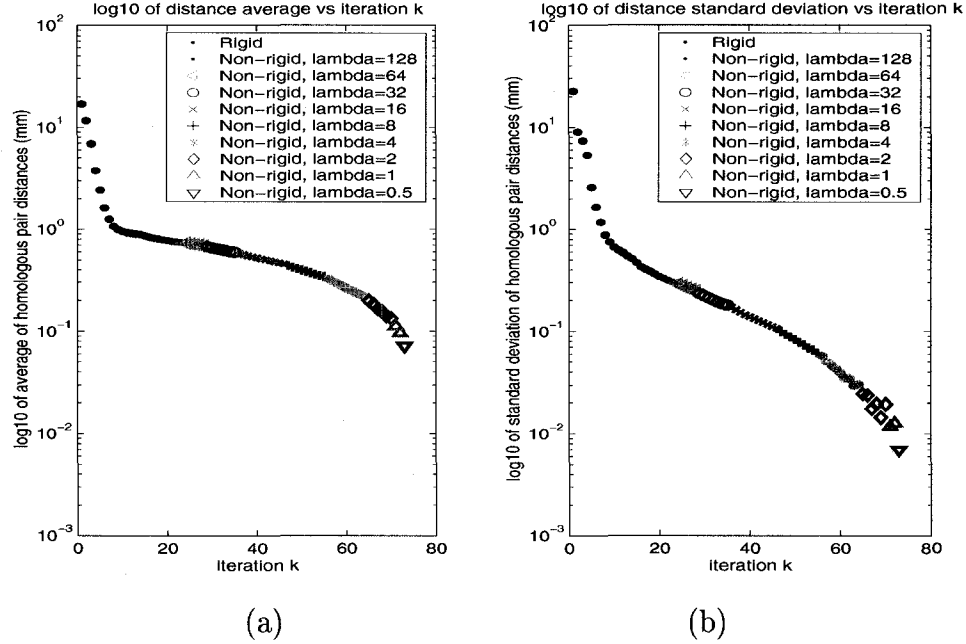


Figure 4.12 : *Statistics of distances between homologous points, assuming an overall transformation $\mathbf{T}_R^M$ (based on expressions 4.8 evaluated up to iteration $K = k$ or $L = k$) applied to range points: (a) $\log_{10}$ of distance average; (b) $\log_{10}$ of distance standard deviation.*

component along the shortest ray from one of the two points to the other surface, and in general a non-null component normal to that shortest ray.

- *Spatial distribution of the distance between two surfaces:* this distance function can easily be visualized as a coloured point whose position is that of a surface point in range or MR space and whose colour is assigned with a mapping of the distance value onto a colour scale, as in figures 4.13 and 4.14. These figures respectively show the gradual improvement of the *forward* and *inverse transformations* over successive iterations k and progressively more local characterization.

- Figure 4.13 illustrates the forward range-MR transformation from start to finish, bringing the range data progressively closer to the MR elastic phantom surface, based on the Closest Point Map computed from the latter surface. The range data, due to their greater density, are often initially

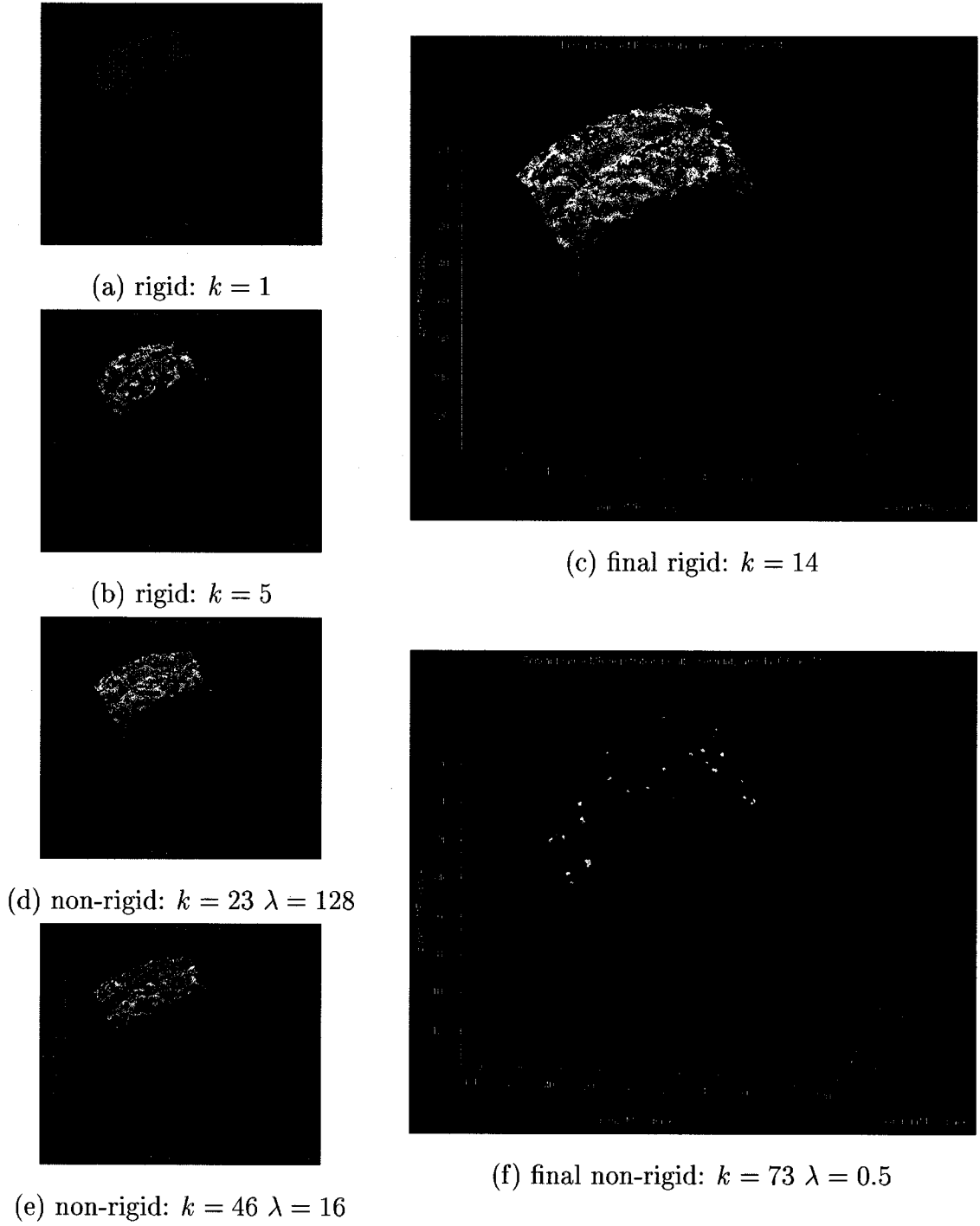


Figure 4.13 : *Evolution of forward range-MR transformation over iterations k and progressively more local characterization. The colour scheme of the range data indicates a homologous MR point within a certain distance d of the range surface: red: $d < 0.5\text{mm}$, yellow: $0.5 \leq d < 2.0\text{mm}$, green: $d \geq 2.0\text{mm}$.*

matched on a many-to-one basis with MR surface data (as in figure 4.8). In order to avoid over-estimating the distance to the MR surface, the distance ascribed to them is determined by the nearest *mutually closest homologous pair*. In this figure,

- * red coincides with a distance $d < 0.5mm$,
- * yellow with $0.5 \leq d < 2.0mm$, and
- * green with $d \geq 2.0mm$.

- Figure 4.14 illustrates the gradual improvement of the inverse MR-range transformation, based on the distance between mutually closest pairs mapped to a logarithmic “hot” colour scale, and rendered at the original point in MR space. Here,

- * black corresponds to $\log_{10}(0.1mm) = -1.0$,
- * saturated red to $\log_{10}(0.67mm) = -0.17$,
- * saturated yellow to $\log_{10}(4.78mm) = 0.67$, and
- * white to $\log_{10}(20mm) \simeq 1.3$.

One observation pertains to the gradual increase of the density of homologous pairs, as the transformation improves (note: points on the MR surface without a homologous range point are not displayed, so that only the blue background is visible). An analogy can be drawn here with a hypothetical scenario involving the registration of surface S_A taken as the top of a circus tent featuring three prominent peaks, and a second surface S_B (a rigid tarpaulin, for the sake of argument) of the same shape, initially suspended well above it from a helicopter. The three peaks of the tent will dominate the Closest Point Map defined high above it: the closest point from S_A , to any point from S_B high above, would be one of the three peaks. As the registration between the two surfaces improves, the contribution to the Closest Point Map of other points, below the peaks on S_A , would gradually have more bearing, until at the end, all points from the overlapping portions of the surfaces would contribute. This phenomenon, featuring

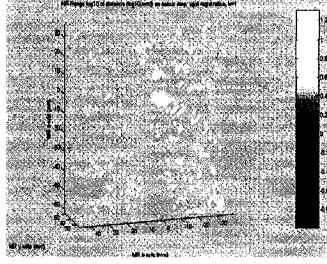
a few peaks dominating the Closest Point Map “above” the top phantom surface and determining the first few iterations of the registration, explains the results in figure 4.14.

- *2D slice illustration:* a qualitative illustration of the surface registration as it evolves over k and progressively more local characterization is provided by the consideration of any given slice in 2D, such the final results of the inverse transformation at $y_R = 79.2$.

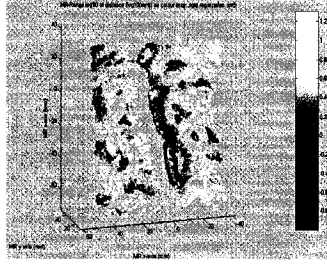
Finally, the impact of the value of the smoothing parameter can be seen by considering the distance statistics and the spatial distributions of distances between homologous pairs, once again, and comparing with results at the usual value. In particular, consider $\epsilon_{N,m} = \bar{\epsilon}_N + \sigma_{\epsilon_N}$ to $\epsilon_{N,m} = \infty$, as shown in figure 4.16.

4.3.2 Validation with an Analytical Displacement Applied to an Anthropomorphic Digital Brain Surface

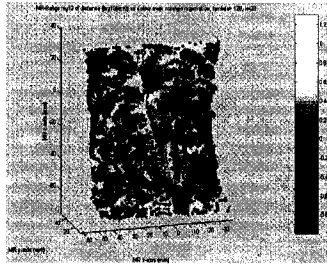
This validation study of the non-rigid surface registration procedure compares the vector displacement function applied to synthetic range image of the brain, obtained from an anthropomorphic digital phantom, with the displacement computed by the registration procedure. As shown in illustration 4.18, the clinical situation is simulated by taking a patch of the known brain surface of this phantom, of isotropic 1mm sampling (see section 7.1 and refer to [74] for further details), and interpolating it in order to replicate the density of the range-sensor CCD. To each synthetic range image obtained in this manner, I apply a “downward” analytical displacement function consisting of a constant plus a slowly varying sinusoid. This study features two synthetic range images, as shown in figure 4.18: a top patch, isolated from the most distal brain surface points of the phantom in a well-centered square area, whose displacement is essentially proximal, as well as a left patch, from the left-most brain surface points in a similarly well-centered square area, whose displacement is towards the right.



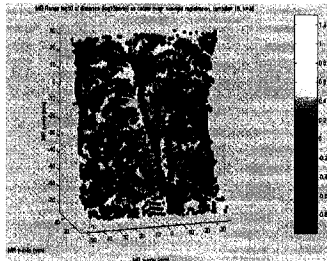
(a) rigid: $k = 1$



(b) rigid: $k = 5$



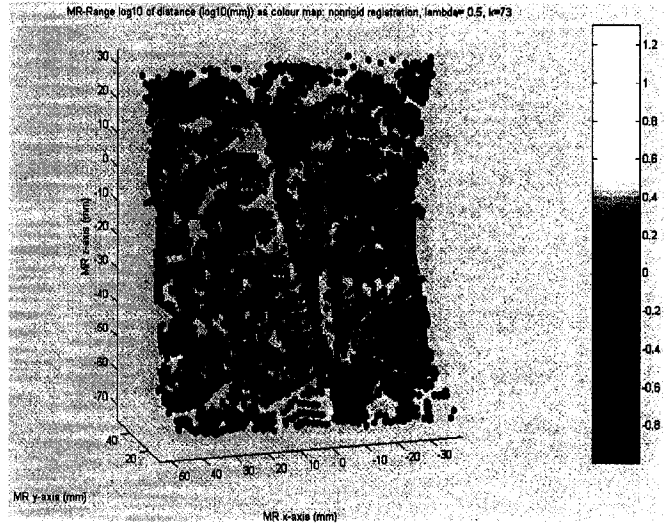
(d) non-rigid: $k = 23 \lambda = 128$



(e) non-rigid: $k = 46 \lambda = 16$

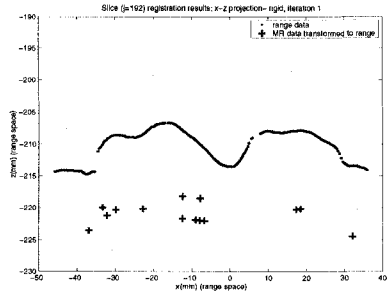


(c) final rigid: $k = 14$

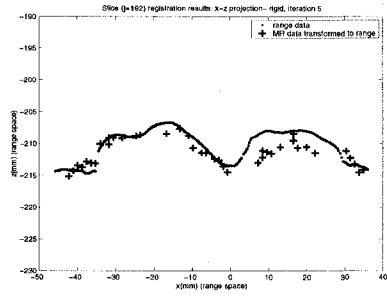


(f) final non-rigid: $k = 73 \lambda = 0.5$

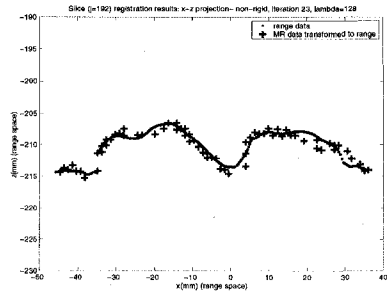
Figure 4.14 : *Evolution of the inverse MR-range transformation over iterations k and progressively more local characterization. The distance between homologous points is shown as a coloured point in MR space. The colour is determined by the logarithmic “hot” colour scale on the right. Black corresponds to -1.0 (0.1mm), saturated red to -0.17 (0.67mm), saturated yellow to 0.67 (4.78mm), and white to 1.3 (20mm).*



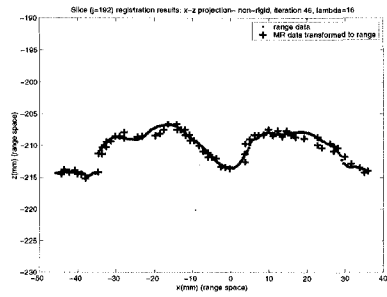
(a) rigid: $k = 1$



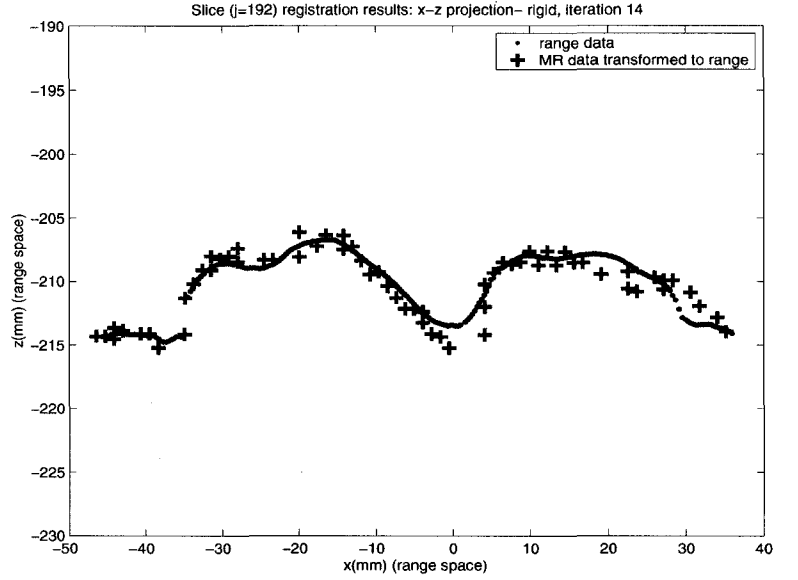
(b) rigid: $k = 5$



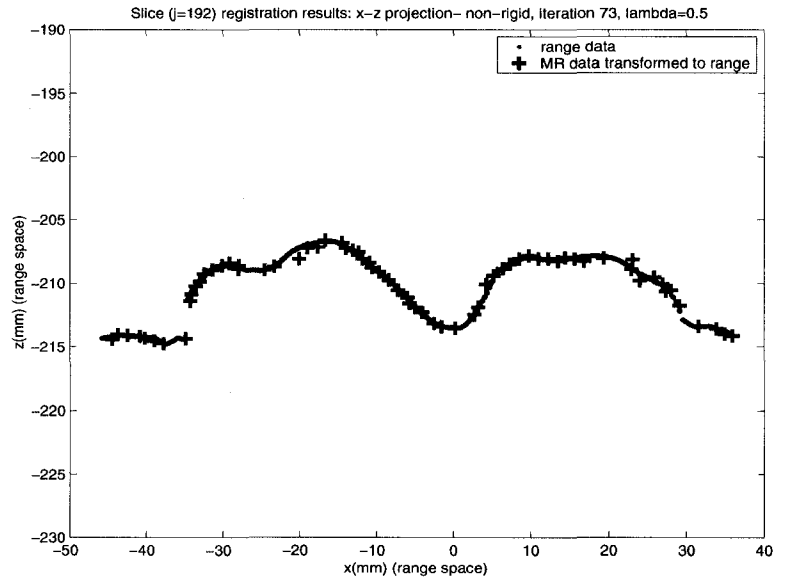
(d) non-rigid: $k = 30 \lambda = 128$



(e) non-rigid: $k = 46 \lambda = 16$



(c) final rigid: $k = 23$



(f) final non-rigid: $k = 73 \lambda = 0.5$

Figure 4.15 : Evolution of inverse MR-range transformation versus iteration k and progressively more local characterization, for the slice $y_R = 79.2$ ($j = 192$), featuring homologous range points and MR brain points mapped back to range space according to $\mathbf{T}_{M,k}^R$ and projected to the $x_R - z_R$ plane.

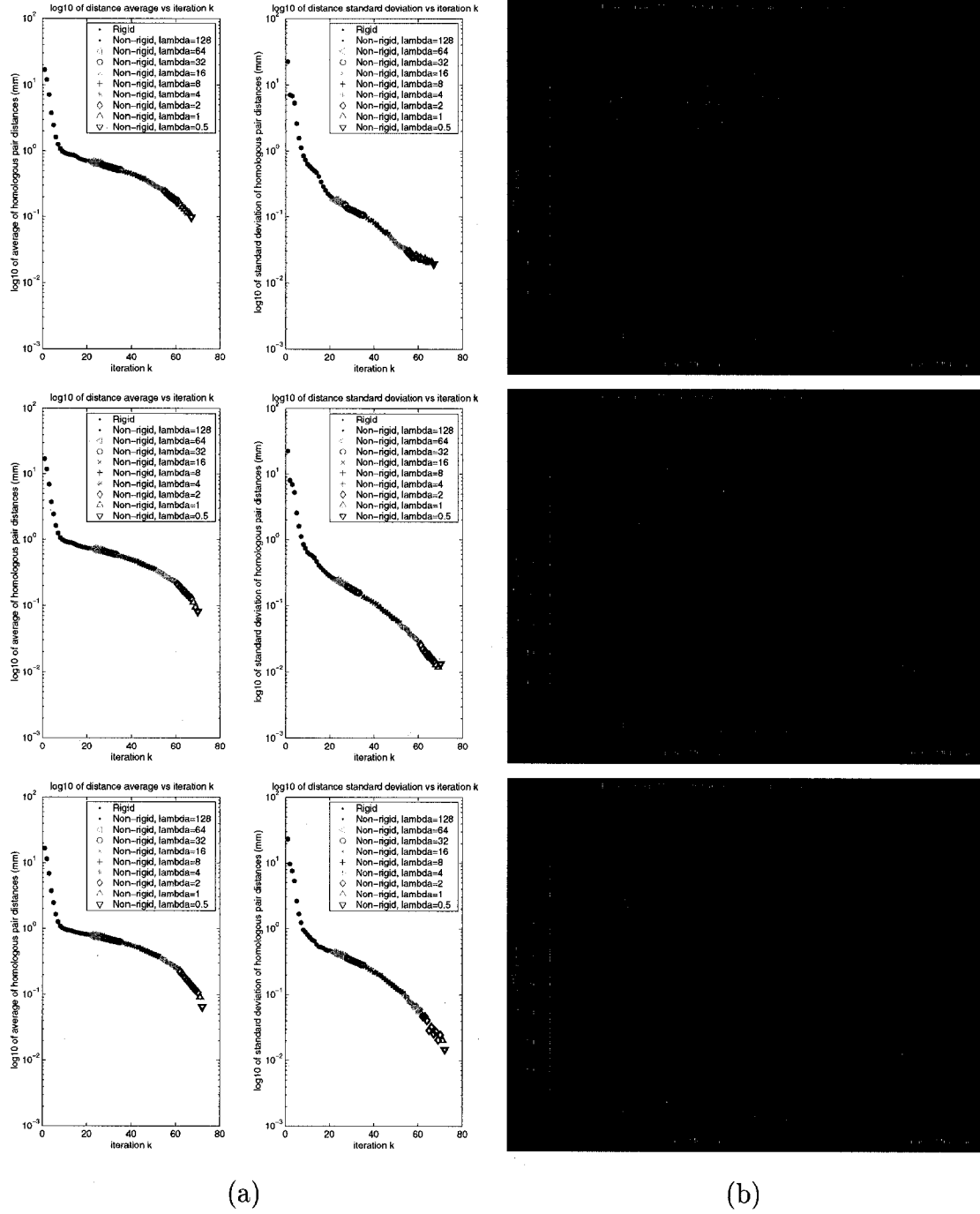


Figure 4.16 : *Influence of smoothness parameter value for forward transformation: (a) distance statistics and (b) final forward transformation results, for $\epsilon_{N,m} = \bar{\epsilon}_N + \sigma_{\epsilon_N}$ (top); $\epsilon_{N,m} = \bar{\epsilon}_N + 2 \times \sigma_{\epsilon_N}$ (middle row); $\epsilon_{N,m} = \infty$ (bottom). Compare with results for $\epsilon_{N,m} = \bar{\epsilon}_N + 3 \times \sigma_{\epsilon_N}$ appearing in figures 4.12 and 4.13 (f).*

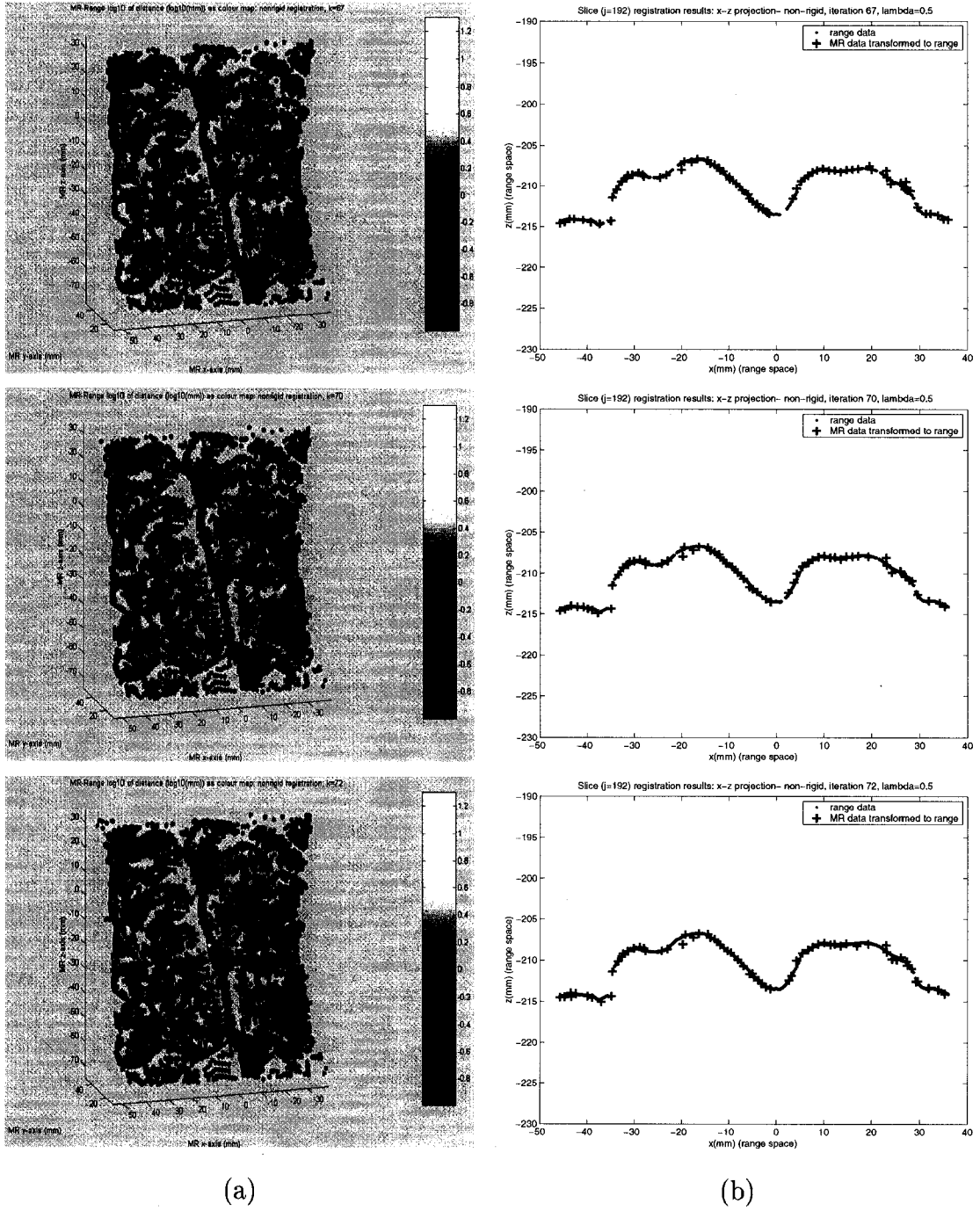


Figure 4.17 : *Influence of smoothness parameter value for inverse transformation: (a) final inverse transformation colour map and (b) 2D slice results, for $\epsilon_{N,m} = \bar{\epsilon}_N + \sigma_{\epsilon_N}$ (top); $\epsilon_{N,m} = \bar{\epsilon}_N + 2 \times \sigma_{\epsilon_N}$ (middle row); $\epsilon_{N,m} = \infty$ (bottom). Compare with results for $\epsilon_{N,m} = \bar{\epsilon}_N + 3 \times \sigma_{\epsilon_N}$ appearing in figures 4.13 (f), 4.14 (f) and 4.15 (f).*

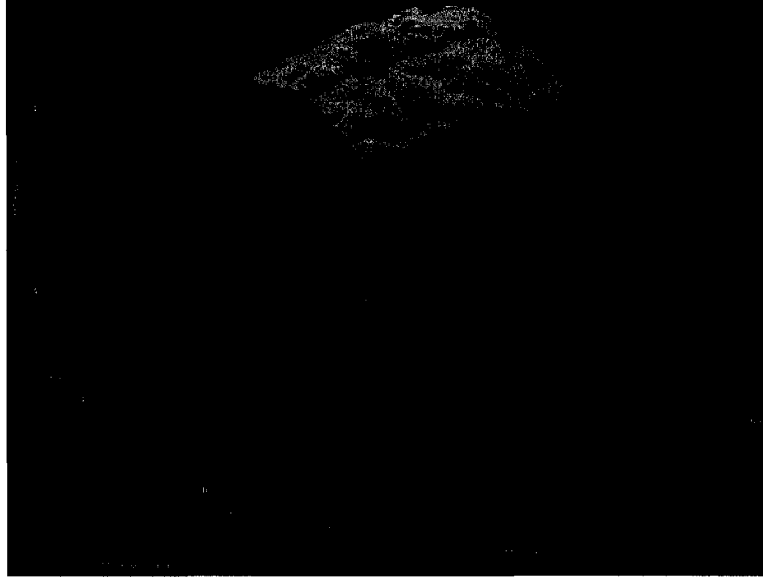


Figure 4.18 : *Two synthetic range images overlaid on original phantom brain surface data, from which they were obtained by interpolation.*

A measure of the accuracy of the registration procedure is then computed from the vector difference between the *pre-deformation* synthetic range patch, and the *post-deformation, post-registration* range patch. Figure 4.19 and table 4.1 illustrate the gradual improvement in displacement estimation, over successive iterations. The deformation function is of the type

$$\mathbf{d} \equiv d_3 = [A + B\cos(2\pi f(x_1 - \bar{x}_1))\cos(2\pi f(x_2 - \bar{x}_2))]\hat{x}_3, \quad (4.11)$$

where x_1 and x_2 coincide with x and y for the top patch study, and with y and z for the left patch study. The period of the sinusoid is chosen to be approximately the width of the patch, and a few combinations of A and B are documented in table 4.1. At the finer scales, small improvements in surface-to-surface distances may no longer be reflected in an improvement the displacement estimation, and there is a risk that spline-based motion characterization could warp the surface too aggressively, essentially unfolding the sulci. To alleviate this risk, the progression of the motion characterization from coarse to fine is gradual, and each non-rigid iteration is in fact a composition of a rigid-body iteration and a spline-based non-rigid motion characterization. As a consequence, the iterative minimization of surface-to-surface distance correlates well

deformation parameters	estimation error (mm)			inter-surface distance (mm)		
Left patch	start	rigid	nonrigid	start	rigid	nonrigid
A= 5 B= 5 f= 1/80	4.2 ± 2.2	1.4 ± 0.6	0.7 ± 0.2	3.6 ± 5.4	0.7 ± 0.3	0.09 ± 0.01
A= 10 B= 5 f= 1/80	9.3 ± 2.1	1.6 ± 0.5	0.8 ± 0.2	4.0 ± 12.2	0.8 ± 0.3	0.09 ± 0.01
A= 15 B= 5 f= 1/80	14.1 ± 2.0	1.8 ± 0.6	0.9 ± 0.3	3.7 ± 13.9	0.8 ± 0.3	0.08 ± 0.01
A= 5 B= 5 f= 1/60	3.3 ± 3.2	1.7 ± 0.9	0.9 ± 0.3	3.0 ± 3.8	0.9 ± 0.4	0.09 ± 0.01
Top patch	start	rigid	nonrigid	start	rigid	nonrigid
A= 5 B= 5 f= 1/80	5.2 ± 2.1	1.2 ± 0.6	0.6 ± 0.2	2.4 ± 2.5	0.7 ± 0.3	0.10 ± 0.01
A= 10 B= 5 f= 1/80	10.6 ± 2.3	9.1 ± 2.3	1.0 ± 0.5	2.8 ± 4.5	0.8 ± 0.3	0.09 ± 0.01
A= 15 B= 5 f= 1/80	15.8 ± 2.5	1.6 ± 0.4	0.6 ± 0.2	2.9 ± 5.5	0.8 ± 0.3	0.09 ± 0.01
A= 5 B= 5 f= 1/60	4.1 ± 3.2	1.5 ± 1.0	0.7 ± 0.3	2.2 ± 1.8	0.8 ± 0.4	0.10 ± 0.01

Table 4.1 : *Analytical deformation statistics. Left column: deformation parameters (in mm, mm and mm^{-1}), as per expression (4.11). Middle column: magnitude of vector difference between estimated and analytical displacement (at the start, after the final rigid iteration and after the final non-rigid iteration). Right column: distance between mutually closest points of the synthetic range and MR surfaces.*

with the minimization of the error in non-rigid displacement estimation.

Estimation error statistics of the analytical deformation study are presented in table 4.1, along with the distance between two surfaces. In all cases, the final non-rigid registration provides, on average, displacement estimation on the order of 1 mm or less. Nevertheless, this error is significantly larger than the final surface-to-surface distance. Given that the average distance between the two surfaces is relatively small, much of the displacement error is likely to be orthogonal to the direction of the true displacement. In no way is any assumption of a strong gravitational component exploited in the deformation estimation, because even if the movement were exclusively downward, one could not assume that the sensor $x - y$ axis were normal to the “downward” direction. Therefore, attempting to model the movement exclusively along z would not characterize it adequately. However, the assumption of a dominant downward component is manifest in the validation displacement function which is along z for the top patch study, and along x for the left patch study. The

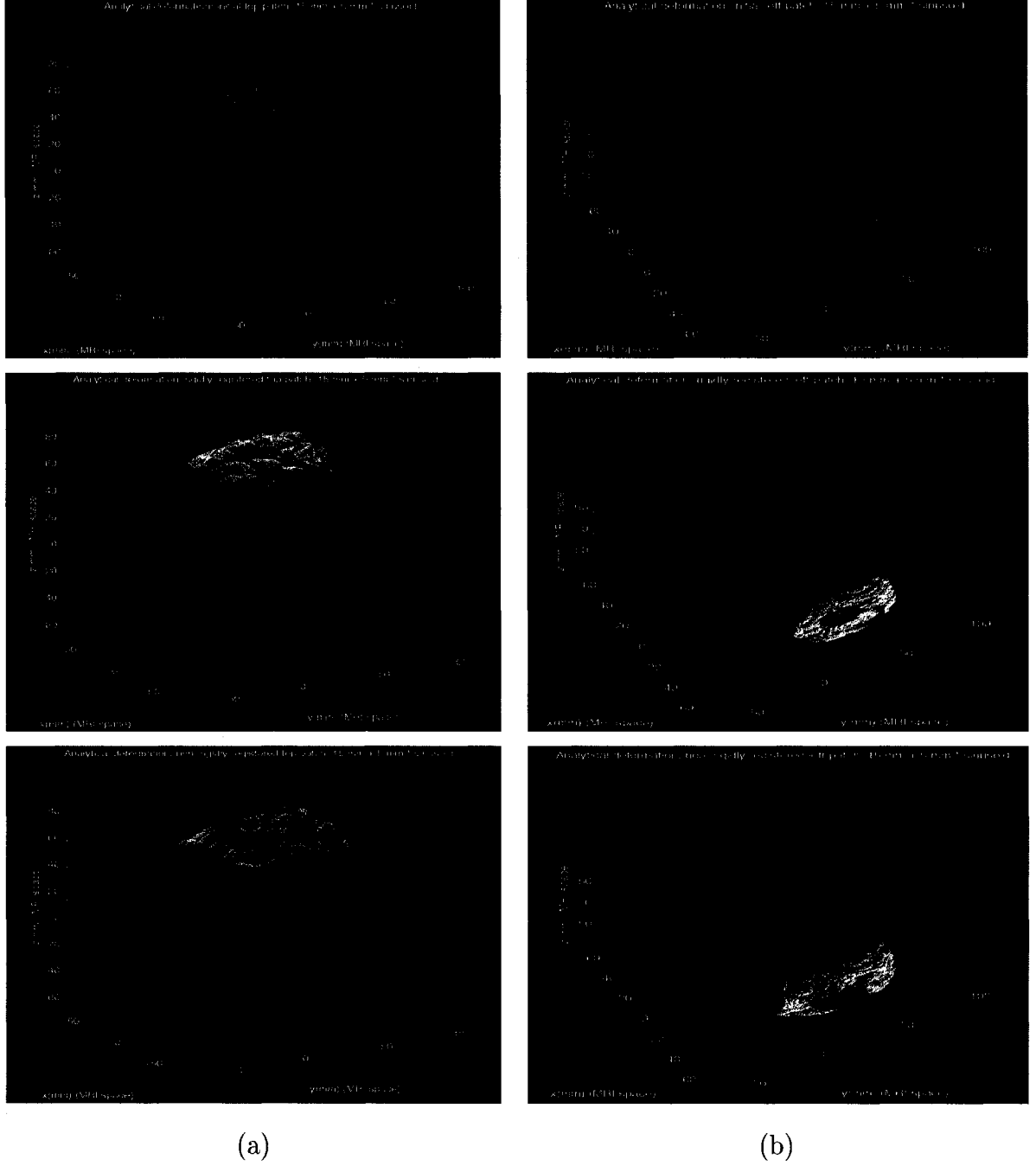


Figure 4.19 : *Analytical deformation results: (a) top study and (b) left study.*

Top row: initial deformation $A + B\cos(2\pi f(x_1 - x_{1,c}))\cos(2\pi f(x_2 - x_{2,c}))$, where $A = 15\text{mm}$ and $B = 5\text{mm}$; middle row: final rigid registration; bottom row: final non-rigid registration. Red signifies that the deformation is estimated with error $\leq 1.0\text{ mm}$, yellow: $1.0\text{ mm} < \text{error} \leq 2.0\text{ mm}$, and green: error $> 2.0\text{ mm}$.

anomalous result of the second top patch study stems from an early exit of the rigid registration stage, which could be prevented by more tolerant logic, but even then the non-rigid stage, consisting of a sequence of compositions of rigid transformation and spline-based displacement, provided reasonable results.

In view of the final non-rigid results in the bottom row of figure 4.19, the spatial distribution of the yellow points, whose estimation error lies between 1.0 and 2.0 mm, suggests a possible limitation of using an Iterative Closest Point scheme to model brain deformation. The estimation is better at ridges and peaks, coinciding with gyri, and conversely is worse in crevasses coinciding with sulci, as well as in sloped areas. One possible interpretation is that the closest point pair constitutes a good homologous set when the surface normals at both points are roughly aligned with the true direction of the deformation, as is the case with anatomically equivalent points from a gyrus. However, a sulcus is characterized by a crevasse surrounded by points on a slope: for a given crevasse point on surface A, the closest point is likely to be on a slope, rather than the corresponding crevasse point on surface B. This geometric limitation is also why the ICP method cannot resolve motion having a strong component tangential to the either surface normal. For example, while this method can resolve a radial increase in scale of a spherical surface, it cannot deal with a pure rotation.

4.4 Conclusions and Future Directions

This chapter presented a technique for estimating non-rigid surface motion, based on the Iterative Closest Point framework. Innovations relating to computational efficiencies were introduced, particularly the use of a precomputed Closest Point Map, and recursive spline filters to model the deformation. Further innovations related to the validation study, through the use of a constitutively realistic brain-shaped deformable phantom, whose deformation is triggered from within, and an analytical deformation function applied to a synthetic cortical range surface, obtained by interpolating the surface coinciding with the outer grey matter boundary of the digital head phan-

tom. The registration results are consistent with other, less efficient ICP techniques, characterized by a gradual reduction of the distance between two surfaces.

The study based on a known deformation suggests that overall the ICP registration technique accurately estimates non-rigid motion, but performs slightly less well in highly sloped areas, such as those in the deep recesses of sulci. This consideration leads to a refinement of the ICP method similar to that proposed by Feldmar and Ayache [107, 108, 109] and reviewed in chapter 2, where only those points which are both close in space and of similar shape are considered homologous. However, it is less obvious how to adapt this shape-based approach to a Closest Point Map. One possibility might involve a second Closest PointMap computed in a modified range space, spanned by (x_R, y_R, K_R) , where K_R is the Gaussian curvature of each point in the range image. The registration could alternate between one forward iteration that takes place in MR space, relating (x_M, y_M, z_M) to $\mathbf{T}(x_R, y_R, z_R)$, and one inverse iteration that seeks homologous pairs between points in modified range space (x_R, y_R, K_R) and $\mathbf{T}^{-1}(x_M, y_M, K_M)$, where K_M is the Gaussian curvature of each MR brain surface point. Alternately, mean curvature H could be substituted for K , in which case H is computed in the context of the brain surface segmentation in MR, as seen in chapter 6.

Chapter 5

Model-based Brain Surface Identification in MRI: Motivation, Model Overview and Surveys

5.1 Introduction: Motivation for Model-based Brain Surface Identification

This chapter begins by providing some motivation for using a semi-automatic model-based approach for anatomical surface identification (also known as *segmentation*). One can further distinguish between two kinds of surface models, namely *physically-based* and *surface evolution* models, and I provide intuitive justification for employing the latter kind. Currently, a prevalent technique for identifying anatomical surfaces in IGS involves identifying a point in a slice corresponding to the tissue type of interest, and based on this “seed”, labelling software identifies the rest of the contiguous voxels which are of similar intensity (based on user-defined intensity thresholds). Occasionally, a low contrast area at the tissue boundary will cause the labelling to bleed onto tissues of the wrong type, which requires special manual intervention in this area to alleviate the bleeding effect, and typically entails a labour-intensive slice-by-slice, rather than 3D, approach [168, 229]. Current tomographic volumes involve up to

256 of these slices, and any slice may require this kind of interaction. Finally, the contiguous labelled voxels within these slices together comprise a volume of interest, coinciding with brain tissue for example, the outer surface of which (identified by the Marching Cubes algorithm [224]) constitutes the boundary of interest.

In contrast, this thesis research proposes a semi-automatic anatomical surface identification stage to replace this procedure. There are four main reasons for choosing this approach.

- An immediate consequence of an automatic or a semi-automatic technique is the *reproducibility* of the results, as opposed to the inter- and intra-subject variability documented for manual identification, even by expert anatomists [430].
- Another immediate consequence is the *reduced user intervention* (for a well-designed algorithm and interface).
- A semi-automatic or automatic technique can incorporate *prior knowledge* into the model in order to avoid common pitfalls, such as low-contrast areas which tend to recur across scan subjects. However, because of the variability of brain anatomy across subjects, particularly where pathology is involved, and of the comparative cost of failure of an algorithm when measured against a small amount of user involvement, I consider a semi-automatic segmentation technique more attractive than a fully automatic one.
- Finally, a semi-automatic technique can easily process tomographic information in a fully *three-dimensional* manner, which offers advantages over a series of 2D slice operations. The justification for an inherently 3D algorithm is that either type of model, physical or surface evolution, features a term that favours a *smooth* contour (in 2D) or surface (in 3D), and a 3D framework produces a boundary that is smooth not only along an intra-slice contour, but across neighbouring slices. This characteristic is critical when using a shape-based registration of a segmented surface with a range image of the same structure, whose sampling coordinate system has no relation with that of the MRI volume.

The rest of this chapter serves two purposes: it reviews the literature relevant to model-based surface identification, and provides an intuitive introduction to the technique adopted and refined in the course of this thesis research.

Magnetic resonance uses an external magnetic field to induce a corresponding magnetization of charged nuclear particles, with hydrogen being the most prevalent, causing them to precess at a characteristic frequency [361, 285]. They can then be temporarily excited by a radiofrequency pulse of the same frequency, altering their magnetization, which in turn can be measured. The signal, which is generated by the return of the various components of the magnetization to their previous equilibrium, is characterized by its intensity and decay rate. The signal is a complex-valued function whose magnitude is retained, and where each component is corrupted by Gaussian noise, with the consequence that the magnitude of the signal exhibits Rayleigh-distributed noise in the background and Rician-distributed noise in the volume occupied by the subject [145, 146, 135]. Furthermore, the signal can be manipulated to enhance contrasts in various aspects of tissue magnetization [361, 285]; i.e.: it can be weighted to reflect *spin-lattice* ($T1$) or *spin-spin* ($T2$) *relaxation* times, or the density of spins themselves (*proton density*, or PD). We then speak of $T1$ -weighted, $T2$ -weighted or PD-weighted images ¹.

Numerous authors have proposed methods for estimating noise and improving SNR in MR volumes [276, 352, 288]. The traditional approach involves subtracting two acquisitions of the same object and computing the standard deviation of the resulting image [276]. More recently, Sijbers [352] suggested a technique based on the cross-correlation of two acquisitions of the same MR image, in a manner which exploits Parseval's theorem to estimate noise and improve SNR directly from k-space data. In his Ph.D. thesis [353], Sijbers suggested and compared 4 methods to improve SNR in MRI, including an anisotropic filter which improves image SNR significantly, despite being optimized for Gaussian rather than Rician noise. In this thesis, anisotropic filtering is also adopted to preprocess the MR volume, and a discussion of filtering is

¹For the sake of brevity, I will occasionally refer to such image volumes as simply $T1$, $T2$ or PD data.

provided in this chapter.

Section 5.2 provides an intuitive explanation of surface evolution models, and a motivation for using them to identify the brain surface in MR data. Surface evolution models [57, 419] express surface identification as a partial differential equation in simulated time t and in space (x, y, z) , from a starting surface (or collection of surfaces) provided by the user. As will be seen, this kind of model possesses certain advantages, such as an ability to merge or break apart as dictated by the underlying image information, a characteristic known as *topological adaptability*. The details of the surface model proposed here, emphasizing original contributions at the implementation and validation levels, are deferred until the next two chapters.

Section 5.3 features a literature survey on model-based surface identification techniques as well as a survey of preprocessing techniques used prior to model computations. In § 5.3.1, the two kinds of surface models are broadly defined and reviewed. Also, model-based surface identification within a MRI volume² requires some kind of force or stopping term which binds the model surface to the relevant anatomical boundary on the basis of image information, such as transitions which can be detected as strong gradients in voxel intensity [250, 419]. The robust computation of this image term in turn imposes the need for a preprocessing stage, and to this effect I review preprocessing techniques used in comparable applications in § 5.3.2. For the reader who is merely interested in a brief synopsis of each survey...

- Surface models are categorized into those based on an explicit representation and on a physical equilibrium between virtual internal properties of the model and image-based external forces [179, 306, 389], and those based on the evolution of an implicit surface expressed as a partial differential equation in time which imbeds the problem into one defined over the complete image domain [237, 419]. Also noted are explicit surface models featuring important contributions to surface representation [83, 381, 254, 373], topological adaptability [252, 85] and

²Conceivably, these approaches could be applied to a CT volume, however soft-tissue delineation on the basis of X-ray attenuation is not as discriminating as that achievable with MRI based on magnetization characteristics. For this reason, MRI data is emphasized rather than CT data.

statistical formulation [78].

- Finally, the computation of an image term which binds a surface model to an anatomical boundary typically requires a preprocessing stage in order to be tolerant to noise. Section 5.3.2 traces the evolution of filtering from isotropic operators, particularly the Gaussian kernel [241], through its implementation via the Heat Equation as proposed by Koenderink [195], to anisotropic diffusion techniques which are expressed as partial differential equations in time and space which are generalizations of the Heat Equation [308, 189].

5.2 An Intuitive Overview of Surface Evolution Models

The anatomical surface segmentation technique proposed in this thesis is an adaptation of a recently published surface evolution model [57, 419], which expresses surface identification as a partial differential equation in time t and space (x, y, z) , evolving from a user-defined initial surface (or collection of surfaces), corresponding to time $t = 0$, to one which coincides with the anatomical boundary of interest. Here, the time parameter t refers to simulated time, whose unit is arbitrary but commonly expressed in seconds. Moreover, x , y and z describe voxel coordinates, although the model also provides for anisotropic spacing. The initial surface may be fully contained in the anatomical boundary of interest, in which case it expands towards the boundary over time. Otherwise the surface fully contains this boundary and contracts toward it. The surface evolution model imbeds the identification of an implicit 3D surface $\mathcal{S}$ into the estimation of an evolving volumetric function $\Psi(x, y, z, t)$, whose isosurface $\{\Psi = 0\}$ within the tomographic volume coincides with $\mathcal{S}(t)$ throughout. This concept of imbedding is illustrated in figure 5.1 for the analogous 2D problem, identifying a contour $\mathcal{C}$ coinciding with an anatomical boundary in a 2D image [295].

An initial contour $\mathcal{C}(t_0)$ is first defined by the user outside the boundary of in-

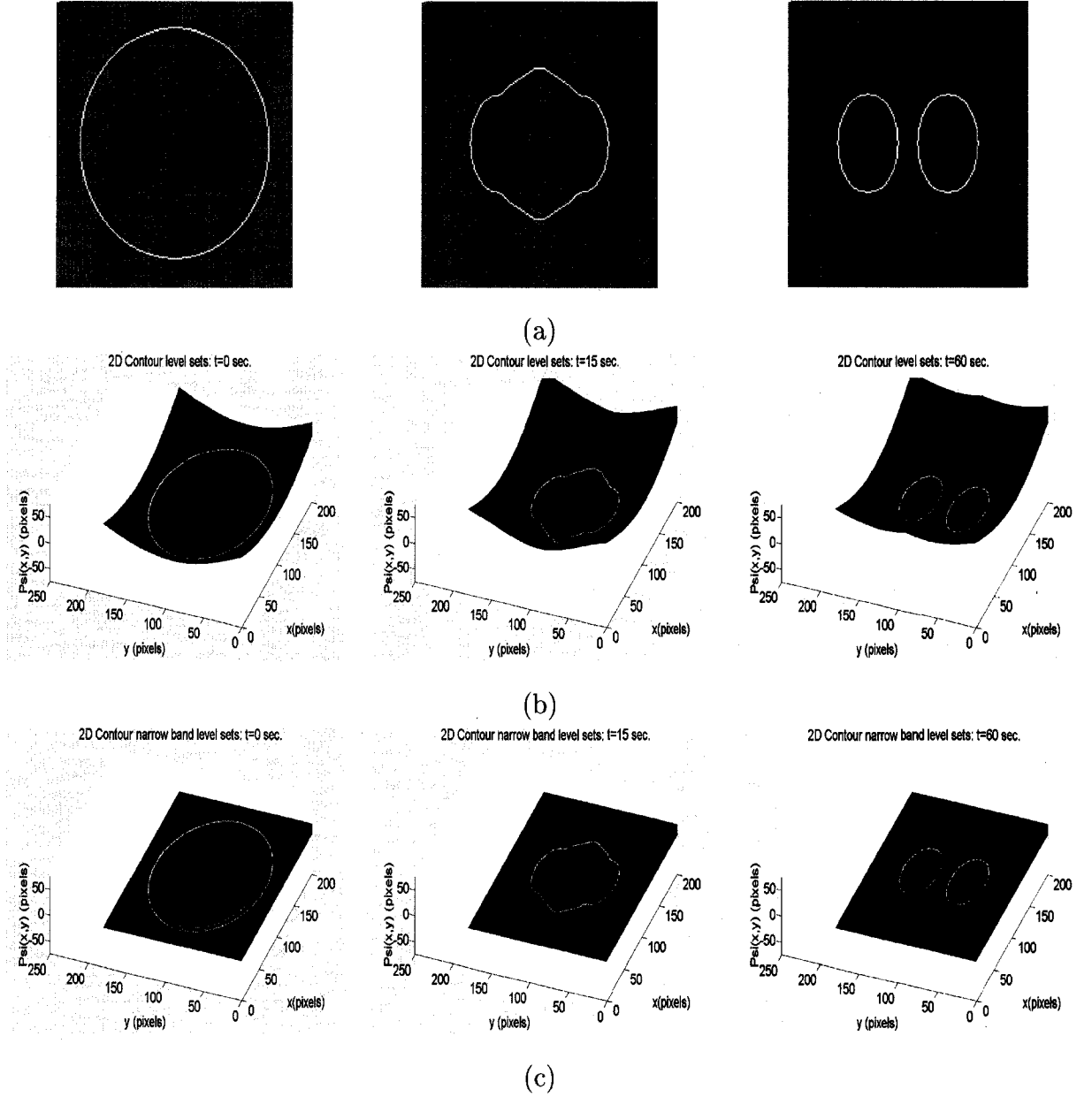


Figure 5.1 : *Imbedding function with 2D contour example, with inward-moving zero-level isocontour, based on synthetic image featuring 2 elliptical bodies: (a) zero-level isocontour (white), overlaid on image at $t = 0$, at $t = 15$ sec., and at $t = 60$ sec.; (b) imbedding function Ψ , with level set $\{\Psi = 0\}$ coinciding with isocontour in (a), at $t = 0$, at $t = 15$ sec., and at $t = 60$ sec.; (c) narrow band implementation of contour evolution model, with identical situation as in (b): the computation of Ψ is restricted within a narrow band near the evolving zero-level isocontour.*

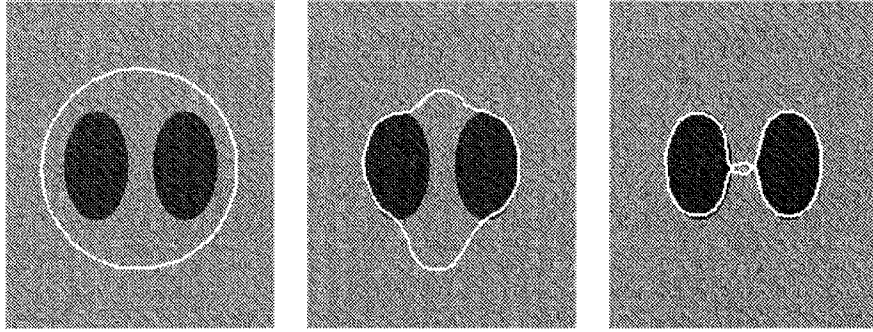


Figure 5.2 : *Identification of elliptical boundaries as in figure 5.1, with a physically-based model without topological adaptation, courtesy of Johan Montagnat [86]: initial configuration, 100th iteration, final iteration.*

terest, and the imbedding function $\Psi(x, y, t)$ is initialized as the *signed distance map* from $\mathcal{C}(t_0)$, as shown in the left column of figure 5.1 (b). The initial contour coincides with the isocontour $\{\Psi(x, y, t_0) = 0\}$. Interior points in the image domain are typically labelled negative, and the exterior points are labelled positive. The imbedding function $\Psi(x, y, t)$ evolves in such a way (figure 5.1 (b), middle column) that its $\{\Psi = 0\}$ isocontour, or *zero-level set*, moves inwards until it coincides with the anatomical boundary of interest, at which point the zero-level set binds to the boundary (figure 5.1 (b), right column), and the resulting contour is precisely this final zero-level set $\{\Psi(x, y, t_f) = 0\}$ (figure 5.1 (a), right column). To reduce the computational load of imbedding a problem over the whole image domain, I adopt the *Narrow Band* approach proposed by Adalsteinsson [2], whereby the computation of $\Psi(\mathbf{x}, t)$ is restricted to a band, a few pixels wide, in the neighbourhood of its evolving zero-level isocontour. This concept is illustrated in figure 5.1 (c) for the 2D contour identification pursued in this discussion.

A contour/surface evolution model possesses certain advantages over physically-based contour/surface models [389, 306], such as its capacity to capture a large variety of topologies from a given initialization and a relative insensitivity to initial conditions [237]. In contrast, physically-based models cannot easily cope with differences in topology with respect to their underlying model. To be fair, significant progress has

been made recently in making physical models more topologically adaptable [252, 85], although at the cost of explicit collision detection, and the extension from 2D to 3D is not as simple as the evolution framework adopted in this thesis research. A comparison of the examples in figures 5.1 and 5.2 illustrates this distinction. Moreover, the tensile properties of physical models may make them ill-suited for capturing highly involuted surfaces (such as that of the inner or outer cortex), as illustrated for the 2D invaginated circle in figure 5.3. The difference between the two approaches stems not merely from a difference in stiffness properties, but is a consequence of the mathematical formulation of the surface evolution model, which features separate smoothing and discontinuity-preserving terms, as will be seen in the next section. From the outset, the justification for choosing a surface evolution model lay in its flexibility in capturing complex shapes, such as the cortex and arbitrarily diffuse pathologies, with minimum user intervention at the initialization stage.

Alternately, the 2D model could just as easily have been initialized with one or more contours completely inside the boundary of interest, and could have employed an outward moving isocontour as shown in figure 5.4. Finally, the method described here fully generalizes to identifying a 3D surface within a volume, requiring an inner or outer user-defined closed surface from which a signed distance map is computed with the volume as its domain, which serves as an initialization for the imbedding function $\Psi(x, y, z, t)$, as shown in figure 5.5.

The surface model proceeds from the signed distance map computed from the user-defined surface, and evolves according to the following expression [58, 419]:

$$\frac{\partial \Psi}{\partial t} = F(x, y, z) \|\nabla \Psi\| \left[\operatorname{div} \left(\frac{\nabla \Psi}{\|\nabla \Psi\|} \right) + \nu \right] + \nabla F \cdot \nabla \Psi, \quad (5.1)$$

This model features a *diffusive* term $\|\nabla \Psi\| \operatorname{div}(\nabla \Psi / \|\nabla \Psi\|)$ which tends to smooth out each isosurface, including the zero-level surface, a *hyperbolic* term $\|\nabla \Psi\| \nu$ which pushes the each isosurface forward in a discontinuity-preserving manner, and two image terms: a *speed* function F which slows down the zero-level surface near anatomical boundaries, and a so-called “*doublet*” term $\nabla F \cdot \nabla \Psi$ which prevents the zero-level surface from overshooting these boundaries. The sign of the ν constant in the hyperbolic

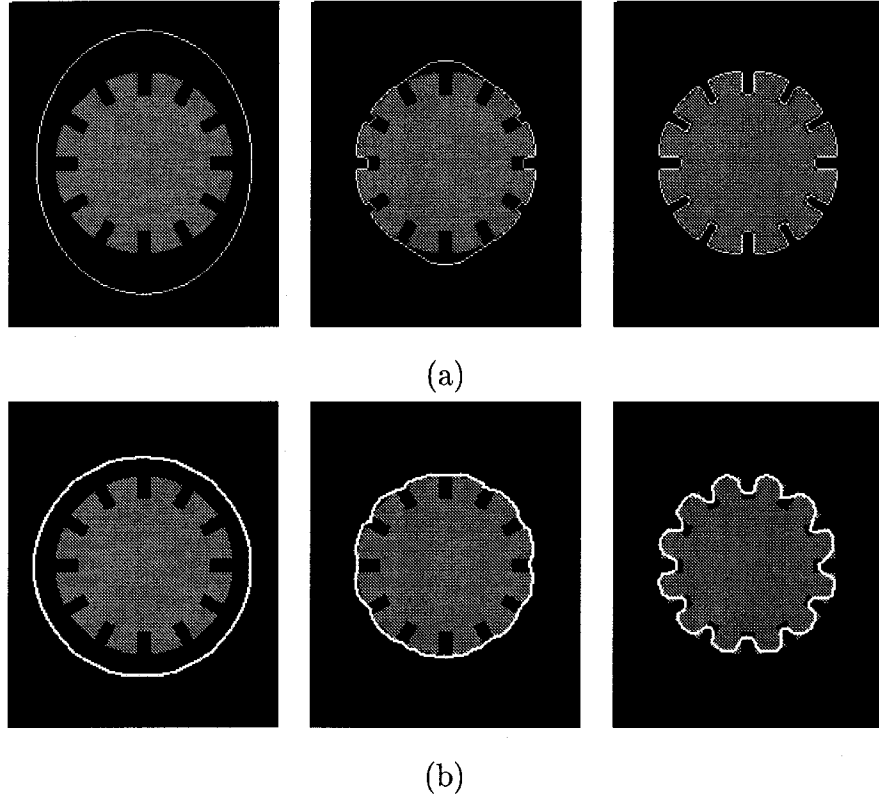


Figure 5.3 : *Demonstration of difference in tensile properties of evolution and physical models, on 2D invaginated circle: (a) contour evolution model at initialization ($t=0$), at 10 sec., and at 30 sec; (b) physical model at initialization, at 8th iteration, and at 30th iteration, courtesy of Johan Montagnat [86].*

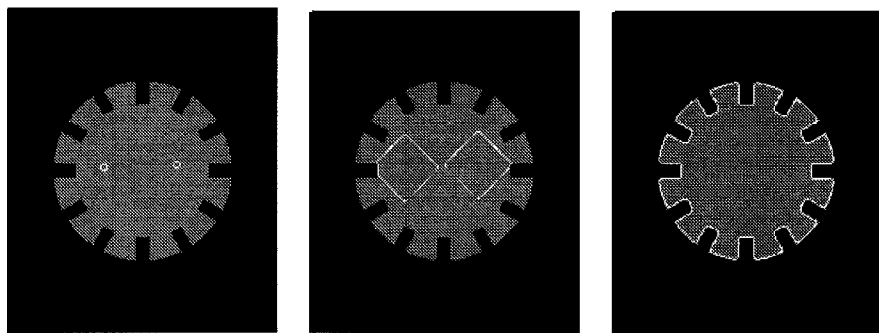


Figure 5.4 : *Contour identification with outward-moving isocontour: zero-level isocontour (white), overlaid on invaginated circle image as in figure 5.3; at $t = 0$, at $t = 10$ sec. and at $t = 60$ sec..*

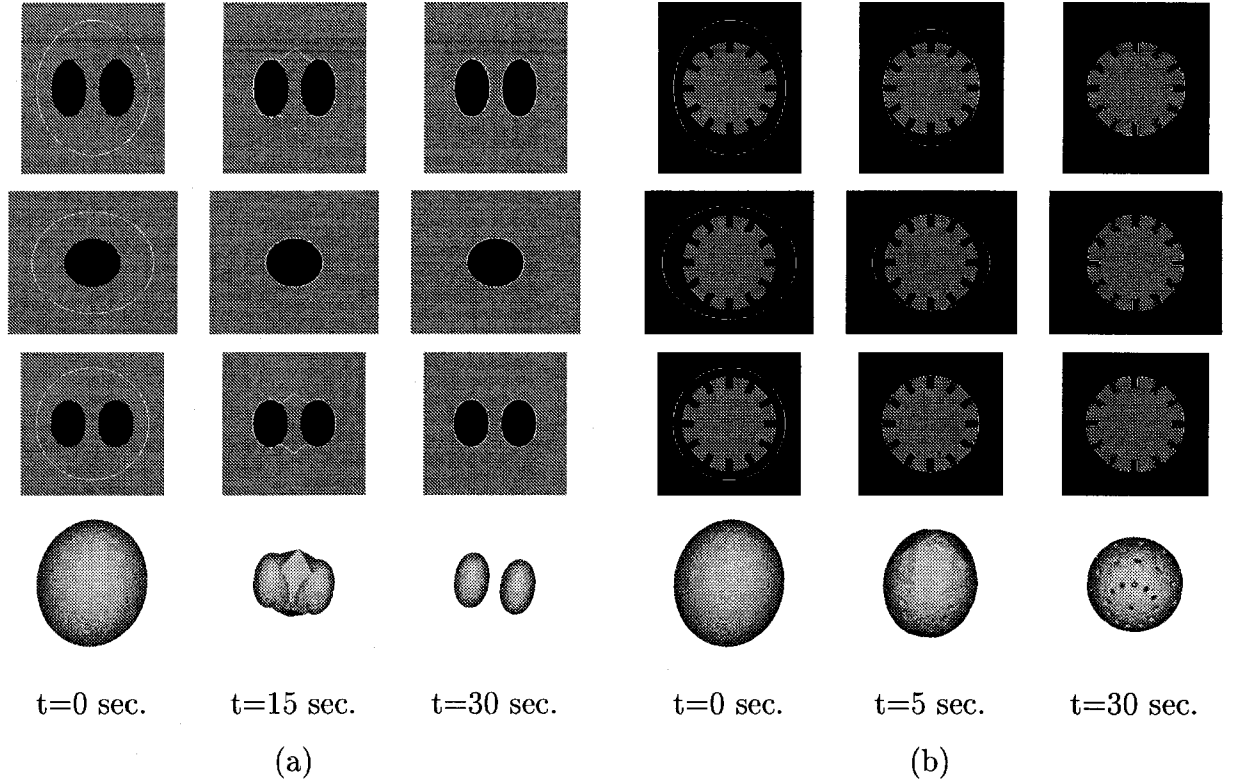


Figure 5.5 : *3D surface identification with synthetic volume featuring (a) a pair of ellipsoids and (b) an invaginated sphere: multiplanar and rendered views of $\Psi(x, y, z) = 0$ at initial, intermediate and final instants.*

term determines whether each isosurface is inward or outward moving. The value of ν is set so that the resolution of the model is kept within the smallest voxel spacing; i.e.: ν cannot be set arbitrarily large, without incurring the risk of overshooting the boundary between successive time discretizations of expression 5.1.

An intuitive illustration of each term in expression (5.1) appears in figure 5.6, featuring once again an analogy of lesser dimension, an evolving 2D contour. These illustrations fully extend to each term of the 3D surface evolution model. In figure 5.6, the diffusive term smooths the corners of the rectangular zero-level contour, nudging it towards an elliptical shape (and ultimately towards a circle, i.e.: tending towards constant curvature κ). The hyperbolic term shrinks the rectangle while retaining its discontinuities. These effects can be seen as being applied to each isocontour in $\Psi(x, y, t)$, including the zero-level contour. The speed term slows the zero-level

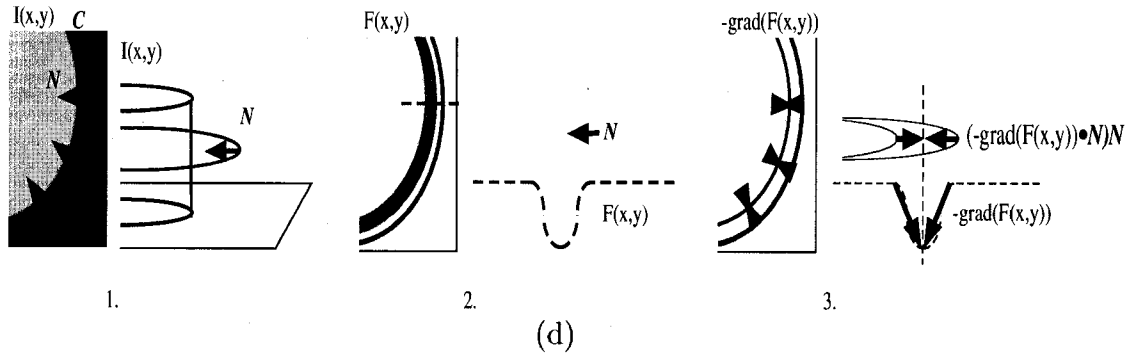
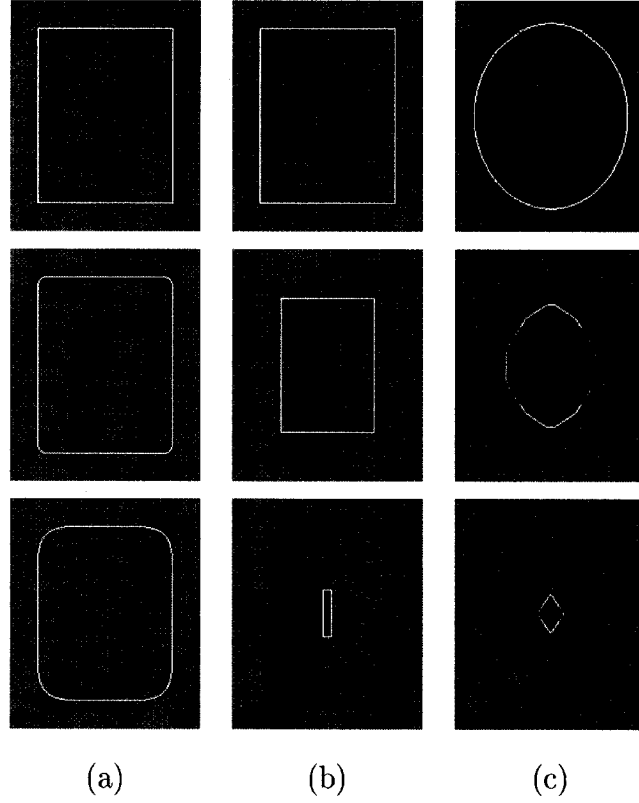


Figure 5.6 : *2D illustration of each term at work: (a) diffusive: at 0 sec., at 10 sec., and at 120 sec.; (b) hyperbolic: at 0 sec., at 10 sec., and at 30 sec.; (c) speed and doublet: the model behaviour is shown above in the absence of these two terms (compare with fig. 5.1): at 0 sec., at 15 sec., and at 30 sec.. (d) Illustration in 2D of doublet term: 1. raw 2D signal $I(x,y)$ with contour C approaching boundary; 2. speed function $F(x,y)$ inversely related to the gradient magnitude of the filtered signal; and 3. doublet term $(-\nabla F \cdot \mathbf{N})\mathbf{N}$: the factor $-\nabla F$ acts as a sink near the minimum, attractive where $(-\nabla F \cdot \mathbf{N})$ is positive, repulsive where $(-\nabla F \cdot \mathbf{N})$ is negative.*

contour down at the tissue boundary in the image, and the doublet term constitutes a local trough, centered at this boundary, which prevents the zero-level contour from overshooting it: in their absence, the zero-level contour never halts at the underlying boundary.

The doublet term is further considered in figure 5.6 (d), and is best explained in 3D surface or 2D contour evolution representation. The level set representation of expression (5.1) corresponds to the evolution of a surface $\mathcal{S}$, along the inward normal direction:

$$\frac{\partial \mathcal{S}}{\partial t} = (FH - \nabla F \cdot \mathbf{N}) \mathbf{N}, \quad (5.2)$$

where H is the mean curvature of $\mathcal{S}$, and $\mathbf{N}$ is its inward normal. If we consider the 2D analogical situation in figure 5.6 (d), we have an inward moving zero-level contour $\mathcal{C}$ whose evolution is given by $\partial \mathcal{C} / \partial t = (F\kappa - \nabla F \cdot \mathbf{N}) \mathbf{N}$, where this time κ is the curvature and $\mathbf{N}$ the inward normal of $\mathcal{C}$. In both the surface and contour expressions, the term $-\nabla F \cdot \mathbf{N}$ corresponds to the doublet term in level set representation. As shown here, the speed term F has a local trough at the boundary of interest, and the vector $-\nabla F$ points toward the minimum of this trough. For an inward moving contour, the dot product $-\nabla F \cdot \mathbf{N}$ outside the boundary is positive, and the doublet term is in the direction of the inward normal and pushes the zero-level contour towards the minimum. Inside the boundary, if the inward moving contour overshoots it, $-\nabla F \cdot \mathbf{N}$ becomes negative, and the doublet term points outwards and draws the contour back towards the minimum. This feature may be needed for a continuously-valued image-based speed function, such as one inversely proportional to a finite gradient magnitude, if this function in practice is never zero-valued.

5.3 A Concise Review of Model-based Surface Identification and Image Filtering

5.3.1 Model-based Surface Identification Techniques

Model-based surface identification in one volume (segmentation) or across a time sequence of volumes (tracking) expresses the identification problem as a model which reconciles the likely shape and/or dynamic behaviour of the surface, based on prior knowledge, with raw image data. Model-based surface tracking methods were surveyed in § 2.3.3. This section provides a brief overview of the literature on model-based surface identification in one volume. Other more exhaustive surveys of surface models are available: McNerney & Terzopoulos [251] emphasized various biomedical imaging applications, whereas Montagnat et al. [271] provided an elegant synthesis of various representations, dynamic equations, and constraints on deformations.

Surface models can broadly be categorized as physically-based and surface evolution-based. For physically-based deformable surface models, the basic idea is to view an object being identified as reaching an equilibrium between internal and external virtual forces. *Internal forces* include inertia, damping and strain, based on the current state which is assumed for a 3D surface. *External forces* are typically determined by image data. The relative motion undergone by the various parts of a surface is a result of the interaction of these two sets of forces, according to formal physical principles. Moreover, physical surface models generally adopt an *explicit* surface representation, in contrast to the *implicit* representation of surface evolution models.

Important early research includes Kass' *snakes* paper [179] and Terzopoulos' symmetry-seeking 3D shape model [391]. Kass expressed the 2D contour detection and tracking problem of intensity images as a minimization of the following energy functional:

$$E_{snake} = \int_0^1 E_{int}(\mathbf{v}(s)) + E_{image}(\mathbf{v}(s)) + E_{con}(\mathbf{v}(s)) ds, \quad (5.3)$$

which is a controlled-continuity spline [388] under the influence of image forces and

external constraint forces. Terzopoulos' model is given by the equation

$$\mu \frac{\partial^2 \nu}{\partial t^2} + \gamma \frac{\partial \nu}{\partial t} + \frac{\delta \mathcal{E}(\nu)}{\delta \nu} = \mathbf{f}(\nu), \quad (5.4)$$

and is analogous to a deformable sheet of elastic material, which is rolled to form a tube, through which passes a deformable spine, also of elastic material. Coupling forces try to make the shape retain its axial symmetry, and extrinsic forces constrain the shape to be consistent with one or more 2D image projections.

Pentland [306, 100] and Terzopoulos [389, 250] independently proposed a finite-element model approach to the numerical solution of deformable surface models, and consequently these surface models are based on the *finite-element equilibrium equation*, which has the form

$$\mathbf{M}\ddot{\mathbf{U}} + \mathbf{C}\dot{\mathbf{U}} + \mathbf{K}\mathbf{U} = \mathbf{R}, \quad (5.5)$$

where $\mathbf{M}$, $\mathbf{C}$, and $\mathbf{K}$ are virtual *mass*, *damping* and *stiffness* matrices, $\mathbf{U}$ is the displacement of the FEM nodes, and $\mathbf{R}$ is the sum of external forces, determined by image data. Pentland and Terzopoulos used a hybrid representation featuring a superquadric ellipsoid [23] upon which is grafted a displacement function, which in turn is estimated by finite-element modeling. This displacement function represents the difference between the simple superquadric shape and the final, more general shape which is more in keeping with image forces. However, if the model is not initialized close enough to the desired surface, short-range image forces may be unable to attract it. To alleviate this problem, Cohen proposed a *pressure force* to inflate the model towards the object surface [70]. In contrast with Pentland and Terzopoulos' approaches, Cohen's physical model was implemented with finite differences.

Refinements to explicit or physically-based surface models can roughly be categorized as follows:

- New formulations for *surface representation*. While Montagnat [271] has exhaustively reviewed the literature on surface representations, this review is restricted to a few elegant formulations. First, the surface models seen so far can broadly be described as *continuous models*. In contrast, some authors have

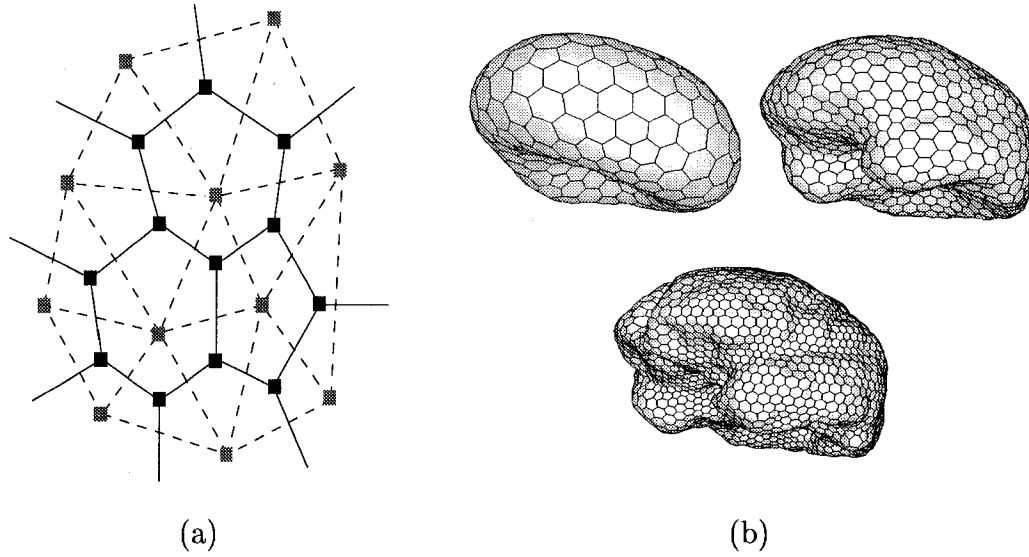


Figure 5.7 : *Left: illustration of 2-simplex mesh, shown in black, with the dual triangulation shown in grey; Right: simplex representation of a brain surface, shown for three levels of meshing resolution (courtesy of Johan Montagnat) [271].*

proposed *discrete models*, in particular the *simplex mesh* [83, 84, 86] and *oriented particle systems* [381]. The simplex mesh is defined as a set of vertices and a specific connectivity function: a k -simplex mesh connects each vertex to exactly $k + 1$ neighbours. Figure 5.7(a) illustrates a 2-simplex mesh (in black), as well as the duality of this representation with a triangulation (in grey). Each vertex submits to image-based external forces and to internal forces dictated by its local geometric relationship with its neighbours, featuring a tangential component which optimizes spacing, possibly involving a prior choice of mesh scale and a normal component which enforces surface smoothness. As shown in figure 5.7(b), the mesh scale of the simplex model is a selectable parameter. The oriented particle system expresses a surface as a set of discrete point masses, each endowed with a local coordinate system, whose interaction is modelled as an energy function which not only integrates “molecular” long-range attraction and short-range repulsion, but also locally enforces coplanarity as well as consistency in surface normal orientation and surface curvature. Other

important *continuous* surface representations are parametrized models exploiting the smoothness properties of B-splines [254, 212] and a Fourier description expressed in terms of orthonormal sinusoidal basis functions [373].

- Improvements to the *flexibility* of physical model formation. An important contribution made recently by McNerney et al. [252] and by Delingette et al. [85, 86] is the incorporation of topological adaptation to physical models. The ability to subdivide or merge is difficult to implement with continuous surface implementations: self-intersection and collision detection logic is far easier to manage with a set of discrete vertices. Consequently, in order to make Terzopoulos' FEM-based surface model topologically adaptive, McNerney exploits *affine cell decomposition*, which represents the spatial occupancy of a contour as a contiguous subset of a regular triangular grid, or that of a surface as a subset of a tetrahedral grid. The merging and splitting of a contour (or surface) is then managed by explicit rules governing its intersections with the edges (faces) of the triangular (tetrahedral) cells occupied by it. In a similar development, Delingette and Montagnat subdivide the image plane or volume into a rectilinear grid, tabulate the spatial occupancy of the contour or surface with a hash table, and implement the splitting and merging of simplex meshes on the basis of self-intersection or collisions detected from grid intersections. In both cases, the extension of topological adaptation from 2D contours to 3D surfaces is quite laborious, in comparison with level-set methods.
- Surface models with a *statistical foundation*. Explicit surface models discussed up to now can be seen as incorporating *generic* prior knowledge in order to stabilize the contour or surface identification. Recently, there has been a trend towards the use of *specific* prior knowledge of object shape. The first approaches of this type were *deformable templates* [222] where the prior model was individually tailored for each application. In contrast, Cootes [78] suggested an approach with potentially more widespread descriptive ability, termed *active shape models*, whose prior knowledge is based on the *principal component anal-*

ysis [117] of training images. 2D contour identification requires a training phase consisting of the following steps:

- *Manual identification* of several (N) training shapes, typically in 2D, each composed of: n points: the i th shape can be expressed as a vector

$$\mathbf{x}_i = (x_{i0}, y_{i0}, x_{i1}, y_{i1}, \dots, x_{in-1}, y_{in-1})^t. \quad (5.6)$$

- Computation of a *mean* $\bar{\mathbf{x}}$ and $2n \times 2n$ *covariance matrix* $\mathbf{S}$ of the N shapes:

$$\bar{\mathbf{x}} = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_i \quad \text{and} \quad \mathbf{S} = \frac{1}{N} \sum_{i=1}^N d\mathbf{x}_i d\mathbf{x}_i^t \quad \text{where} \quad d\mathbf{x}_i = \mathbf{x}_i - \bar{\mathbf{x}}. \quad (5.7)$$

- Computation of the *modes of variation* $\mathbf{p}_k$, $k = 1 \dots 2n$, of the points of the shape, namely the eigenvectors of $\mathbf{S}$ such that

$$\mathbf{S}\mathbf{p}_k = \lambda_k \mathbf{p}_k, \quad (5.8)$$

and finding a small subset of $\mathbf{p}_k$, coinciding with the largest eigenvalues, which accounts for a significant proportion of the shape variability.

A new shape to be identified is seen as a linear combination of these modes,

$$\mathbf{x} = \bar{\mathbf{x}} + \mathbf{P}\mathbf{b} \quad \text{where} \quad \mathbf{P} = (\mathbf{p}_1, \dots, \mathbf{p}_t) \quad \text{and} \quad \mathbf{b} = (b_1, \dots, b_t)^t, \quad (5.9)$$

once the pose and scale are ascertained (in practice, pose, scale and the weights $\mathbf{b}$ are estimated iteratively). While this method is attractive for medical imaging, its training phase is extremely labour intensive, which makes it difficult to apply to 3D surface identification. Moreover, the assumption of a 1-1 point mapping, between the average shape and each training shape, and between the average and unknown data shapes, may break down under extremely variable anatomy, such as the human cortex.

In contrast to the preceding models, those developed by Caselles et al. [56] and by Malladi et al. [237] have a geometric, rather than physical, interpretation, and are generally based on the *evolution equation of a curve or surface*. Moreover, as seen in

§ 5.2, the contour or surface is *implicit*: it is viewed as the zero level of an imbedding function $\Psi(x, y)$ or $\Psi(x, y, z)$ defined over the whole 2D or 3D image domain. Caselles and Malladi first proposed an imbedding function $\Psi(x, y)$ which evolves according to the equation:

$$\frac{\partial \Psi}{\partial t} = F(g(I)) \|\nabla \Psi\| [\nu + \kappa(\Psi)] . \quad (5.10)$$

In other words, the original contour evolution model features the image-based *speed term* F which approaches zero near anatomical boundaries, the *diffusive* smoothing term based on the *curvature* $\kappa = \text{div} \left(\frac{\nabla \Psi(x, y)}{\|\nabla \Psi(x, y)\|} \right)$ of the contour, and a *hyperbolic* motion term ν which nudges the contour inward or outward, based on the sign of ν . This model does not feature the doublet term, which further attracts the model to anatomical boundaries and which appears in expression (5.1). Expression (5.10) is fully extensible to identifying surfaces in a 3D volume, given the smoothing term $H = \text{div} \left(\frac{\nabla \Psi(x, y, z)}{\|\nabla \Psi(x, y, z)\|} \right)$ which corresponds to the *mean curvature* of the $\{\Psi = 0\}$ isosurface [58].

Recent improvements include *weighted surface-area* [57, 187, 419] and *volume* [350] minimizing evolution equations, the application of the *Fast Marching* [341] techniques to accelerate the computation of the surface evolution model [238], and the modeling of *two coupled surfaces* which evolve to coincide with the inner and outer cortex [422]. The weighted surface-area evolution, or geodesic, model features the now-familiar doublet term which is added to the classical evolution model expression (5.10), producing expression (5.1), the model adopted and refined in this thesis research. As emphasized in appendix A, the original Caselles/Malladi model shrinks the zero-level isosurface in a manner which decreases most rapidly a *Euclidean* surface-area functional,

$$A(t) \equiv \int_0^1 \int_0^1 dS \text{ where } dS \equiv (\|\mathcal{S}_u\|^2 \|\mathcal{S}_v\|^2 - \langle \mathcal{S}_u, \mathcal{S}_v \rangle^2)^{1/2} dudv . \quad (5.11)$$

whereas the geodesic model decreases an *image-weighted* surface-area functional,

$$A(t)_F \equiv \int_0^1 \int_0^1 dS_F \text{ where } dS_F \equiv F dS , \quad (5.12)$$

where F is the image-based speed function, thereby attracting the isosurface to a minimum of a potential energy functional which factors in image gradients. Siddiqi

et al. [350] have proposed another evolution model which shrinks a 3D isosurface in a manner that most quickly minimizes an image-weighted volume functional. This evolution equation replaces the hyperbolic term in equation (5.1), which is seen as a somewhat artificial term based on Cohen's inflation force [70], with a new inflation term which is derived from a formal minimization.

Next, Malladi et al. proposed a fast 3D medical segmentation technique based on the Fast Marching (FM) method [341], which can be used to solve a surface evolution equation of the type $\Psi_t + F(x, y, z)\|\nabla\Psi\| = 0$, known as a Hamilton-Jacobi equation (which does not include a diffusive smoothing term). This technique solves the surface evolution equation by considering the *arrival time* $T(x, y, z)$ of the surface,

$$\|\nabla T\|F = 1 \quad \text{where} \quad \nabla T = \left(\frac{\partial}{\partial x} + \frac{\partial}{\partial y} + \frac{\partial}{\partial z}\right)T. \quad (5.13)$$

Here ∇T , the gradient of the arrival time of a surface at a point in the image domain, is inversely proportional to its speed of the front at that point. The FM method produces a physically reasonable *weak* solution from an *upwind* numerical approximation, which in turn relies on *min-heap* sorting and a *narrow band* implementation, as discussed in § 6.4.1, in order to achieve computational efficiencies. I adopt this method to efficiently compute a distance map from a surface (refer to § 6.4.1), assuming a constant unit speed $F = 1$, whereas Malladi opts for an image-based speed

$$F(\mathbf{x}) = \exp \left[-\alpha \|\nabla \tilde{I}(\mathbf{x})\| \right], \quad (5.14)$$

which produces a surface that quickly captures the anatomical boundary of interest, albeit requiring a final smoothing stage.

Finally, Zeng uses coupled equations of two implicit surfaces [422],

$$\begin{aligned} \Psi_{in_t} + F_{in}\|\nabla\Psi_{in}\| &= 0 \\ \Psi_{out_t} + F_{out}\|\nabla\Psi_{out}\| &= 0, \end{aligned} \quad (5.15)$$

to determine the evolution of her model towards the inner and outer cortical surfaces respectively. The speed terms F_{in} and F_{out} are comprised of an image-based function

g and coupling factor h :

$$\begin{aligned} F_{in} &= g(p_{GM-WM}(x, y, z)) h(\Psi_{out}) \\ F_{out} &= g(p_{CSF-GM}(x, y, z)) h(\Psi_{in}). \end{aligned} \quad (5.16)$$

The image-based term estimates the probability of the voxel being on the WM/GM boundary, in the first case, or on the GM/CSF boundary, in the second case, and sets the speed to 0 as this probability tends to 1. The coupling factor maintains a non-zero speed provided that each surface is within some desirable range of the other, as the two evolve, and halts the corresponding zero-level isosurface otherwise.

5.3.2 Image Filtering Techniques

As discussed in the introduction of this chapter, model-based surface identification within an MRI volume requires an image-based force or stopping term to bind the model surface to the relevant anatomical boundary. This can take the form of an intensity-based term which benefits from separating noise from true tissue magnetization characteristics, or a gradient-based term whose derivative computation requires a preprocessing stage in order to be tolerant to noise. While a surface model can overcome minor low-contrast areas or corrupted pixels, partial volume effects and artefacts typically either are an inherent limitation of the technology or benefit from addressing the problem by choosing appropriate image acquisition parameters [147]. This survey is therefore restricted to noise reduction, and discusses existing image preprocessing techniques used for inverse problems such as edge-detection, feature extraction, registration and segmentation.

The classical approach to preprocessing, particularly in edge detection, is isotropic filtering. This typically takes the form of a convolution with an isotropic (near-) Gaussian kernel [241, 54, 270] (or its Laplacian, for edge-detection). Marr & Hildreth [241] proposed a theory of edge detection whose frequency and spatial localization constraints are conflicting and shown to be optimally met by integrating a convolution with a Gaussian kernel into this detection. Canny [54] and Poggio et al. [313] independently suggested near-Gaussian operators, from a rigorous optimization perspec-

tive. Moreover, these operators have subsequently been implemented as recursive filters [87, 88, 89, 403], vastly improving their computational efficiency.

The transition from Gaussian isotropic filters to anisotropic filters is based on work by Koenderink [195], by Hummel [165], and by Perona and Malik [308]. While these authors discussed 2D implementations, each framework carries over to volumetric data. Koenderink and Hummel have both described the Gaussian filter as the solution of the *Heat Conduction Equation* (also known as the *Diffusion Equation*) which is expressed as

$$\psi_t = \nabla^2 \psi \equiv \psi_{xx} + \psi_{yy}, \text{ in 2D.} \quad (5.17)$$

Koenderink's motivation was to study descriptions of an image at various scales simultaneously, as a family. He posited that the family of images should be *causal*, *homogeneous* and *isotropic*. The first criterion implies that no spurious contour should be generated as resolution is diminished (or as t increases), and the latter two imply that the blurring must be space invariant (independent of x and y). These criteria lead to the Heat Equation formulation, where the simulated time of the equation is related to Gaussian scale by the expression $2\sigma^2 = 4t$. Hummel [165] subsequently observed that Koenderink's causality criterion only leads to a Gaussian convolution under fairly severe restrictions.

Perona and Malik [308] dropped the homogeneity and isotropy criteria in favour of *immediate localization* and *piecewise smoothing*, which imply that the boundaries of regions (of consistent intensity) in a 2D image should be sharp and coincide with "semantically meaningful" boundaries in the physical world, and that intra-region smoothing should occur preferentially over inter-region smoothing. This was accomplished with a model which was shown not to generate spurious contours as resolution is diminished, and featuring a spatially variant conduction coefficient $c(x, y, t)$, which yields anisotropic diffusion:

$$\psi_t = \nabla \cdot (c(x, y, t) \nabla \psi) = c(x, y, t) \nabla^2 \psi + \nabla c \cdot \nabla \psi \quad \text{where} \quad \psi(x, y, 0) = I(x, y). \quad (5.18)$$

According to the localization and smoothing criteria, c should be set to 1 within each region and to 0 at the boundaries, but the region boundaries are not known a priori.

However, the position of a boundary can be estimated with the magnitude of the gradient of the previous time iteration, i.e.:

$$c(x, y, t) = g(\|\nabla\psi\|) \quad \text{where, e.g.: } g(s) = e^{-(s/K)^2} \text{ or } g(s) = 1/\left[1 + \left(\frac{s}{K}\right)^2\right], \quad (5.19)$$

This framework has subsequently been extended to volumetric MR data by Gerig et al. [125].

The formulation of expression (5.18) has subsequently led a number of researchers to investigate the mathematical properties of equations of this type [60, 5, 189, 392, 59]. Once again, this research has been carried out on 2D images but is extensible to volumetric data. As pointed out by Catté et al. [60], the Perona model suffers from two shortcomings, namely a potential for unbounded oscillations in the presence of noise and for highly similar pictures having vastly dissimilar edge maps. Catté replaced the gradient in expression 5.19 with a Gaussian smoothed estimate of the gradient $c(x, y, t) = g(\|\nabla G_\sigma * \psi\|)$ and proved the existence and uniqueness of the model. Alvarez et al. [5] suggested a class of models which have a clearer geometric interpretation and more stability in the limit as $\sigma \rightarrow 0$ than Catté's, the simplest of which is characterized by the level-sets of its solution moving in the normal direction with a speed proportional to their mean curvature.

Finally, Kimia & Siddiqi [189] reinforced the link between the Level Set framework and anisotropic diffusion. By considering each iso-intensity level in a 2D image as a contour C , they have cast image smoothing as an exercise in curvature-dependent shape smoothing, along the lines of the “reaction-diffusion” framework proposed by Kimia et al. [190]. This framework features *curve shortening flow*, $\partial C/\partial t = \kappa \mathbf{n}$, whose solution is made tolerant of singularities, or *shocks*, by imbedding it into a higher dimension function, initialized as a distance map from the contour, which evolves according to the expression:

$$\psi_t = \kappa \|\nabla\psi\| \quad \text{where} \quad \kappa = \frac{(\psi_{xx}\psi_y^2 - 2\psi_{xy}\psi_x\psi_y + \psi_{yy}\psi_x^2)}{(\psi_x^2 + \psi_y^2)^{3/2}}. \quad (5.20)$$

Furthermore, each iso-intensity level can evolve, according to its curvature, independently of other levels, which entails that no distance map need be computed for each

intensity level, and that the raw image itself can be used to initialize ψ in expression (5.20). Finally, recent research in diffusion equations in image processing can be found in [392] and [59].

5.4 Conclusions

This chapter began by providing some motivation for adopting a model-based approach for brain surface identification, emphasizing its reproducibility, reduced user intervention, incorporation of prior knowledge, and three-dimensional framework. A distinction was made between two kinds of surface models, namely physically-based and surface evolution models. Section 5.2 then offered an intuitive description of surface evolution models, which are characterized by a partial differential equation in simulated time and space of a function $\Psi(x, y, z, t)$ defined over the whole image domain. Surface evolution models are also characterized by an imbedding framework whereby the evolving surface of interest $\mathcal{S}(t)$, initialized from the user-defined surface at $t = 0$ and contracting inward or expanding outward to the anatomical boundary of interest, is defined implicitly and coincides with the $\{\Psi = 0\}$ isosurface for all values of t . This imbedding allows surface evolution models to merge and break apart, as dictated by the nature of the image information coinciding with anatomical boundaries.

Subsequently, relevant surveys were presented on MR signal and noise modeling, on surface models of both physical and evolution types, and on image filtering techniques commonly used prior to inverse problems, such as edge detection and segmentation. The MR signal was described as a complex-valued function whose magnitude is retained, and where each component is corrupted by Gaussian noise, with the consequence that the magnitude of the signal exhibits Rayleigh-distributed noise in the background and Rician-distributed noise in the volume occupied by the subject. Surface models were categorized into those based on an explicit representation and on a physical equilibrium between virtual internal properties of the model and image-based external forces, and those based on the evolution of an implicit surface.

Finally, the survey on image filtering traced its evolution from the isotropic Gaussian operator, through its Heat Equation implementation, to anisotropic diffusion techniques which are generalizations of the Heat Equation.

Chapter 6

Model-Based Brain Surface Identification in MRI: Implementation Details

6.1 Surface Evolution Model

To recapitulate the overview presented in the previous chapter, the anatomical surface identification technique presented here is a model-based technique, known as a surface evolution model, which is characterized by embedding the identification of a 3D surface $\mathcal{S}$ into the evaluation of a function Ψ whose domain spans the image volume:

$$\{X \in \mathbb{R}^3 \text{ such that } \mathcal{S}(t) \equiv X : \Psi(X, t) = 0\} . \quad (6.1)$$

The mathematical equation expressing the movement of the surface model is as follows [58, 419]:

$$\frac{\partial \Psi}{\partial t} = F(x, y, z) \|\nabla \Psi\| \left[\operatorname{div} \left(\frac{\nabla \Psi}{\|\nabla \Psi\|} \right) + \nu \right] + \nabla F \cdot \nabla \Psi , \quad (6.2)$$

where $\operatorname{div}(\nabla \Psi / \|\nabla \Psi\|)$ represents the mean curvature H of an isosurface of the function Ψ . This model features a *diffusive* smoothing term $\|\nabla \Psi\| \operatorname{div}(\nabla \Psi / \|\nabla \Psi\|)$, a *hyperbolic* motion term $\|\nabla \Psi\| \nu$, and two image terms: a *speed* function F and a

doublet term $\nabla F \cdot \nabla \Psi$ which serve to bind the evolving surface to the anatomical boundary of interest. Moreover, the model is initialized by one or more user-defined 3D surfaces which fully contain, or are fully contained by, the anatomical boundary. The model evolves in such a manner that the zero-level isosurface is nudged inwards or outwards until the image terms bind it to the boundary. The embedding function Ψ is initialized as a *signed distance map* from the user-defined surface(s). This *geodesic* model, so-called because the contour/surface evolves along the shortest path to the minimum of a potential energy function, was originally derived by Caselles [57] and Tannenbaum [187, 419] and their respective collaborators, and this derivation is outlined in appendix B.

In most existing implementations of surface evolution models, the speed term is strictly a function of the image gradient, i.e.:

$$F = \frac{1}{1 + \|\nabla \tilde{I}(x, y, z)\|^n}, \quad (6.3)$$

where typically $n=1, 2$ or 3 and $\tilde{I}(x, y, z)$ usually is a Gaussian-filtered image (MR) volume. However, as we will see in the next section, this approach has some limitations when it comes to identifying the brain surface, such as a lack of T1 contrast (in the absence of contrast agents) between grey matter and the sagittal sinus and between grey matter and muscle tissue, particularly near the eyes, possibly entailing a bleeding effect outside the brain surface. To alleviate this problem, the model is endowed with higher-level anatomic information, namely by first running a *voxel classification* algorithm on the image volume(s), and to compute a speed function which restricts the surface model to expand only within white and grey matter (WM and GM respectively), according to the following discrete relation:

$$\begin{aligned} F(x, y, z) &= 1.0 \text{ if } \{(x, y, z) \in WM\} \text{ or } \{(x, y, z) \in GM \text{ and near } WM\}, \\ &= 0 \text{ otherwise.} \end{aligned} \quad (6.4)$$

In practice, this implementation features the following stages, as illustrated in figures 6.1 and 6.2:

- Preprocessing of the MR volume(s) for the subsequent computation of the

image-based speed term. Given that one would want to eventually *register* the identified cortical surface, it is preferable to preprocess the MR volume with an anisotropic diffusion operator rather than an isotropic filter, to limit unwanted blurring across relevant anatomical surfaces. The preprocessing stage is Kimia & Siddiqi’s *curvature flow* [189], which is implemented within the level-set framework, as described in § 7.2.

- An image-based speed term computation stage designed to allow the expansion of the model through white matter, to capture its boundary: $F(x, y, z) = 1.0$ if $\{(x, y, z) \in WM\}$ and $F(x, y, z) = 0.0$ otherwise.
- The identification of the WM boundary, which starts as a set of small spheres, positioned by the user in white matter, as the initial zero-level isosurface at $t = 0$. This kind of initialization from classifier WM training points, which are idealized as small spheres and eventually merge together as they expand, is an immediate benefit of a topological adaptive surface model. The surface evolution model presented here is then computed over (simulated) time by alternating between the following two methods:
 - The initialization of $\Psi(t_{l+1})$ as a signed distance map from the $\{\Psi(t_l) = 0\}$ isosurface at the previous iteration, which is computed with the Fast Marching method. This method is briefly described and validated in § 7.4.
 - The numerical implementation of expression (6.2) as a finite difference equation at the voxel $\mathbf{i} = (i, j, k)$ and at time $t_{l+1} \equiv (l + 1)\Delta t$, namely:

$$\begin{aligned} \Psi(\mathbf{i}, (l + 1)\Delta t) = & \Psi(\mathbf{i}, l\Delta t) + \\ & \Delta t \left\{ F(\check{\mathbf{i}}) \left\| (\Psi_x(\mathbf{i}), \Psi_y(\mathbf{i}), \Psi_z(\mathbf{i})) \right\| [H(\Psi(\mathbf{i})) + \nu] + \right. \\ & \left. (F_x(\check{\mathbf{i}}), F_y(\check{\mathbf{i}}), F_z(\check{\mathbf{i}})) \cdot (\Psi_x(\mathbf{i}), \Psi_y(\mathbf{i}), \Psi_z(\mathbf{i})) \right\} . \end{aligned} \quad (6.5)$$

The numerical implementation of the second method follows standard classical level sets procedure. The label $\check{\mathbf{i}}$ denotes the voxel on the $\{\Psi(t_l) = 0\}$ isosurface which is closest to the voxel of interest $\mathbf{i}$. Since the image information only coincides with the zero-level, it must be “extended” from

zero-level isosurface voxels to other voxels where expression (6.5) is evaluated.

Note that the narrow band restriction applies to both methods, whereby each is computed only within a thin volumetric shell close to the evolving $\{\Psi(t) = 0\}$ isosurface.

- A recomputation of the image-based speed term, according to expression (6.4), to allow expansion of the surface model through grey matter “close” to the WM boundary, to capture the outer brain surface.
- The expansion of the $\{\Psi(t) = 0\}$ isosurface to coincide with the outer cerebral and cerebellar surface, again based on alternating between signed distance map computation and the numerical evaluation of the surface evolution model.
- Post-processing of the outer brain surface to produce a **Closest Point Map**, for the subsequent non-rigid registration with the intra-operative range-based surface. This is simply the Fast Marching method alluded to in stage 1 of the surface model, applied not to computing a distance map from the evolving zero-level isosurface within a narrow band, but to computing distance and closest point labels from the final brain surface everywhere in the volume. As seen in chapters 3 and 4, the identity of the brain surface point closest to each voxel is relevant to the subsequent range-MRI surface registration in the OR.

The rest of this chapter is divided as follows. Details of the tissue class-based speed function are given in § 6.2. The choice of voxel classifier is discussed in § 6.3. It is preferable to preprocess the MR volume with an anisotropic diffusion operator rather than an isotropic filter, to limit unwanted blurring across relevant anatomical surfaces which could adversely affect the subsequent registration. Isotropic blurring is similar to a partial volume effect, in that voxels blurred in this manner contain heterogeneous information about neighbouring tissues. Consequently, limiting blurring normal to anatomical boundaries is also beneficial to classifier performance, as seen in § 7.3. The preprocessing stage is Kimia & Siddiqi’s *curvature flow* [189], which is implemented

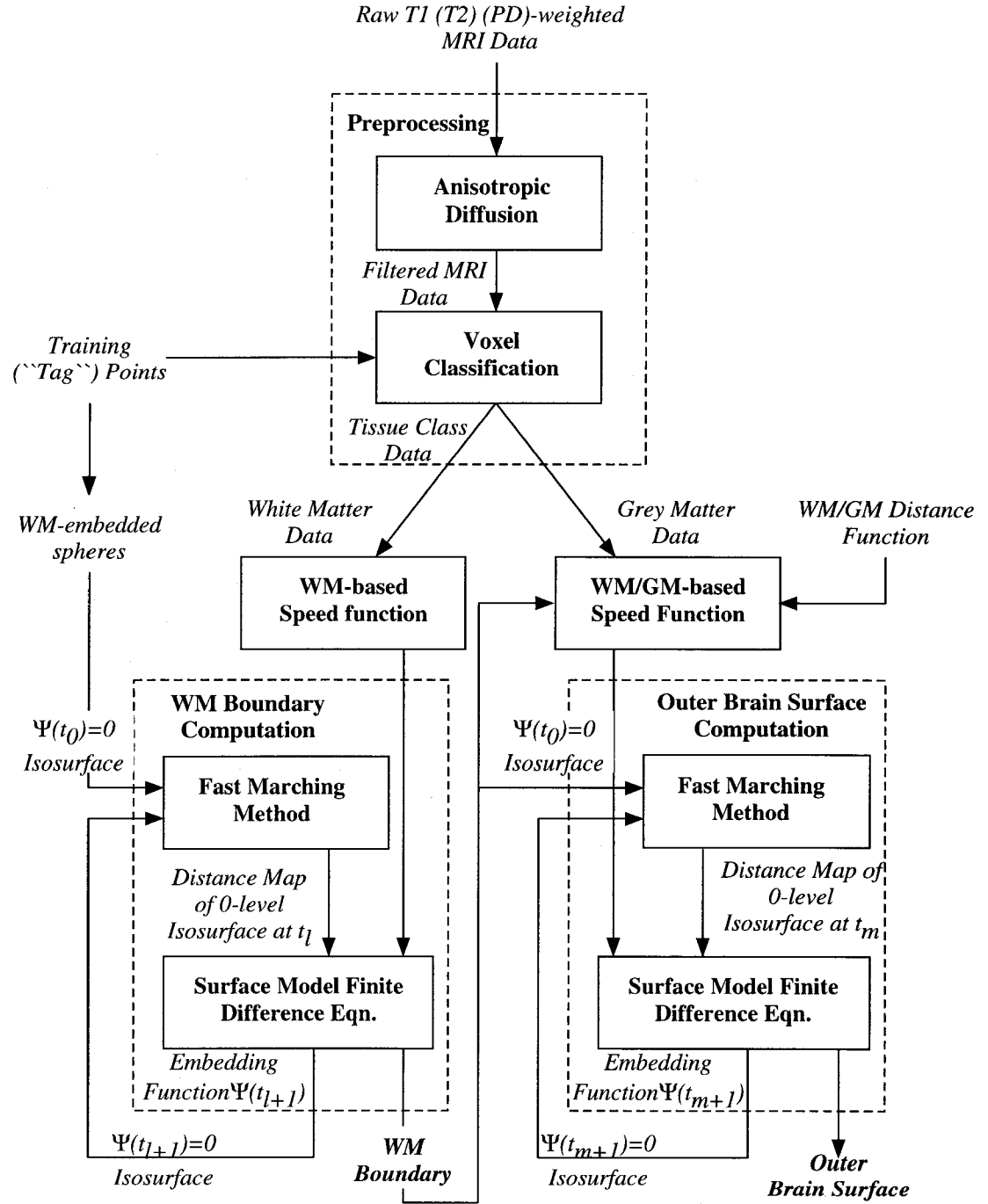


Figure 6.1 : Algorithmic description of the tissue class-based surface evolution model framework presented in this thesis.

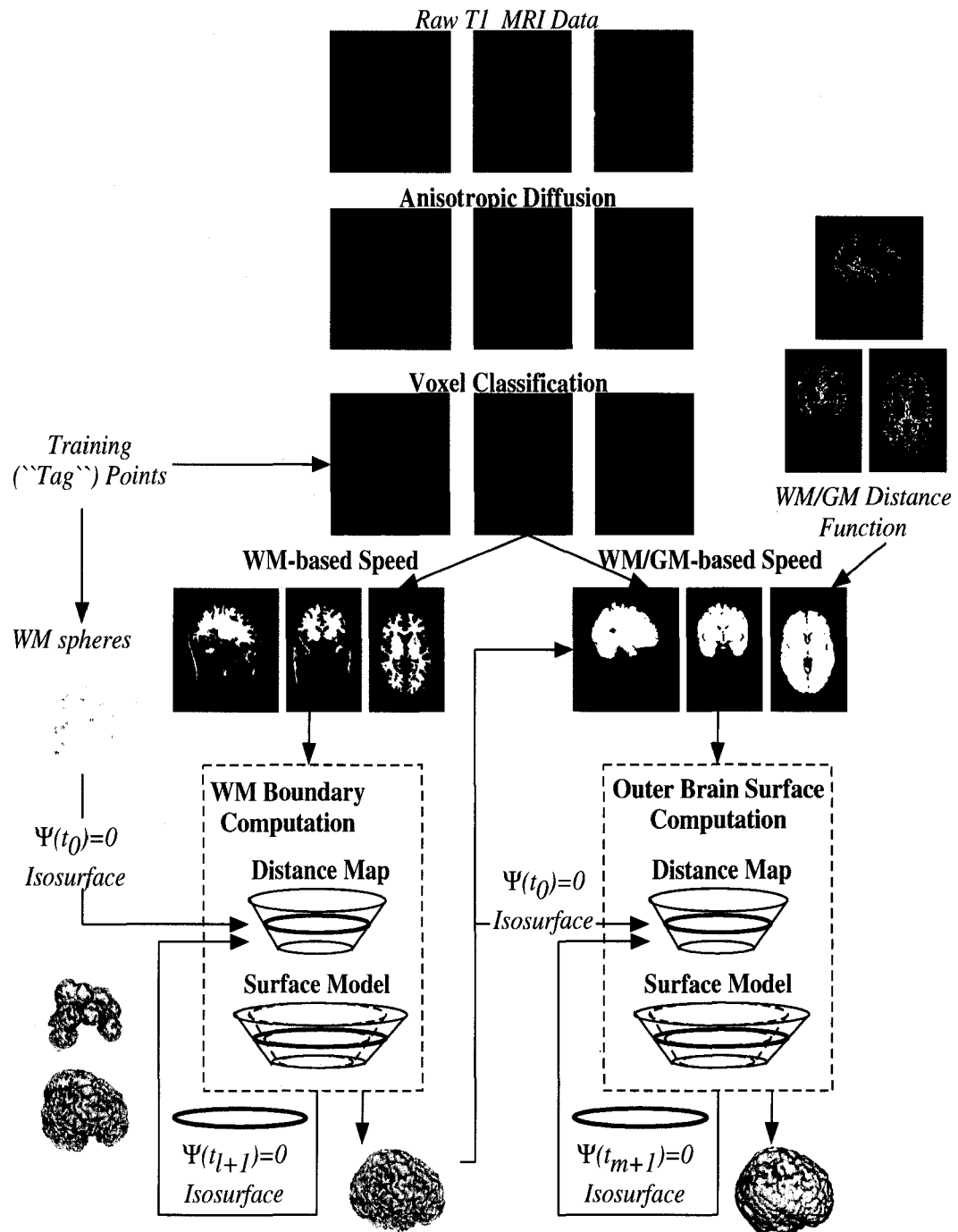


Figure 6.2 : Pictorial description of the tissue class-based surface evolution model framework presented in this thesis.

within the level-set framework, as described in the survey on image filtering in § 5.3.2. Finally, I discuss specifics of the numerical implementation of the surface model in § 6.4. In § 6.4.1 I elaborate on the *Fast Marching* (FM) method used to compute a distance map from a zero-level isosurface, which is the first stage of the surface model. This section emphasizes that the FM algorithm not only initializes the embedding function $\Psi(x, y, z)$ of the surface model, but has other uses as well. In particular the FM method also yields from the final outer brain surface $\mathcal{S}(t_f)$ a Closest Point Map, which is used in the surface registration.

The basic tools for validating all aspects of the computation of the surface model, in the form of a MRI simulator derived from first principles and implemented with an extensible object-oriented software architecture, and a realistic digital brain phantom based on an expertly labelled single-subject average of 27 scans, are discussed at the beginning of chapter 7. The results of the validation of the anisotropic diffusion, of the voxel classification, of the Distance and Closest Point Map, and of the surface model itself, are also deferred until chapter 7.

6.2 Integrating Anatomical Information into the Model: the WM/GM Distance Function

Surface evolution models have recently been demonstrated on volumetric MR brain data [238, 422]. These techniques are founded on an outward moving zero-level isosurface initialized completely within the brain (typically a few small spheres embedded in white matter). An inward-moving surface initialized at the skin surface is conceivable, but not likely to produce good results at the base of the brain. This area features muscle tissue distal to the brain, which is characterized by T1-weighted intensity and contrast similar to GM tissue, so that an inward moving surface model could inappropriately be halted on the muscle tissue boundary. Hence the predominance of outward-moving approaches initialized from within.

However, the problem faced by this approach is the prevention of the evolving

surface from becoming “snagged” in the gradients not coinciding with the outer brain surface. One example of this situation is the white-grey matter boundary, as illustrated figure 6.3(a). A simple solution to this snagging problem would appear to be to only consider very strong intensity gradients, coinciding with the outer GM and lesion boundaries, in order to halt the moving surface model. Strong gradients could be assessed on the basis of a threshold on gradient magnitude. But automatically finding such a threshold is elusive, as is guaranteeing that there won’t be gaps in the halting force on the cortical surface, where the local intensity gradient falls below this threshold. The bleeding effect that ensues, through such gaps, is illustrated in figure 6.3(b). While the diffusive smoothing term can alleviate this bleeding effect, its action is limited to very small low-gradient regions. As the surface evolution model moves across small low-contrast areas, the bleeding effect produces a zero-level isosurface that is locally characterized by a very large Mean Curvature, and the diffusive term tends to push the isosurface back so as to reduce this curvature. However, for larger gaps the hyperbolic term pushes the surface forward in a manner that cannot be overcome by curvature-dependent smoothing. This behaviour is a consequence of the topological adaptability of the model. Moreover, a surface initialized with a combination of spheres embedded in both white and grey matter might do better, but it would still run into low contrast areas near the eyes and sagittal sinus, once again causing the surface to bleed through.

In short, the exclusive use of gradient information to halt a surface evolution model, in the absence of strong tensile properties, has limitations. One way to overcome these limitations is by identifying first the grey-white matter boundary, and then using it to anchor the identification of the outer brain surface. Zeng [422] and MacDonald [230] both proposed surface models, each of which features two evolving surfaces bound to each other by a distance constraint and that expand to coincide with the inner and outer surfaces of a healthy cortex. This elegant idea confers on their respective models higher-level anatomical information, which stabilizes their convergence to inner and outer cortical surfaces. However, the assumption of the white matter boundary always being contained in a layer of grey matter may break down in

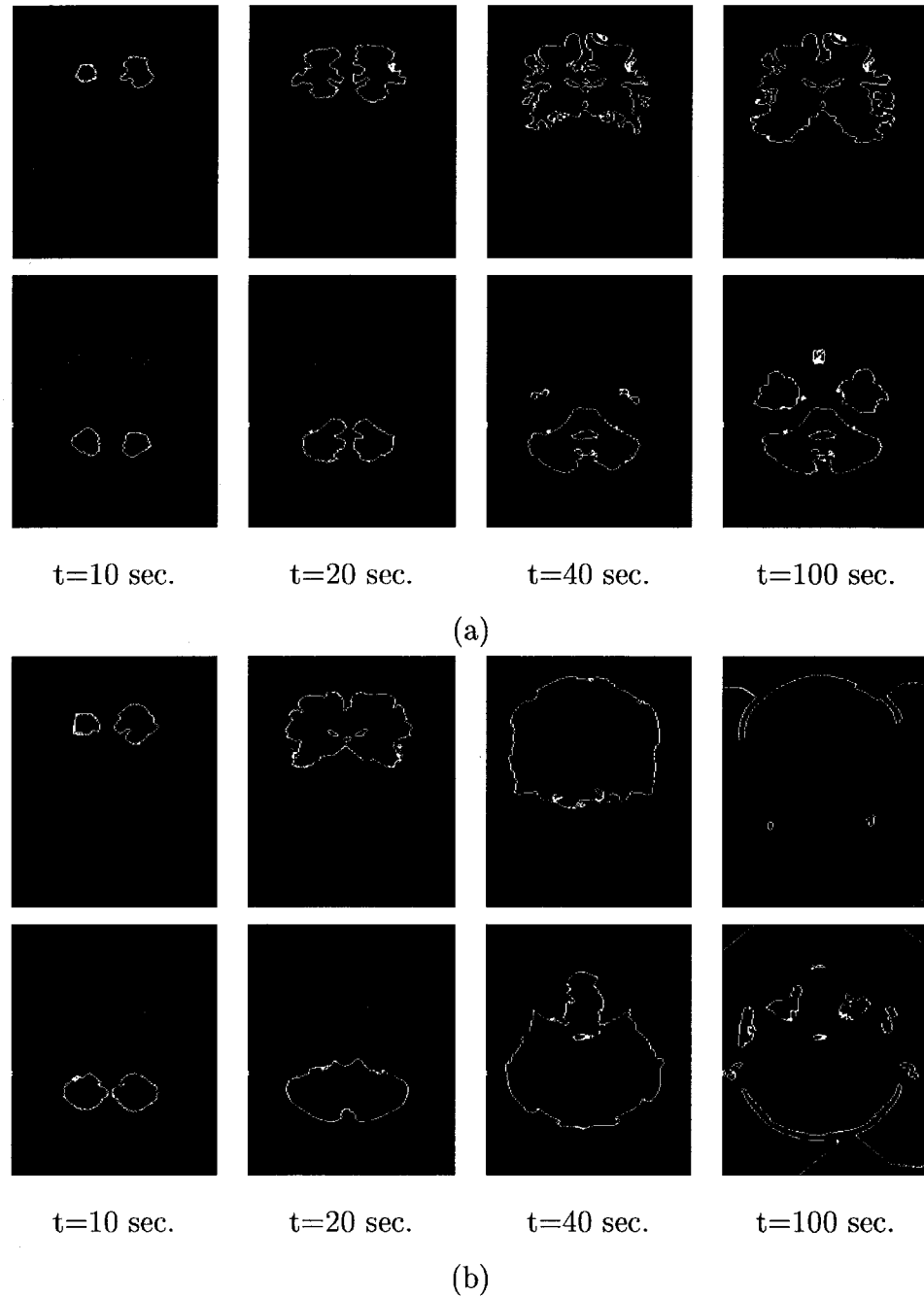


Figure 6.3 : *Axial and coronal views of evolving cortical surface model with its image term solely based on gradient magnitude. (a) With a low gradient threshold (allowing points of low gradient magnitude to halt the zero-level surface), the surface may converge to the WM boundary rather than the outer brain boundary. (b) With a higher threshold, the surface can bleed through the outer brain boundary, and never actually converge to a solution.*

the presence of pathologies. A lesion embedded in white matter presupposes null grey matter thickness and violates both models. In addition, the tight distance threshold which is appropriate for the outer cortex, useful to prevent the surface model from bleeding into the eyeballs or sagittal sinus, is inappropriate for the cerebellum and for the area near the ventricles, which are characterized by comparatively large expanses of grey matter. In other words, a constant distance threshold, enforcing an allowable range of thicknesses of grey matter, is not likely to identify the cortical, ventricular and cerebellar surfaces with equal effectiveness¹.

For these reasons, I chose to provide the surface model with higher-level anatomical knowledge by integrating the results of a *voxel classification algorithm* [117, 67, 196] into the image-based speed term. To achieve this, a classifier maps an image voxel, represented for D modalities as a D -dimensional feature vector in $\mathbb{R}^D$, $\mathbf{x} = (x_1, x_2, \dots, x_D)$, onto a tissue class C_j , where $j = 1, \dots, c$, and c is the assumed total number of tissue classes. This concept is illustrated by figure 6.5. The presence of pathology suggests that each separate patient data set requires a separate training stage, which precludes a classifier that relies on a large training set. A *minimally supervised* technique is sought, whereby an expert can provide anatomical knowledge by simply clicking on a few (5-20) training points for each class. The classifier training points for white matter can in turn also serve as the centers of the outward moving spheres used to initialize the surface model. Spheres centered at grey matter points could also be recruited, but the outer brain surface could not then be anchored by the WM boundary. The model moves outward from embedded WM points to the WM boundary, determines GM points “close” to this WM boundary, and moves to the outer boundary of these GM points. This method has the following advantages:

- Integrating tissue class information into the speed term constrains the model in a manner comparable to the methods of Zeng and MacDonald: the outer brain surface *evolves only in voxels classified as white matter or as grey matter and*

¹Additionally, from a theoretical standpoint, a surface evolution model based on coupled equations does not share the optimality qualities of the single-surface model; i.e.: its flow is not along a geodesic path to a potential energy minimum, so stable convergence to a solution is not guaranteed.

deemed close to white matter, thereby *limiting undesired bleeding effects*. Moreover, there is no reason for this distance tolerance, which enforces grey/white matter boundary proximity, to be a constant over the whole volume. Such a **WM/GM Distance Function** (WDF), in order to be useful over a large variety of individual anatomies, should be simple and robustly computed. In order to prevent artefacts in the brain surface, it should also be smooth over the whole volume. As illustrated in figure 6.4, this function is set to some maximum value ε_{in} at the centroid, and tends to a minimum parameter ε_{out} , which also varies spatially.

- This technique, which exploits a speed function rich in anatomical information, is highly *flexible*. Pathologies are easily dealt with, provided they can be discriminated from white and grey matter [43], perhaps through the use of a contrast agent [48]: voxels of a class other than the preceding two are characterized by a null speed function value, causing the zero-level isosurface to halt at that tissue boundary².
- This approach to constituting a speed function also provides an efficient way of taking available *multiple scan modalities* into account: the classifier simply maps the parameter space encompassing T1-, T2- and/or proton density-weighted volumes to a tissue class volume spatially coinciding with this image domain.
- This approach is easily adapted to *other anatomical applications*, at the cost of a different voxel classification training, and alternate use of voxel class information, possibly involving some sort of distance constraint between two classes.

The ability to cope with conflicting requirements in white to grey matter boundary proximity is handled by the nature of the WDF:

- It is tightest at the cortex (above the centroid along the proximal-distal axis), particularly near the sagittal sinus.

²For large pathologies which would prevent the WM from anchoring the outer brain surface evolution, both the pathology and WM could be used together to stabilize the evolution of the zero-level surface through GM.

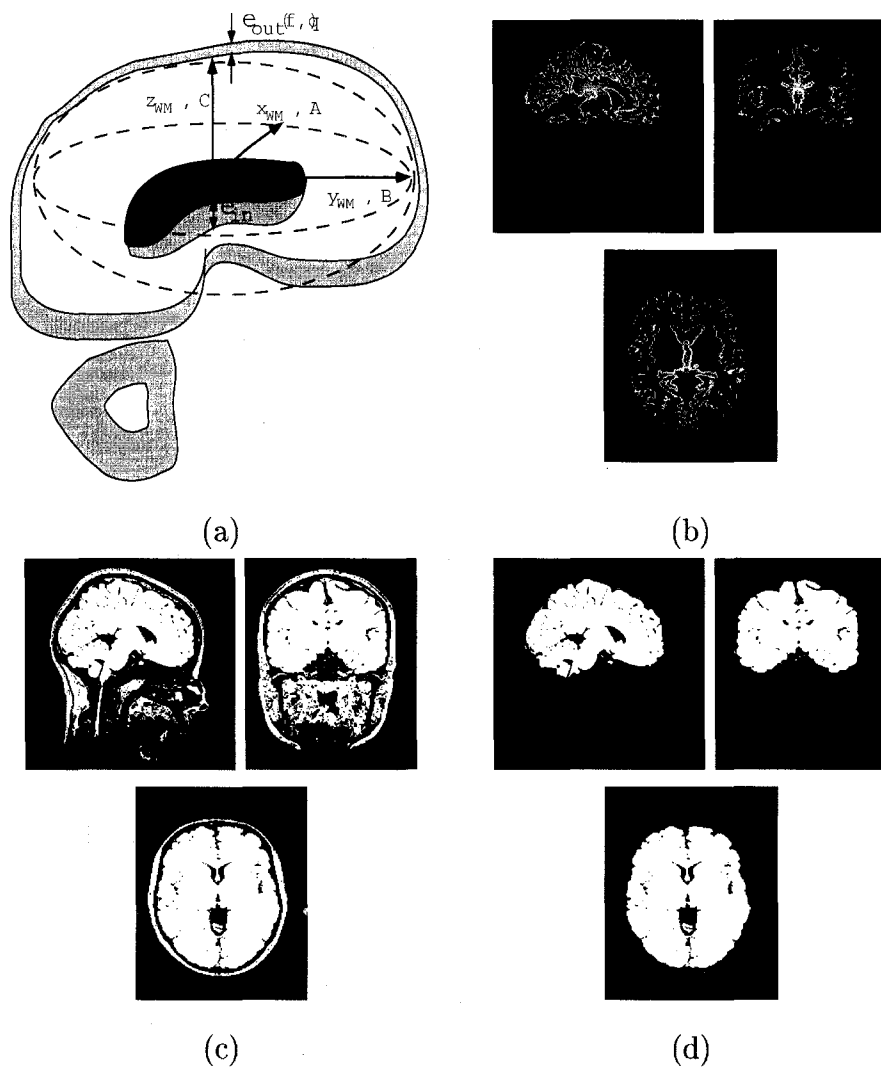


Figure 6.4 : *WM/GM Distance Function: (a) illustration of parameters used in computation, featuring ellipsoid roughly equated with WM data and used for normalizing $\epsilon_{WM/GM}$, as well as inner and outer thresholds, overlaid on simplified sagittal plane of anatomy; (b) WDF mapped to a grey-level intensity with WM contours overlaid in white; (c) speed function based on membership of voxels classified as WM or GM, using T1 alone without a contrast agent (note that there is little discrimination from T1 alone between GM and muscle, skin or sagittal sinus voxels); (d) speed function based on WM/GM voxels, as in (c), but excluding GM voxels which exceed the WDF shown in (b).*

- It is somewhat looser at the base of the brain where the grey matter tends to be thicker than in the cortex.
- It is more permissive still at the cerebellum, where the grey matter is thickest.
- It is loosest at the centroid of white matter, where bleeding of the surface model is less a concern, and which coincides with grey matter in the area of basal ganglia, neighbouring the ventricles [55].

In particular, this function linearly decreases with the distance from the WM centroid to the voxel of interest. For the sake of computing this proximity function, I assume that the brain can be idealized as an ellipsoid, and in turn this function can be normalized by an *ellipsoidal radius* ρ along the ray from the WM centroid to the voxel of interest. The semi-axes of this ellipsoid can be robustly estimated.

Using spherical coordinates $(\theta(\mathbf{x}), \phi(\mathbf{x}), \rho(\mathbf{x}))$, where $\mathbf{x} = (x, y, z)$, we have the following linear expression for the WDF:

$$\begin{aligned} \varepsilon_{WM/GM} &= a(\phi, \theta)\rho + b \text{ where} \\ a(\phi, \theta) &= [\varepsilon_{out}(\phi, \theta) - \varepsilon_{in}] / \rho_{max}(\phi, \theta) , \quad b = \varepsilon_{in} \text{ and} \\ \rho_{max}(\phi, \theta) &= \sqrt{[A^2 B^2 C^2] / [B^2 C^2 \sin^2 \phi \cos^2 \theta + A^2 B^2 \sin^2 \phi \sin^2 \theta + A^2 C^2 \cos^2 \phi]} . \end{aligned} \tag{6.6}$$

When $\rho = 0$, we have $\varepsilon_{WM/GM} = \varepsilon_{in}$, and as $\rho = 0$ approaches ρ_{max} , $\varepsilon_{WM/GM}$ tends to $\varepsilon_{out}(\phi, \theta)$. The expression is clamped for values of $\rho > \rho_{max}$, so that $\varepsilon_{WM/GM}$ is limited by ε_{in} and $\varepsilon_{out}(\phi, \theta)$. The expression for ρ_{max} is that of the radius of an ellipsoid, of semi-axes A , B and C , which can be derived from the expressions for x , y and z in spherical coordinates. The minimum threshold $\varepsilon_{out}(\phi, \theta)$ is set to one of 3 values, depending on where in the brain the voxel is located (with respect to the white matter centroid). This relative position has a relevance to anatomy, due to the consistent alignment of patients within the scanner: the upper brain corresponds to $\{\phi \leq \pi/2\}$, the frontal lower brain to $\{\phi > \pi/2, 0.0 < \theta \leq \pi\}$, and the cerebellum to $\{\phi > \pi/2, -\pi < \theta \leq 0.0\}$. The value of ε_{in} is constant and sets a high tolerance near the ventricles. Furthermore, if one simply partitioned the brain volume into 3 regions, based on the value of $\varepsilon_{out}(\phi, \theta)$, one would have a piecewise function which

would cause image artefacts at the boundaries of these regions. For this reason, there are small transition zones where the WDF is a weighted sum of the two neighbouring functions, based on local weights which are linear in ϕ or in θ . Finally, the semi-axes A , B and C of the ellipsoid which are implicit in the ellipsoidal radius $\rho_{max}(\phi, \theta)$ are obtained by computing the first and second moments of WM voxels with respect to x , y and z and by equating each second moment to the corresponding moment of inertia of an ellipsoid about a given axis. For example, the moment of an ellipsoid about the axis coinciding with semi-axis C is given by $1/5M(A^2 + B^2)$ [371], where M is the number of WM voxels circumscribed by the WM/GM boundary. Three such equations determine the three unknowns A , B , and C .

It is worth noting that the WM boundary, to which proximity of GM voxels is established, corresponds to the surface identified by first running the surface model only on WM voxels, as opposed to the boundary of all WM positives determined by the classifier. The latter approach would leave the surface model susceptible to isolated points misclassified as WM (false WM positives), which would drastically reduce the effectiveness of the coupling, particularly if only one modality were used. False WM-class voxels can occur near the eyeballs, for example. The 2-stage procedure presented here ensures that only voxels classified as WM which are contiguous with user-selected WM points are retained to anchor the identification of the outer cerebral and cerebellar surfaces. A distance map is then computed from the resulting inner cerebral and cerebellar surfaces (based on the method discussed in § 6.4.1), and the WM/GM Distance Function described above is used to restrict the expansion of the zero-level surface through nearby GM voxels, starting from the isosurface coinciding with the WM/GM boundary.

6.3 Semi-automatic Voxel Classification

As alluded to in the previous section and shown in figure 6.5, a classifier maps an image voxel, represented for D modalities as a D -dimensional feature vector in $\mathbb{R}^D$, $\mathbf{x} = (x_1, x_2, \dots, x_D)$, onto a tissue class C_j , where $j = 1, \dots, c$, and c is the assumed

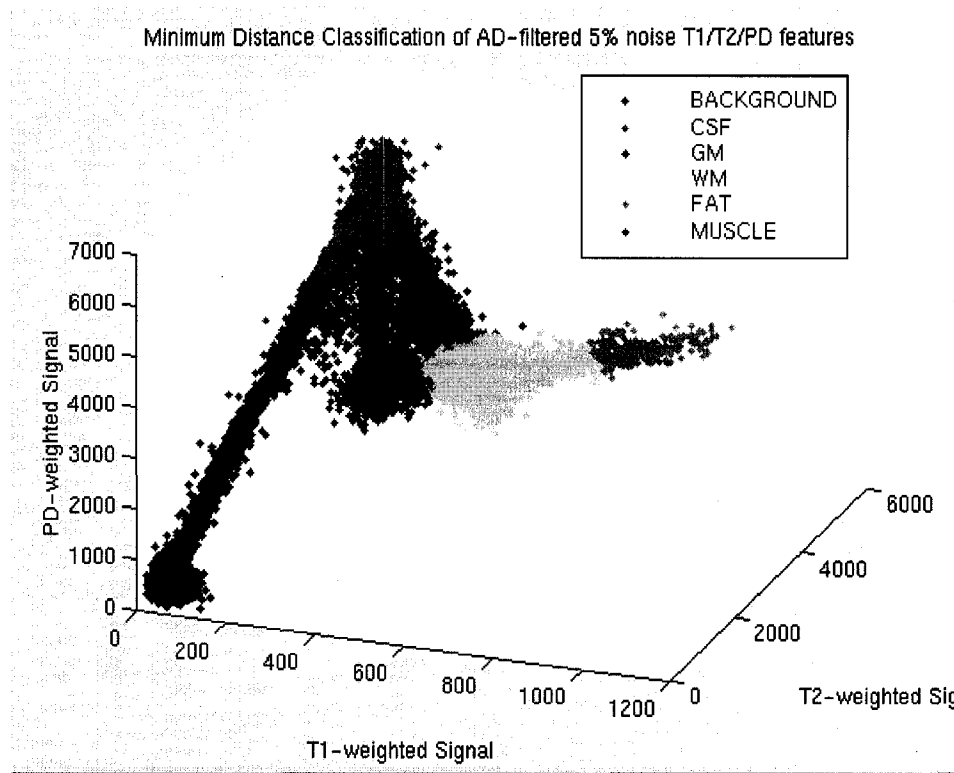


Figure 6.5 : *Illustration of voxel classification, for T1-, T2- and PD-Weighted feature space, based on the middle coronal slice of simulated data with 5% noise (see discussion in § 7.1), preprocessed with anisotropic diffusion ($t = 0.5\text{sec}$).*

total number of tissue classes. My motivation here is to select a method that can exploit multiple modalities, which does not degrade too badly if only one (typically T1) modality is available (which is often the case in a clinical setting, although the scanned tissues may feature contrast agents), and which is able to deal with the presence of pathologies.

A classifier can be either *supervised* or *unsupervised*. Supervised classification techniques [196, 117] include *Minimum Distance* (MD), *Bayesian*, *k Nearest-Neighbour* (kNN) and *Artificial Neural Network* (ANN) classifiers, all of which categorize points in feature space according to some optimality condition, such as minimal Euclidean or covariance-weighted (Mahalanobis) distance to class centroids, given a feature space which has been preprocessed on the basis of training points of known class. Some of

these methods, such as ANN, require extensive training, typically hundreds of points, to be usable. Unsupervised classification, or clustering, iteratively assigns a membership value u_{ik} to each feature point $\mathbf{x}_k$ with respect to each class C_i . This process groups points into clusters in feature space, measures the quality of these clusters in terms of an objective function, and seeks to optimize this objective function over successive iterations, generally without any prior knowledge of feature space. This membership can be either *discrete*, $u_{ik} \in \{0, 1\}$, as in the case of the *Hard C-Means* classifier, or a *continuous* function, $u_{ik} \in [0, 1]$, as in the case of the *Fuzzy C-Means* technique.

In the INSECT³ multiple sclerosis study conducted at the MNI [430], the training phase of the classification methods is expressed as a set of expertly labelled tag points, typically numbering in the hundreds and given in a standard anatomically-relevant coordinate space named *Talairach space* [386, 73]. These tag points can then be warped to any individual’s brain scan and used to train the classifier applied to that scan [196]. However, lesions occurring in an IGNS setting can appear anywhere in the brain and be arbitrarily large, making a standardized set of tag points erroneous, and making Talairach-to-native space warping unfeasible. Therefore, a semi-automatic classification scheme used in IGNS precludes such extensive supervision, since *one training stage is required for each classification stage*. Otherwise this semi-automatic approach would become nearly as laborious as a manual one. Nevertheless, some performance advantages of supervised over unsupervised classification algorithms [196] provide impetus for using at least some form of supervision.

Consequently, this thesis investigates a variation of Fuzzy C-Means (FCM) clustering, called the *semi-supervised FCM* (ssFCM) algorithm, which involves *limited supervision* (tested with as few as 3-5 points of each tissue type) and which in fact has been used to classify voxels of brain scans featuring pathologies [29]. I compare the ssFCM performance with the *Minimum Distance* (MD) classifier, which can function with a relatively small training set and constitutes the first iteration of the ssFCM classifier. Consideration of the two allows us to easily establish the benefit of

³Intensity Normalized Stereotaxic Environment for Classification of Tissues.

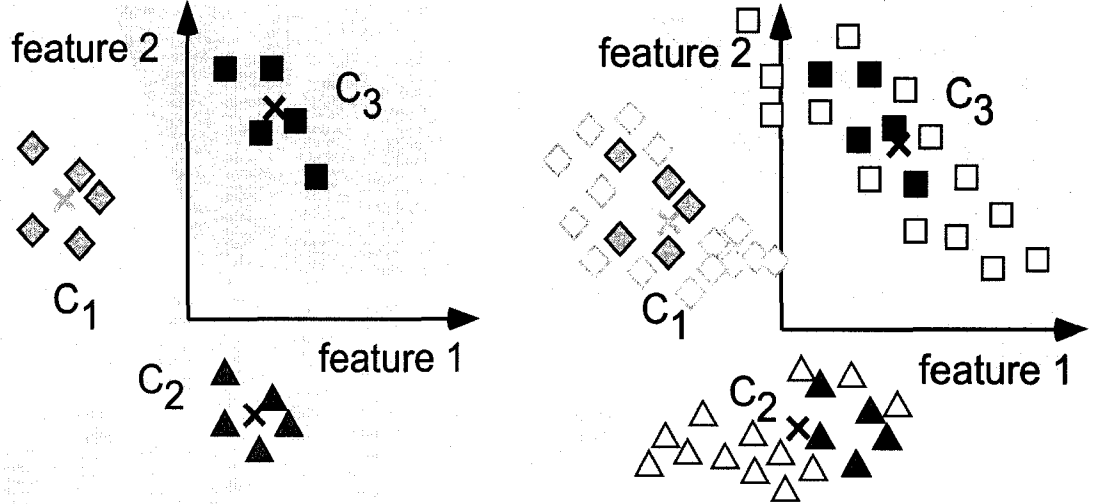


Figure 6.6 : *Simplified illustration of partitioning of parameter space by (a) the Minimum Distance classifier, where the centroids and feature space partition is computed from training points, and (b) the semi-supervised Fuzzy-C Means classifier, which iteratively refines the MD-partitioned feature space on the basis of both training and putatively classified data points. Training points are shown in solid colours, data points are shown as white-centered shapes and centroids of each class are marked with an X.*

clustering after the initial partition of parameter space based on Euclidian distance to the centroid. As shown in figure 6.6, the MD classifier is characterized by a partitioning of parameter space on the basis of distance to the centroid of the training points of each class. I use a Euclidean metric, normalized according to the range of values in the T1, T2 and PD volumes. Other distance metrics, such as the Mahalanobis distance [97], could also have been used. The ssFCM classifier refines this partitioning of parameter space by adapting it to clusters of points amongst *all*, not just training, data.

Formally, the ssFCM classifier involves, for all points $\mathbf{x}_k$ in parameter space, computing a set of fuzzy membership values $\mathbf{U}_t \equiv \{U_{i,t}(\mathbf{x}_k)\}$ to the corresponding classes

$\{C_i\}$ at iteration t , based on the position of each class centroid $\mathbf{v}_i$.

$$U_{i,t}(\mathbf{x}_k) = \left[\sum_{j=1}^c \left(\frac{\|\mathbf{x}_k - \mathbf{v}_{i,t-1}\|_A^2}{\|\mathbf{x}_k - \mathbf{v}_{j,t-1}\|_A^2} \right)^{2/(m-1)} \right]^{-1}, \quad (6.7)$$

where $\|\mathbf{x}_k - \mathbf{v}_{i,t-1}\|_A^2$ represents the norm of a vector weighted by a positive semi-definite matrix $\mathbf{A}$ which in this application normalizes the metric for the different ranges of T1, T2 and PD values. The parameter $m \in [0, \infty)$ is a weighting exponent which controls the abruptness of the fuzzy class partition. At $t = 0$, the centroids are computed from training points $\mathbf{t}_k$ only. Subsequent centroid computations take into account all points, with training points weighted strongly [29]:

$$\mathbf{v}_{i,t} = \left(\frac{\sum_{k=1}^{n_t} w_k (U_{i,t}(\mathbf{t}_k))^m \mathbf{t}_k + \sum_{k=1}^{n_x} (U_{i,t}(\mathbf{x}_k))^m \mathbf{x}_k}{\sum_{k=1}^{n_t} w_k (U_{i,t}(\mathbf{t}_k))^m + \sum_{k=1}^{n_x} (U_{i,t}(\mathbf{x}_k))^m} \right) \quad \text{where} \quad (6.8)$$

$n = n_t + n_x$. The algorithm iterates, recomputing the class memberships of each voxel, as well as the class centroids, until there is no appreciable change in the class partition function $U_t(\mathbf{x}_k)$, or until a fixed number of iterations is reached. After the last ssFCM iteration, the results can be *hardened*, whereby each voxel is finally assigned to the class coinciding with its highest-ranking fuzzy membership value.

6.4 Numerical Implementation of the Surface Model

6.4.1 Computing Signed Distance with the Fast Marching Technique

As previously discussed, the embedding function Ψ is initialized as a signed distance function from the user-defined initial $\Psi = 0$ isosurface⁴. This section describes the use of the Fast Marching (FM) technique to compute the signed distance to the surface and closest surface point to each voxel. It should be emphasized here that the Fast Marching technique is one of many which could have been employed for distance

⁴While it is not absolutely necessary to initialize Ψ with a signed distance map, the initial function must be Lipschitz continuous, and the use of a signed distance map is common in the literature, because of its convenience.

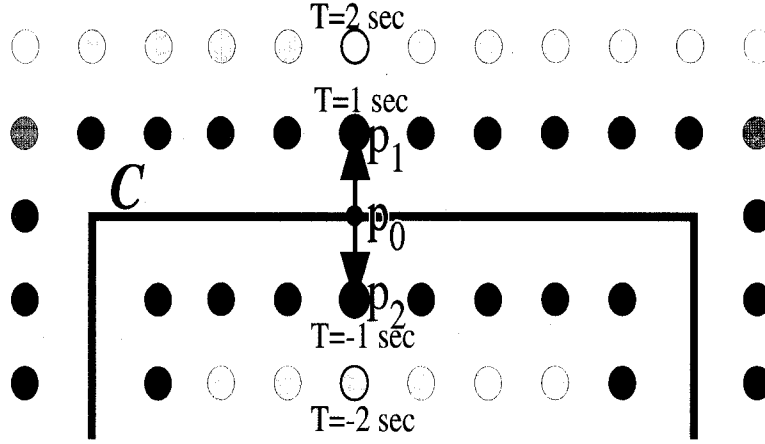


Figure 6.7 : *Computation of Distance and Closest Point Map for 2D contour C , featuring the “arrival time” of the contour (whose value is symbolized by a shade of blue: the darker the blue, the lesser the arrival time): p_0 is the closest contour point for p_1 and p_2 .*

computation and surface label propagation. This thesis research adopts a propagative distance computation technique (see also [320, 80]) because such a technique can easily be modified to propagate the label of the closest surface point, in contrast with a method based on a transform operator [40, 299]. The computation of distance from a surface can be idealized, as shown for a simple 2D contour illustration in figure 6.7, as equivalent to determining the “arrival time” of an outward moving surface (or contour), assuming it has unit speed. In other words, if we describe the initial surface (contour) as a front moving with unit speed, and picture individuals with a stopwatch at each voxel (pixel) in the volume (image), those who are at a point traversed by the front at time $t = 1$ are in effect 1 pixel away from the original surface (contour). Those which see the front at $t = 2$ are 2 pixels away, and so on. Similarly, we can describe the distance to inside points in terms of the arrival time of an inward moving front originating at the initial surface.

The notion that the arrival time of a surface within a volume is related to its speed, provided that its motion is monotonic, is the basis for Fast Marching (FM) method. The FM method has been used to model the motion of implicitly defined interfaces, in a manner similar to surface evolution methods presented so far, with the

added condition that this motion must be monotonic (e.g.: there can be no diffusive smoothing). The motion of the surface or contour can be expressed as the *Eikonal equation*⁵, which has a long history in the geometric optics and classical mechanics literature:

$$\|\nabla T\|F = 1. \quad (6.9)$$

Here, T is the arrival time of the surface, and this expression states in effect that the gradient of this arrival time is inversely proportional to the speed of the surface's evolution. This expression leads to a highly efficient and numerically stable method for estimating the monotonic motion of a surface, and assuming uniform unit speed, for numerically estimating the distance map to a surface. The method uses *upwind differencing* [331] to ensure numerical stability. This approach can be likened to a front being swept outward, whose motion estimation is based on points where the front has been, and whose motion has already been computed, rather than on points where the front is going. The improvement made by the Fast Marching method is to introduce order in the selection of the grid points, with a heap-sort [340], ensuring this upwind computation and starting at the original surface at $T = 0$. Details of this procedure appear in [343], pp. 86-100. I assign a negative sign to interior points, to produce the signed distance map required to initialize Ψ .

The distance and closest point computation is exploited on several levels (see illustrations drawn for 2D contour analogy on figure 6.8):

- The computation of the *sign of a voxel*, with regard to the initial ($t = 0$) closed surface, based on the sign of the dot-product two vectors: 1. the vector between the voxel and its closest surface point, and 2. the orientation of the normal of the original surface, which is given analytically. For $t > 0$, the sign is obtained by a heuristic more robust than normal computation.
- Implementation of the *Narrow Band*, within which computation of the evolution

⁵In 1D, this equation is derived in [343]: considering $T(x)$ the arrival time and using *distance = rate $\times$ time*, we have $1 = F \frac{dT}{dx}$. In multiple dimensions, ∇T is orthogonal to the level sets of T , and its magnitude is inversely proportional to speed.

equation is restricted, near the evolving surface at $\{\Psi(t) = 0\}$. This is a simple distance threshold beyond which no point is processed in the FM heap-sort.

- *Image extension*: image information only coincides with the $\{\Psi(t) = 0\}$ isosurface, and must be extended elsewhere in the domain of $\Psi(t)$, within the narrow band. If a point $\mathbf{p}_0$ on the zero-level isosurface coincides with an anatomical boundary, this should not only prevent $\Psi(\mathbf{p}_0)$ from changing, but freeze points on other isosurfaces $\{\Psi(t) = c, \text{ where } c \neq 0\}$, whose closest point on $\{\Psi(t) = 0\}$ is $\mathbf{p}_0$.
- Determining the *proximity of GM voxels to WM*. This is a simple Fast Marching distance computation from the final WM/GM boundary. The resulting distance from the WM/GM boundary is compared against the WM/GM Distance Function shown in figure 6.4 (a) in order to decide whether the speed function is 1.0 or 0.0.
- Finally, *the Euclidean distance and closest surface point computation is exploited to accelerate the registration stage later on*, as we compute a Closest Point Map from the final identified cortex surface, everywhere in the scan volume⁶. As is emphasized in chapters 3 and 4, the surface registration technique makes use of closest-point pairs to (over-)determine an iterative sequence of transformations between MRI-derived and range-sensed cortical surfaces.

6.4.2 Numerical Implementation of Surface Evolution Equation

Once again, the surface evolution model is evaluated with the following expression:

$$\begin{aligned} \Psi(\mathbf{i}, (l+1)\Delta t) = & \Psi(\mathbf{i}, l\Delta t) + \\ & \Delta t \left\{ F(\check{\mathbf{i}}) \left\| (\Psi_x(\mathbf{i}), \Psi_y(\mathbf{i}), \Psi_z(\mathbf{i})) \right\| [H(\Psi(\mathbf{i})) + \nu] + \right. \\ & \left. (F_x(\check{\mathbf{i}}), F_y(\check{\mathbf{i}}), F_z(\check{\mathbf{i}})) \cdot (\Psi_x(\mathbf{i}), \Psi_y(\mathbf{i}), \Psi_z(\mathbf{i})) \right\} . \end{aligned} \quad (6.10)$$

⁶If a tie exists for two closest surface point to a voxel, this voxel is labelled as such, which modulates its use in the subsequent registration.

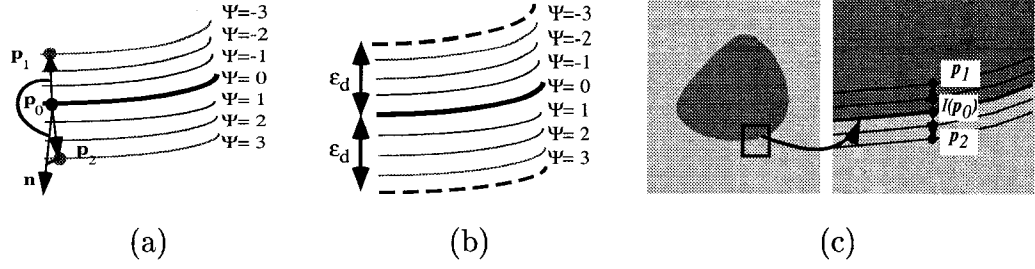


Figure 6.8 : *Applications of Distance and Closest Point Map, for 2D contour example: (a) the sign computed for the signed distance map is determined by the sign of the dot-product between the normal and the vector from the closest surface point to the current point of interest; (b) the restriction of the computation of the evolution equation to a narrow band close to the zero-level also exploits the Fast Marching distance computation; (c) the extension of image information from a contour point to other points closest to it: the contour point $\mathbf{p}_0$, being on an anatomical boundary, stabilizes the value of $\Psi(\mathbf{p}_1)$ and $\Psi(\mathbf{p}_2)$.*

The image term F is evaluated at the closest $\{\Psi(t_l) = 0\}$ isosurface point $\tilde{\mathbf{i}} = (\tilde{i}, \tilde{j}, \tilde{k})$, since the image information only coincides with the zero-level. This is better visualized when considering the analogous 2D problem, where the image plane on which the halting force is based stops the computation when the $\{\Psi = 0\}$ isocontour coincides with the anatomical 2D boundary of interest. Both the diffusive term $\|(\Psi_x, \Psi_y, \Psi_z)\| H$ and the doublet term $(F_x, F_y, F_z) \cdot (\Psi_x, \Psi_y, \Psi_z)$, where

$$H = \frac{[\Psi_{xx}(\Psi_y + \Psi_z) + \Psi_{yy}(\Psi_x + \Psi_z) + \Psi_{zz}(\Psi_x + \Psi_y) - 2\Psi_x\Psi_y\Psi_{xy} - 2\Psi_x\Psi_z\Psi_{xz} - 2\Psi_y\Psi_z\Psi_{yz}]}{[\Psi_x^2 + \Psi_y^2 + \Psi_z^2]^{3/2}}, \quad (6.11)$$

are based on derivatives evaluated with *central differences* [206], e.g.:

$$\Psi_x(i, j, k) = [\Psi(i+1, j, k) - \Psi(i-1, j, k)] / [2h]$$

$$\Psi_{xx}(i, j, k) = [\Psi(i+1, j, k) - 2\Psi(i, j, k) + \Psi(i-1, j, k)] / [h^2], \text{ etc.} \quad (6.12)$$

Once again, the hyperbolic term is computed with upwind differences, which for an outward moving front gives:

$$\|(\Psi_x, \Psi_y, \Psi_z)\| \nu = \nu [(max(D_x^- \Psi, 0))^2 + (min(D_x^+ \Psi, 0))^2 + (max(D_y^- \Psi, 0))^2 + (min(D_y^+ \Psi, 0))^2 + (max(D_z^- \Psi, 0))^2 + (min(D_z^+ \Psi, 0))^2], \quad (6.13)$$

where, e.g.: $D_x^- \Psi = [\Psi(i, j, k) - \Psi(i - 1, j, k)]/h$ and is known as the *backward* difference, $D_x^+ \Psi = [\Psi(i + 1, j, k) - \Psi(i, j, k)]/h$ and is referred to as the *forward* difference. Finally, the size of the time step is also restricted, for the sake of stability, by the Courant-Friedrichs-Levy condition: $\Delta t/h \leq 1$ [218, 343], where h is the smallest inter-voxel spacing. The reader should note that first-order differencing is employed here for the simplicity of its implementation, but higher-order discretizations of the model, in both time and space, are also possible, and are characterized by good computational stability (see [343, 218] for further details).

6.5 Conclusions

This chapter provided implementation details about the surface evolution model proposed in this thesis. In particular, the model implemented here uses an image-based speed function which is determined not by gradient information, but by the results of a voxel classification technique. The use of a classifier allows the surface model to proceed in two stages, one which identifies the outer white matter surface, and one which flexibly uses this surface to anchor the final brain surface identification, to prevent unwanted bleeding effects. The flexible coupling is implemented with the WM/GM Distance Function, where the value which maintains the final brain surface close to the outer WM surface is a function which varies according to spherical coordinates. A classifier-based image speed function also provides a seamless mechanism for exploiting multispectral data, that is any combination of T1-, T2- and PD-weighted data.

The choice of the classifier is a trade-off between the requirement of some supervision for such a critical application, and the realization that the presence of pathology makes exhaustive supervision, particularly based on a tag file warped to the individual data set, virtually inapplicable. In short, a classifier which uses only limited supervision, and which also performs reasonably well with T1-weighted data alone, is sought, leading to a comparative study of the Minimum Distance and semi-supervised Fuzzy C-Means methods. Finally, some details of the numerical implementations were

also featured, namely the Fast Marching method used to compute a signed distance function, to manage the narrow band and to identify the closest surface point in the domain of interest, as well as the numerical differencing which implements the surface evolution equation in space and simulated time.

Chapter 7

Model-Based Surface Identification in MRI: Validation

7.1 Validation Tools: MRI Simulator and Realistic Digital Brain Phantom

This chapter presents the validation of the various stages of the brain surface identification technique presented so far with a quantitative approach based on realistic synthetic data, where ground truth is known, and qualitative tests with *in vivo* subject data. The latter data come in the form of healthy subject scans obtained from the ICBM¹ database, as well as MNI patient scans featuring pathologies. While patient data can be exploited quantitatively by comparing image processing results against manual analyses performed by experts, the latter process is known to be labour-intensive and suffers from both intra- and inter-expert variabilities [200, 430]. In contrast, the quantitative validation stage presented in this chapter makes use of two software tools which were developed at the MNI and which obviate these variabilities: a MRI simulator specifically designed and implemented by Kwan et al. for the purpose of image analysis evaluation [200, 199], and a high-resolution, volumetric, anthropomorphic digital brain phantom elaborated by Collins et al. [74].

¹International Consortium for Brain Mapping [400].

The design of the digital brain phantom is motivated by the limitations of *in vivo* studies, namely the implausibility of establishing with certainty the exact boundaries of tissues in a tomographic scan (without slicing and staining the underlying anatomy), and those of previously existing phantoms, both physical and digital, whose geometrical simplicity has limited their predictive ability and anatomical relevance [74]. This phantom is based on a single-subject average dataset, compiled from 27 scans [155], which was semi-automatically classified and subsequently manually corrected by a trained neuroanatomist. Each voxel is labelled as one of 10 classes, namely²: *Background, Grey Matter, White Matter, Cortico-Spinal Fluid, Fat, Muscle, Skin, Skull, Glial Matter, and Other*. The phantom comes in two forms, respectively conveying *discrete* and *fuzzy* tissue membership. While the latter is judged to represent partial volume effects most realistically, I have chosen to work with the discrete version because of the requirement of precise boundaries in characterizing the accuracy of the various stages of the surface model.

Finally, the healthy-subject digital phantom is the starting point for a second phantom featuring pathological tissue, which based on the manual re-labelling of some grey and white matter voxels. These *Pathology* class voxels are given relaxation time values which are culled from the review paper of Bottomley et al. [43]. Bottomley describes T_1 and T_2 values for specific pathologies present in various tissues, including brain tissue. A relationship of the type

$$T_1 = Af^B, \quad (7.1)$$

between T_1 and NMR frequency f is characterized, along with average T_2 values (which are not seen as frequency-dependent) based on a fitting of data points taken from an exhaustive literature survey, for miscellaneous tumours, astrocytomas, meningiomas, gliomas and edemas.

²However, the T_1 , T_2 and PD values of the latter 4 classes mirror those of 4 of the first 6, which for classification studies is equivalent to a labelling according to 6 tissue classes. The realistic assignment of magnetization properties of the latter 4 classes was not judged to be essential to the applicability of the phantom [75]. For this reason, all illustrations of the phantom in this thesis feature 6 classes, distinguished on the basis of magnetization properties.

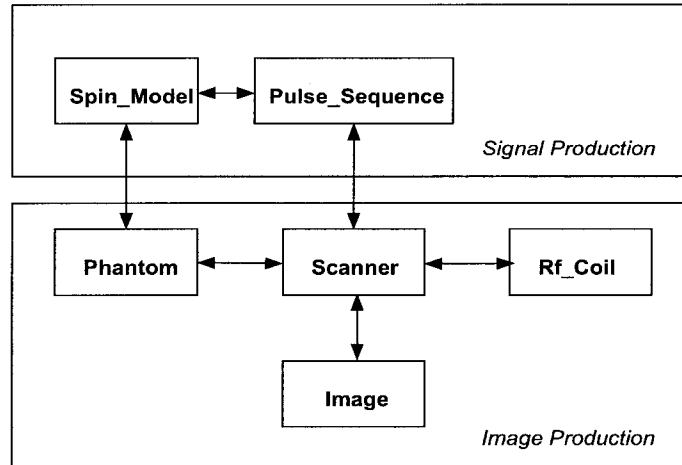


Figure 7.1 : *An overview of the simulation process (courtesy of G. Bruce Pike [198]).*

The MRI simulator features *Signal* and *Image Production* modules, the former of which is comprised of *Spin Model* and *Pulse Sequence* objects, while the latter is comprised of *Phantom*, *Scanner*, *Rf Coil* and *Image* objects [199]. The architecture of the system, emphasizing the interaction between objects, appears in figure 7.1, while an illustration of the 3D brain phantom and typical T1-, T2- and PD-weighted images resulting from it appear in figure 7.2. The Signal model is dedicated to computing tissue intensities on the basis of the magnetization properties, made explicit in the Spin Model, and of the signal resulting from Pulse Sequence object. The Image Production module is controlled by the Scanner object and generates tomographic images from computed tissue intensities, and features imaging effects such as noise, RF inhomogeneity, and partial volume. The Phantom object synthesizes the patient anatomy in the scanner as described above, the RF Coil object models signal reception, and finally the resulting tomographic scan is stored in the Image object.

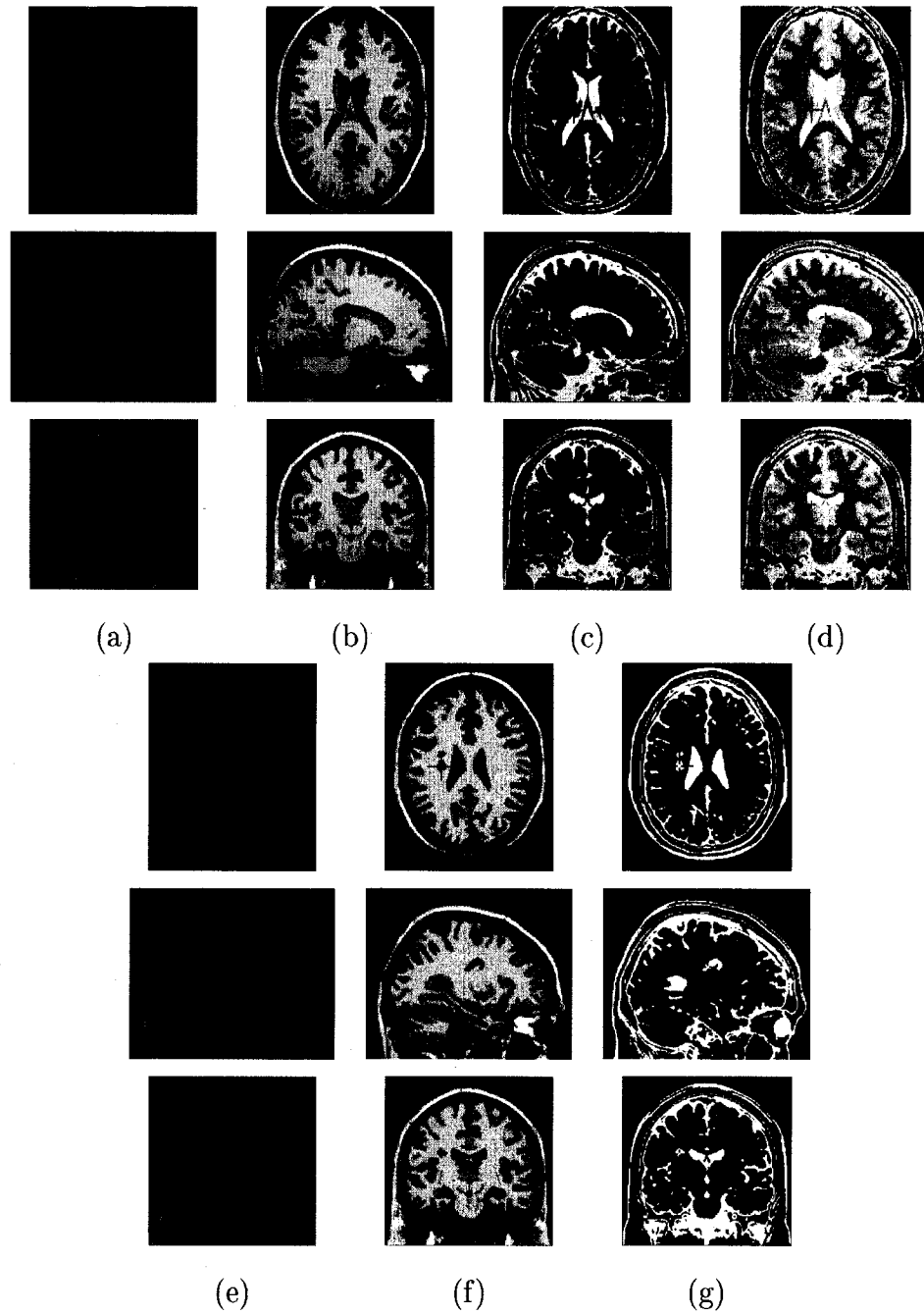


Figure 7.2 : (a)-(d) 3D digital healthy brain phantom and simulated MR data (noise 3%, inhomogeneity 0%): (a) phantom, (b) T1-, (c) T2- and (d) PD-weighted volumes. (e)-(g) Analogous phantom as in (a), with synthetic astrocytoma inserted: (e) phantom, (f) T1- and (g) T2-weighted volumes, based on data in Bottomley [43]. No PD information on brain pathologies was found in the literature.

7.2 Anisotropic Diffusion Validation

The validation of the anisotropic diffusion preprocessing comprises a comparison of the Kimia technique [189] adopted here,

$$\begin{aligned}\psi_t &= H \|\nabla \psi\| \text{ where} \\ H &= [\psi_{xx}(\psi_y + \psi_z) + \psi_{yy}(\psi_x + \psi_z) + \psi_{zz}(\psi_x + \psi_y) \\ &\quad - 2\psi_x\psi_y\psi_{xy} - 2\psi_x\psi_z\psi_{xz} - 2\psi_y\psi_z\psi_{yz}] / [\psi_x^2 + \psi_y^2 + \psi_z^2]^{3/2},\end{aligned}\quad (7.2)$$

with Gaussian filtering, as well as a series of tests to determine an optimal scale factor under typical imaging conditions. In contrast with the surface model application, the function $\psi(x, y, z)$ used here is not initialized with a signed distance map, but with the raw image data. The smoothed output is simply the value of ψ at the simulated time duration, based on a finite differences discretization of expression (5.20). In other words, *the same evolution framework serves both to filter the raw data and identify the brain surface*. In the former case, only the diffusive term is enabled and the input is the raw image data. In the latter case, all four now-familiar terms of expression 6.2 are used, and the input is a user-defined surface, along with image gradient or tissue class data. The comparison presented here exploits Koenderink's Diffusion Equation formulation for Gaussian filtering [195],

$$\psi_t = \nabla^2 \psi, \quad (7.3)$$

which allows us to easily look at the two formulations at identical scales, characterized by the time duration of each diffusion. Once again, the relationship between the spatial extent of the convolution implementation and the time duration of the diffusion implementation of the Gaussian filter is given by [195, 165]:

$$2\sigma^2 = 4t. \quad (7.4)$$

Figure 7.3 provides a qualitative illustration of the difference between anisotropic and Gaussian diffusion, over a range of scales. The quantitative comparison between the two types of filtering consists of two tests:

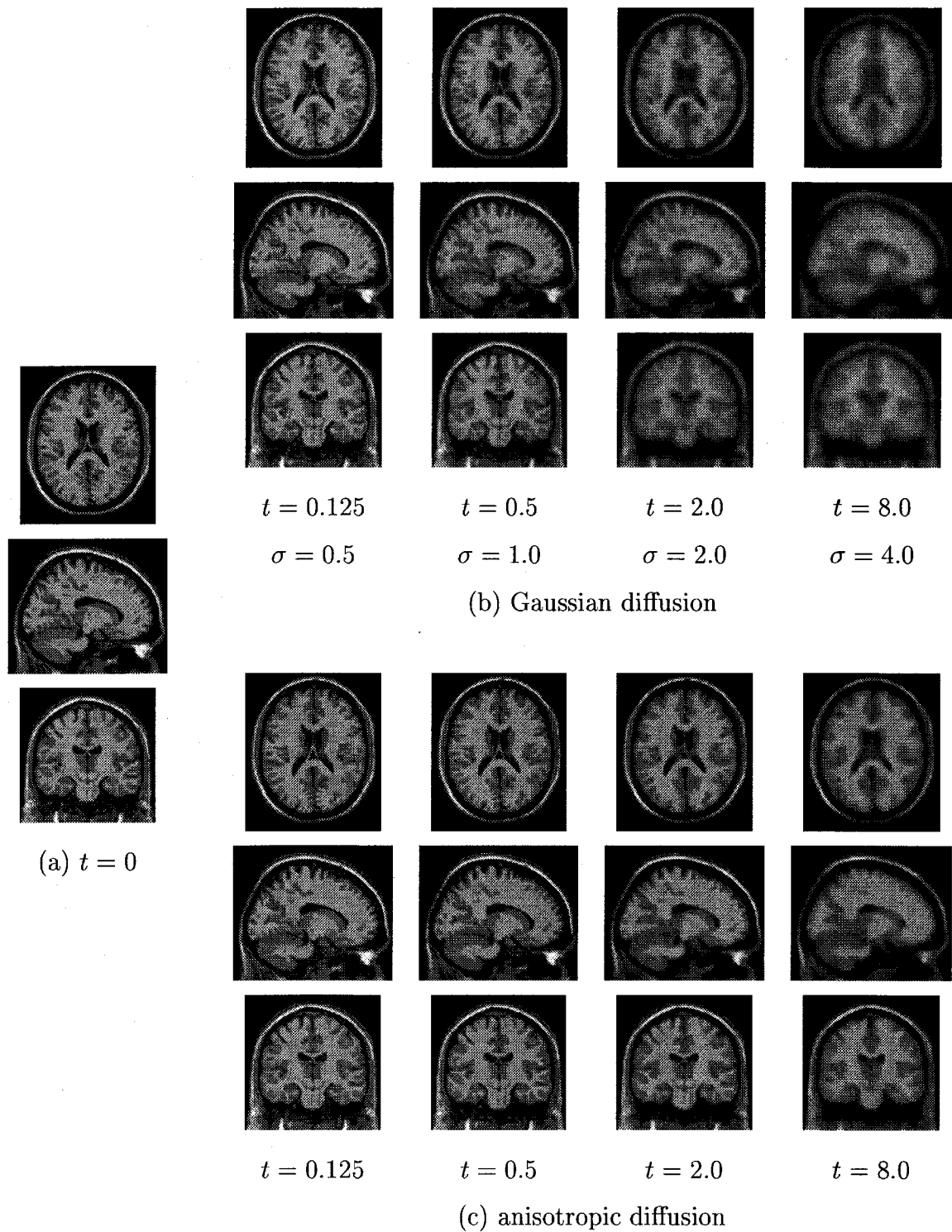


Figure 7.3 : *Illustration of the effect of the scale parameter, for (a) simulated T1-weighted volume (with null noise and inhomogeneity values) on (b) Gaussian and (c) anisotropic diffusion.*

- First, an *evaluation of the strength and localization of the edges*, which is ultimately an indication of the position of the brain boundary ultimately determined by the surface model detected in simulated T1-weighted volumes (under various noise and RF inhomogeneity conditions). The ground truth in these tests is the known position of the brain boundary in the underlying digital phantom.
- Second, an *evaluation of the effect of pre-filtering on voxel classification accuracy*, based on various combinations of T1, T2 and PD features (again over a range of noise conditions), given the known class membership of the voxel of the underlying digital phantom.

The two comparison tests are performed over a range of scales. The edge test is an indication of how much blurring takes place across anatomical boundaries and ultimately how accurately the image force stops the moving zero-level isosurface. The classification test is an indication of the fidelity of the preprocessed noisy T1, T2 and PD data to the noiseless case, both at, and away from, anatomical boundaries, and an indication of implications of this fidelity for the subsequent voxel classification. A detailed discussion of the classification pre-filtering test is deferred to the validation of the voxel classifier, appearing in § 7.3.

The edge test uses as a starting point the position of all tissue boundaries of the digital phantom, assumed to be ground truth, and computes three measurements from the filtered T1 volume, under different noise, inhomogeneity and diffusion scale characteristics, in the immediate neighbourhood of this boundary point, as illustrated in figure 7.4:

- The average value of the *maximal gradient amplitude*, in this immediate neighbourhood in the T1 volume of known boundary points. The value has no absolute significance, but provides a way to compare the stopping power of the image force, particularly if the magnitude of the gradient is considered in its expression. The larger the magnitude of the gradient, the greater the stopping power of the model.

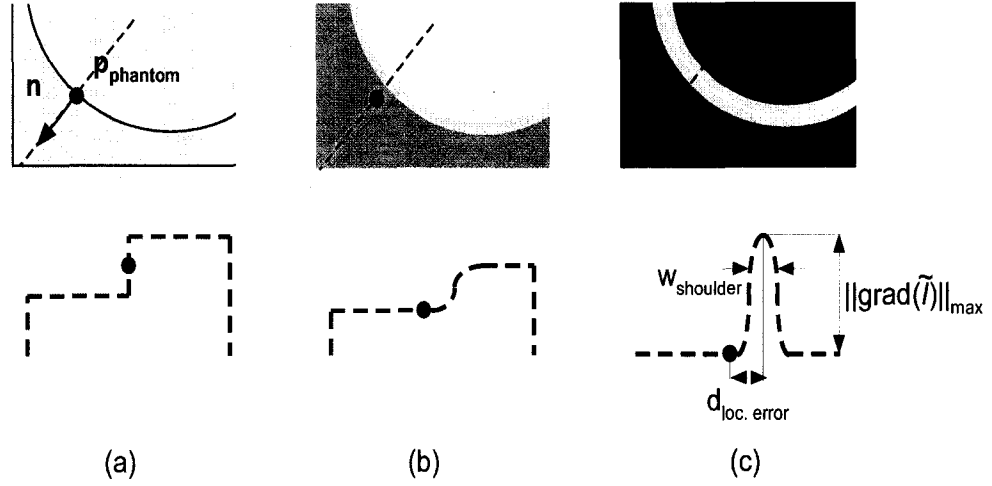


Figure 7.4 : *2D Illustration of measurements for validation and choice of scale of anisotropic diffusion: (a) phantom class data, with 1D profile along normal direction, through known boundary point; (b) filtered simulated T1-weighted MR data, with 1D profile as in (a); (c) magnitude of the gradient of filtered T1 data, with 1D profile as (a) and (b), with measurements shown.*

- The average *localization error* between the maximum gradient point, detected in the neighbourhood of the known boundary point along the direction normal to this boundary, and the position of the boundary point (as determined on the digital phantom). Were the boundary a step edge along one axis, and perfectly planar along the other axes, then there would be no reason to believe such a localization error would occur, because the Gaussian convolution kernel is radially symmetric³.

³Nalwa and Binford [278] have shown that edge detection with Gaussian pre-filtering is unbiased for an ideal step edge, but for a generalized step edge of height S featuring different slopes k_1 and k_2 on opposite sides of the transition. Moreover, curvature-dependent anisotropic diffusion would not modify a noiseless straight boundary. However, when dealing a curved boundary, where the convolution is not necessarily symmetric, one expects an offset between the known boundary and the detected edge point.

- The average *width of the shoulders* of the gradient magnitude profile, along the direction normal to the boundary, is an indication of the uncertainty of the stopping area of the surface model. Ideally, this stopping area should be relatively compact, and close to the local maximum of the gradient magnitude which corresponds to a boundary point. This width is defined as the distance between the two shoulders, or voxels whose gradient magnitude is 80% or less of the local gradient maximum, in the directions along and opposite the boundary normal. From the gradient maximum, the algorithm takes unit steps in the normal and anti-normal directions, consider the closest neighbouring voxel, and stops when it finds the shoulder. The vectors from each boundary point to both its shoulders are projected onto the normal direction, and the total length of these projections is the measurement which is averaged for all known boundary points.

Measurements statistics for validation and choice of scale of anisotropic diffusion, under different noise conditions, are compiled under 0% RF inhomogeneity conditions. The tests were repeated with 20% RF inhomogeneity, and resulting values for average localization error and shoulder width were affected less than 0.5%; the average gradient magnitude had variations on the order of 1%, but these variations were found to be consistent for the two types of diffusion and therefore of judged to be of little interest. It is assumed that that in vivo scans can be compensated accurately for RF inhomogeneity by the nonparametric nonuniformity correction method described in [360].

The reader will have noticed that noise values are presented in terms of a percentage rather than a SNR value, in accordance with the simulator noise model used by Kwan [199]. The noise is specified therein as a percentage noise standard deviation σ over WM average signal intensity A : $\frac{\sigma}{100A}$, for each quadrature component. Were we dealing with pure Gaussian-distributed noise added to the magnitude signal, the SNR would be the inverse of the noise value $\times 100$: $\frac{A}{\sigma}$; i.e.: 1% noise equivalent to SNR=100, 5% noise equivalent to SNR=20, and so on. However, the distribution is Rician, and converting from % to SNR requires a compensation for this distribution

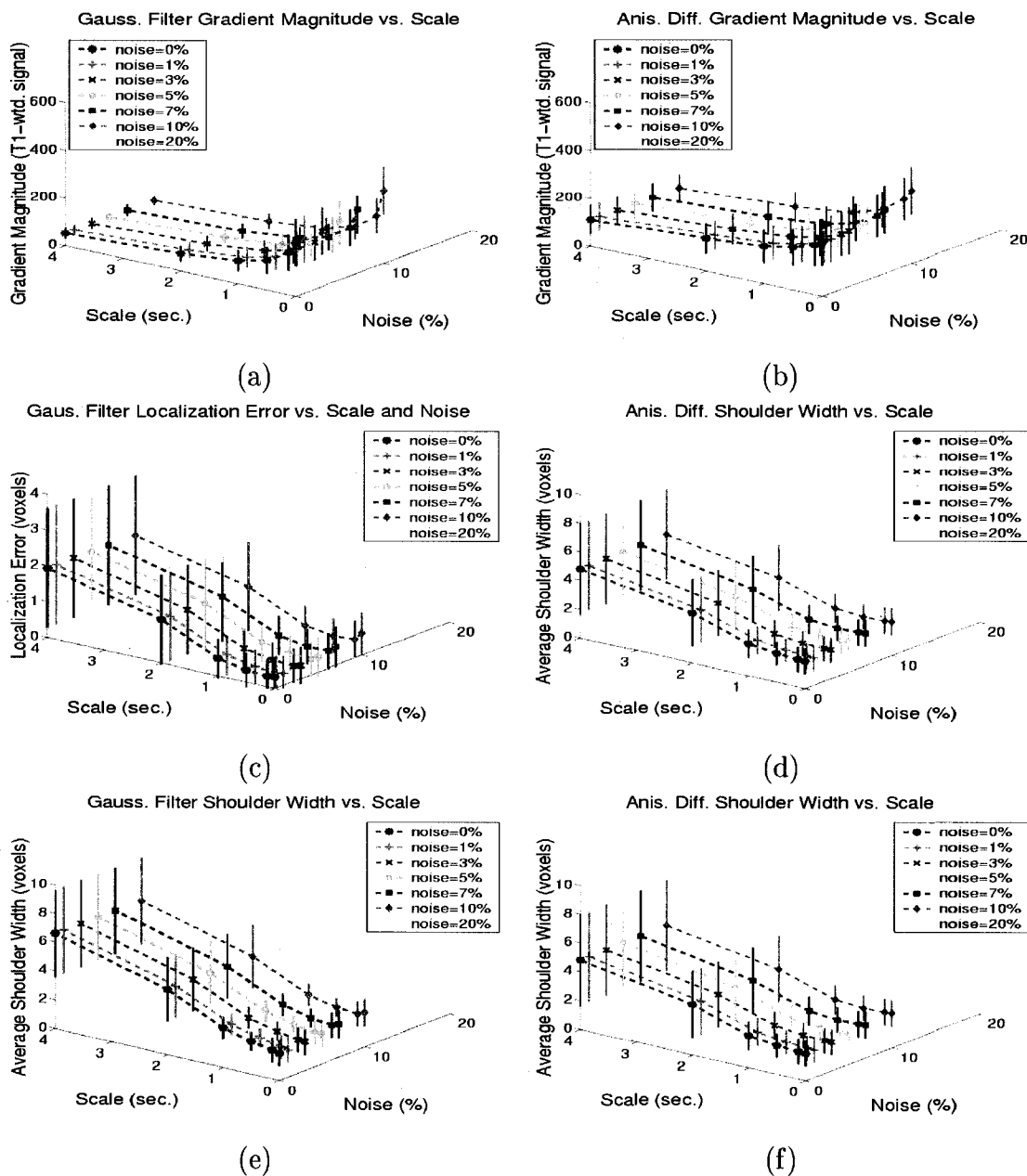


Figure 7.5 : *Measurement statistics for validation and choice of scale of anisotropic diffusion, under different noise and 0% RF inhomogeneity conditions. Average and standard deviation of: 1. local maximum gradient magnitude under (a) Gaussian filtering (GF) and (b) anisotropic diffusion (AD); 2. localization error of local gradient maximum from known boundary position under (c) GF and (d) AD; 3. shoulder (80% of gradient magnitude maximum) width value along the boundary under (e) GF and (f) AD.*

(see [199] pp.73-95). A 3% noise level is seen as representative of a typical scan.

Comparison of figures 7.5 (a) and (b) indicates that the magnitude of the local gradient maximum is considerably more stable over scale for anisotropic diffusion than Gaussian filtering, for all noise conditions of 0 – 10%. The 0-scale intercept of the noiseless curve is seen as an ideal here, since there is no noise to corrupt the computation of derivatives. Under heavier noise conditions, local gradient maximum values differ significantly from the 0-scale noiseless value: at low scales in particular, the gradient magnitude is much higher, due to the effect of noise spikes on derivative computation. At high scales, the average local gradient maximum tends to a value lower than the unfiltered, noiseless value: consideration of a progressively larger volume in the filtering, particularly in the isotropic (Gaussian) case, causes the underlying step edge in T1 values to be drowned out.

The performance of the two types of diffusion with regards to gradient maximum localization error is illustrated in figures 7.5 (c) and (d), and the comparison of the shoulder-to-shoulder width statistics appears in (e) and (f). The gradient maximum localization error statistics for the two filters are highly comparable at all scales, virtually over all noise levels. Only at medium to high scales is there a slight advantage to anisotropic diffusion, although the Gaussian is more stable in the presence of very high noise amplitudes. A comparison of the uncertainty of the edge position, as measured by shoulder-to-shoulder width, favours the anisotropic diffusion virtually at all scales, at low to medium noise levels, as expected.

7.3 Validation of Semi-Automatic Voxel Classification

The objectives of this validation study are to determine the feasibility of using either the semi-supervised Fuzzy C-Means (ssFCM) or Minimum Distance (MD) algorithm to map a T1 value or a feature vector, T1-T2 or T1-T2-PD, to the correct tissue class C_i (which is known a priori for this study, by virtue of the digital phantom), and

in particular to choose one method over the other, based on the available modalities and on assumptions about noise (I assume once again that RF inhomogeneity can be corrected as in [360]). The effect of noise and of the choice of filtering method (including a choice of scale), on classifier performance is also featured here. The results of the validation study of classification appear in figures 7.6 to 7.9.

Figure 7.6 provides a comparison of the two filtering options as they relate to the Minimum Distance voxel classification. This study uses Minimum Distance rather than semi-supervised Fuzzy C-Means because of the unpredictable effect of clustering on the partitioning of parameter space, given that I want to isolate the effect of noise distribution and of pre-filtering method and scale. It is clear that under severe noise, namely at 20% and higher, Gaussian diffusion outperforms anisotropic diffusion. However, I would not expect to use a classifier with any success under those conditions. Under less severe conditions, anisotropic diffusion is comparable at very small scales and is otherwise superior to Gaussian filtering, for 1 to 3 modalities. For both diffusion options, a scale value between $t = 0.125$ and $t = 0.5$ produces optimal results, for a noise value within 10%. Beyond that noise level, the classifier is generally unreliable.

An important observation about figure 7.6 is that for higher scales, $t \geq 2.0$, and at low noise values ranging from 0 to 5%, the classification doesn't appear to gain much from the added consideration of T2- and PD-weighted scans. This anomaly can be cleared up by considering erroneous voxels summed in accumulators at the pixels in one plane, say the sagittal plane, as shown in figure 7.7. In this case, the intensity of the pixel is proportional to the number of misclassified voxels at the same (y, z) coordinate, over all x values. It is clear that misclassified voxels in the multi-modality case are grouped in bands, coinciding with the borders of tissue classes, and that much of the misclassification in this case is related to border effects. It stands to reason that these border effects would worsen with increasing scale, which is seen on all combinations of modalities. Moreover, the consideration of more than one modality improves the classification of brain tissue voxels, particularly at finer scales. This is borne out by considering separate classification statistics for white

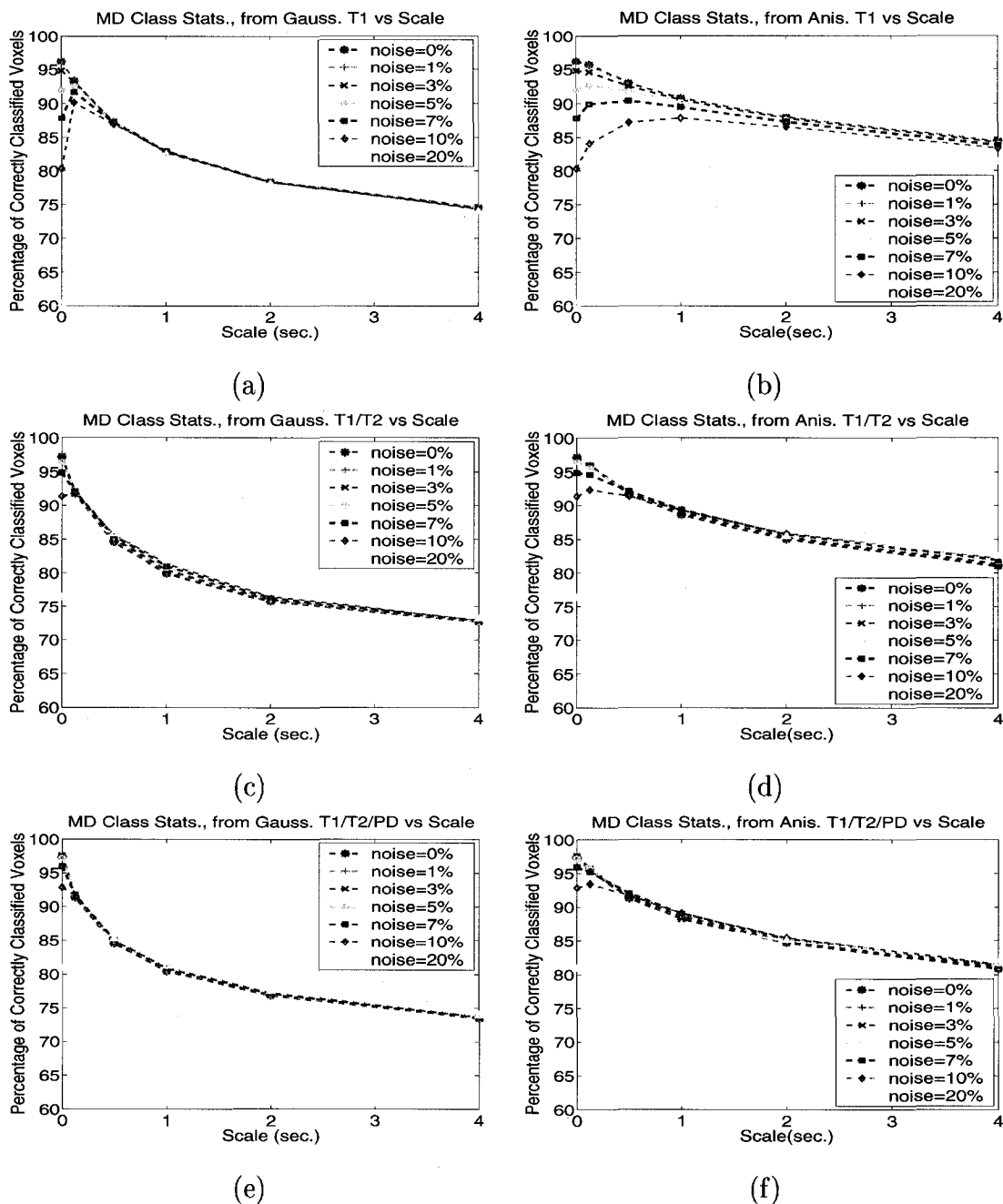


Figure 7.6 : *Illustration of sensitivity of MD classification to noise levels, to pre-processing method and scale, and to number of modalities: percentage of correct voxels, 1. based on T1-weighted scan alone, at various noise levels, under (a) Gaussian filtering (GF) and (b) anisotropic diffusion (AD); 2. based on T1- and T2-weighted scans, under (c) GF and (d) AD; 3. based on T1-, T2- and PD-weighted scans, under (e) GF and (f) AD.*

and grey matter voxels, which appear in figure 7.8. In particular, the use of T2 data to complement T1 data clearly improves the classification of GM voxels. The T1 value of the GREY MATTER class, at 900, allows very little separation from the MUSCLE/SKIN class, at 833, based on that modality alone. It is particularly when only T1 data are available that some form of coupling of the outer brain surface to the WM/GM boundary is necessary in the surface model.

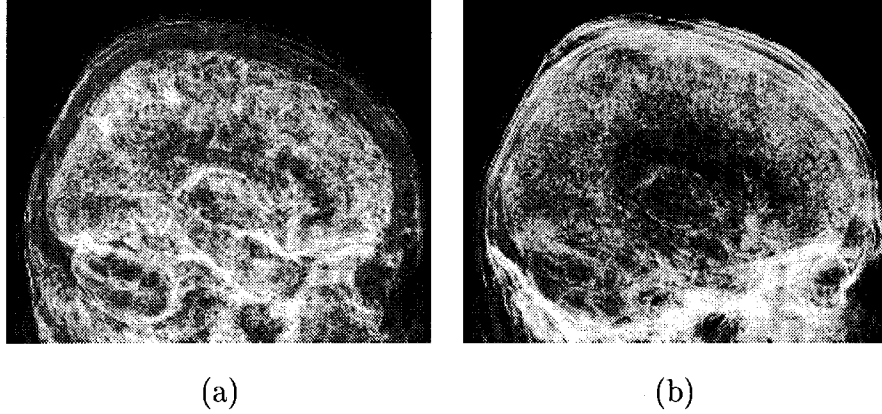


Figure 7.7 : *Spatial distribution of misclassified voxels, based on accumulators in the $y-z$ (sagittal) plane. The brighter the pixel, the larger the number of misclassified voxels having that x position. Results for (a) T1-weighted scan alone and (b) T1- and T2-weighted scans.*

Finally, figure 7.9 illustrate the fragility of the convergence properties of the semi-supervised Fuzzy C-Means algorithm, which in turn justifies the decision to ultimately opt for the Minimum Distance classifier. Figures (a), (b) and (c) illustrate the classification results at various noise levels and at various scales of anisotropic diffusion, respectively based on a T1-weighted scan alone, on a T1/T2 feature and on a T1/T2/PD feature. These graphs can be compared with figures 7.6 (b), (d) and (f) respectively. Only at very low scales and noise levels does ssFCM outperform MD, and then only slightly. The tuning of the ssFCM parameters (w_k and m in expression (6.8), as well as halting conditions) was done with simulated data of all three modalities under noiseless conditions, and with no pre-filtering. Each parameter was selected in the midst of an optimal band about which there seemed to be little variation in the classification results in the noiseless/filterless case. Figure (d) illustrates

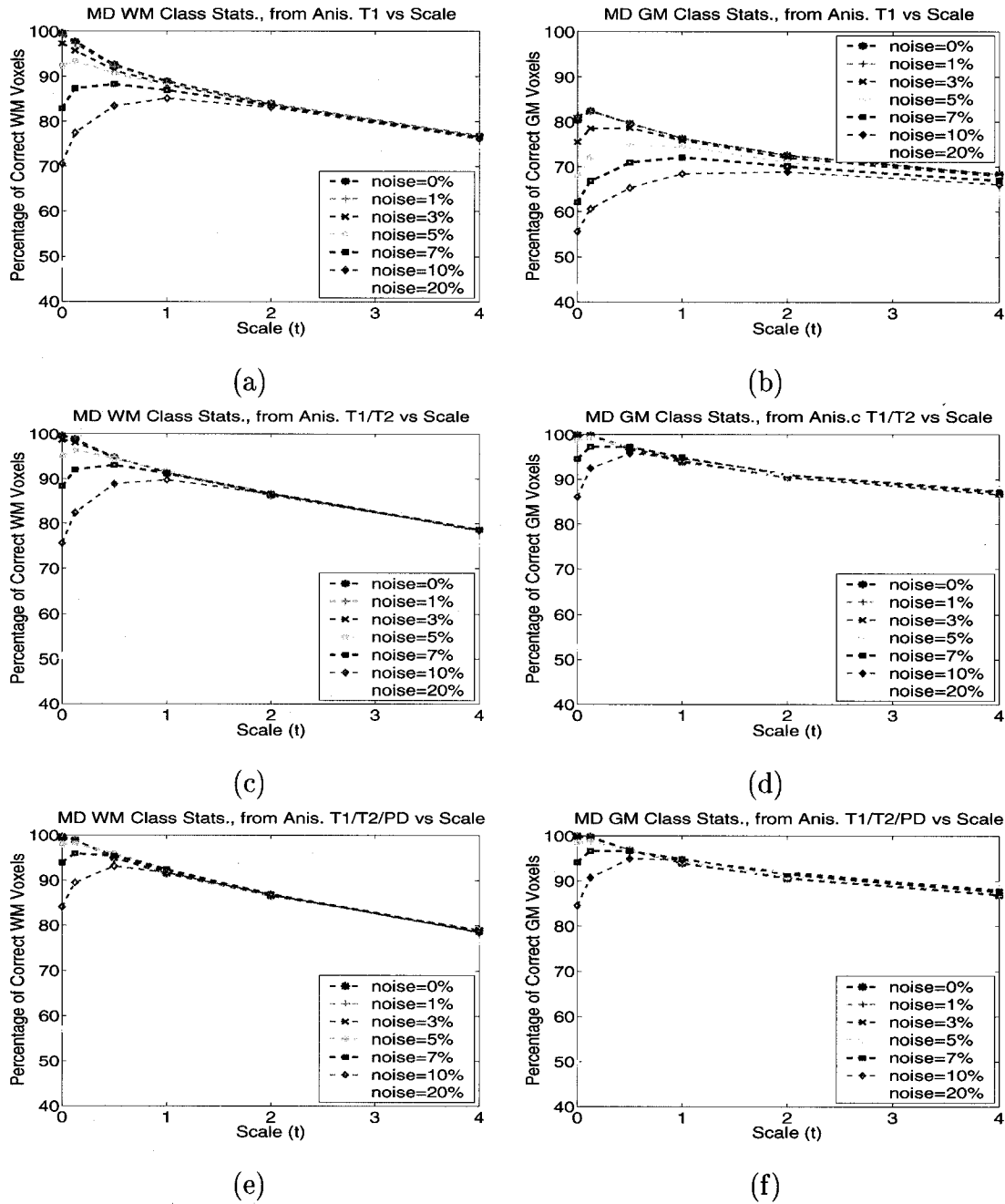


Figure 7.8 : Percentage of correct voxels of white and grey matter of MD classifier, at various noise levels, scales and combinations of modalities (using anisotropic diffusion): (a) WM and (b) GM results based on T1-weighted scan alone; (c) WM and (d) GM results based on T1- and T2-weighted scans; and (e) WM and (f) GM results based on T1-, T2- and PD-weighted scans.

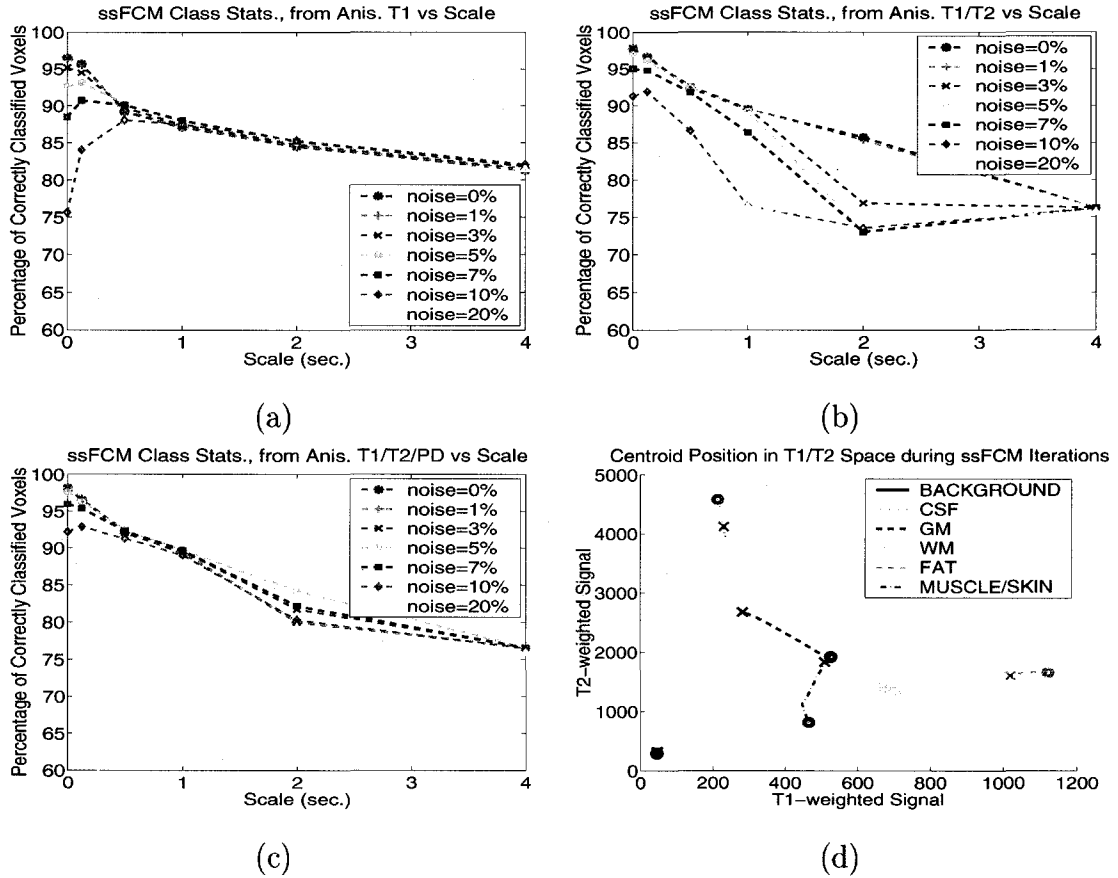


Figure 7.9 : *Illustration of sensitivity of FCM classification to noise levels, to preprocessing method and scale, and to number of modalities: percentage of correct voxels based on (a) T1-weighted scan alone, (b) on T1- and T2-weighted scans, and (c) on T1-, T2- and proton density-weighted scans, all after anisotropic diffusion; (d) illustration of migration of centroids over successive iterations caused by border effects at noise=5% and scale $t=2.0$ sec., where a "o" illustrates the position of the class centroid at the first iteration (coinciding with MD classification) and an "x" indicates the final (25th) iteration. Note the final position of MUSCLE/SKIN class centroid nearly coinciding with the initial GM centroid position.*

the ssFCM-related migration of class centroids, from points marked “O” to points marked “X”, based on T1- and T2-weighted data with 5% noise and a preprocessing stage of anisotropic diffusion with scale $t = 2.0$ (this coincides with the dip in graph (b) at this scale value). Border effects between CSF and grey matter, which worsen with increasing scale, can draw the GM class centroid towards the CSF area of parameter space, causing in turn a significant portion of GM voxels to be misclassified as WM or MUSCLE.

In summary, this validation study confirms the choice of anisotropic diffusion as optimal preprocessing, based on subsequent classifier performance. This study suggests a choice of scale between $t = 0.125$ and $t = 0.5$, which would correspond to the $\sigma = 0.5$ to $\sigma = 1.0$ range in a Gaussian framework. Finally, the fragility of semi-supervised Fuzzy C-Means technique, particularly in response to border effects, justifies the final decision to adopt the Minimum Distance classifier. Subsequent to this study, it has come to my attention that recent classification methods improve on a strictly feature-based approach by making use of higher-level information in the form of separate classes which are explicitly characterized by partial volume effects [66, 345], or a classification which exploits local neighbourhood information, through a Markov Random Field formalism [31, 219, 144] which expresses a preferred spatial configuration as a low Gibbs energy state⁴. Furthermore, work on classification methods which flexibly exploit global spatial information, such as contiguity with training points or prior assumptions about the scale of contiguous tissues of each type, along with feature space, is currently under way [13].

⁴Of course, brain tissue classification may also benefit from even more detailed high-level information, in the form of a co-registered atlas [112] or statistical average [102], but the presence of a large imbedded pathology may render such a method difficult to apply in this context: the warping problem is undermined by the uncertainty of the anatomical correspondences in the tissues in, and immediately surrounding, the pathology.

7.4 Fast Marching Distance Computation

Validation

The first stage of the validation is the comparison between the distance and closest point results of the Fast Marching (FM) technique with a gold standard obtained by a *brute force* (BF) *exhaustive search* for the discrete brain boundary point closest to each voxel. In other words, the BF technique finds the true closest surface point for a given voxel by explicitly determining the Euclidean distance between this voxel and each discrete brain boundary point. The BF search then associates with this voxel the distance to, and the label of, the true closest surface point, and this result is compared with the distance and closest point label obtained by the Fast Marching technique described in § 6.4.1⁵.

The brain boundary in question is obtained by first binarizing the digital phantom volume, by setting to 1 the grey, white and glial matter voxels and setting to 0 all other voxels, and then computing the gradient everywhere. Because the volume is binary, the gradient should be null everywhere except on the boundary. Moreover, I restrict my attention to the inner edge points (outside edge points also have non-zero gradients, due to the computation of derivatives with central differences). The choice of inner or outer boundary is unimportant, as this validation is only concerned with the discrepancy between the FM and BF techniques for an anatomically representative boundary.

The distance map for each technique is shown in figures 7.10 (a) and (b), in tri-planar grey level image format, where a maximum distance of 100 voxels in the volume maps to a grey-level of 192 (out of a possible 255), and where the boundary itself is overlaid in white. The difference between the two maps is shown qualitatively in tri-planar format in figure 7.10 (c), where 255 represents a maximum difference

⁵On a 200 MHz Pentium, the BF approach took more than 2 weeks to produce the equivalent map computed in less than 10 minutes by the FM method on a similar machine, for the $181 \times 217 \times 181$ scan volume comprising the digital phantom, although improvements to the exhaustive search have proposed, such as the k-D tree [424].

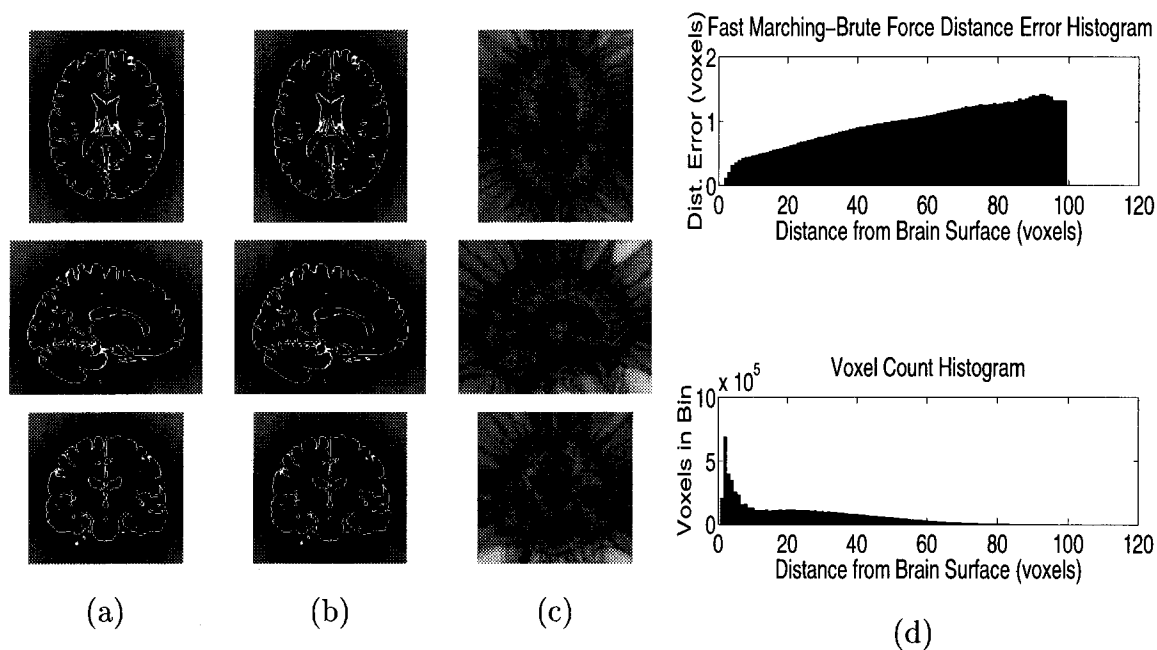


Figure 7.10 : (a) Distance Map to brain surface produced by Fast Marching (FM) technique, with distance values ranging in the volume from 0 to 100 pixels are mapped to grey-levels 0 to 192, and brain boundary appearing in white; (b) Distance Map to brain surface produced by Brute Force (BF) technique, with same distance to grey-level mapping as in a; (c) difference between FM and BF Distance Maps, shown as grey-level image: a voxel error ranging from 0.0 to 2.0 voxels maps to 0-255; (d) difference between FM and BF, averaged over voxels of similar distance to the brain surface: the average FM-BF difference value for each bin appears in the top histogram, whereas the number of points summed in each bin is shown in the lower histogram.

of 2.0 voxels. Moreover, the top histogram appearing in figure 7.10 (d) provides a quantitative illustration of the difference between the two techniques: the difference value at each voxel was binned according to the reference (Brute Force) distance to the brain surface (0.0-1.0 voxel, 1.0-2.0 voxels..., 99.0-100.0 voxels) and the average computed for each bin. The size of the summation is shown in the bottom histogram in figure 7.10 (d). As one would expect, particularly for a distance computation which is based on a propagative front emanating from the original surface, the error is cumulative over distance, and monotonically increasing (except for the very tail which factors in very few data points). It should be emphasized that the Fast Marching technique employed here is based on the consideration of the 6-neighbours of the central voxel under consideration in the minheap, and that better results are possible with the consideration of 26-neighbours, at the cost of considerably slower computation.

However, perhaps more relevant than the discrepancy in computed distances between the two methods is the comparison of the closest point labels found by each, which is illustrated by figure 7.11: the agreement between BF and FM in closest point labels is better than 99%. The distance value itself is essentially only used in managing the narrow band. In contrast, the label of the closest surface point is used for computing the sign of the signed distance function which initializes the surface model and for extending the image information away from the zero-level isosurface, as illustrated in figure 6.8. Also, these labels are used in the registration of the MRI brain surface with the range-based brain surface later on, through the Closest Point Map.

Moreover, it is geometrically feasible for a voxel to have 2 or more closest surface points, and this implementation of the Fast Marching technique has an explicit check for this case⁶, whereupon the segmentation uses the first label (one label is required everywhere in the narrow band for the image extension and for the sign computation)

⁶Recently, Siddiqi and his collaborators [351, 44] developed a technique for the simple and robust detection of shock points coinciding characterized by 2 or more closest contour/surface points, based on the application of the divergence theorem applied to a Hamiltonian system.

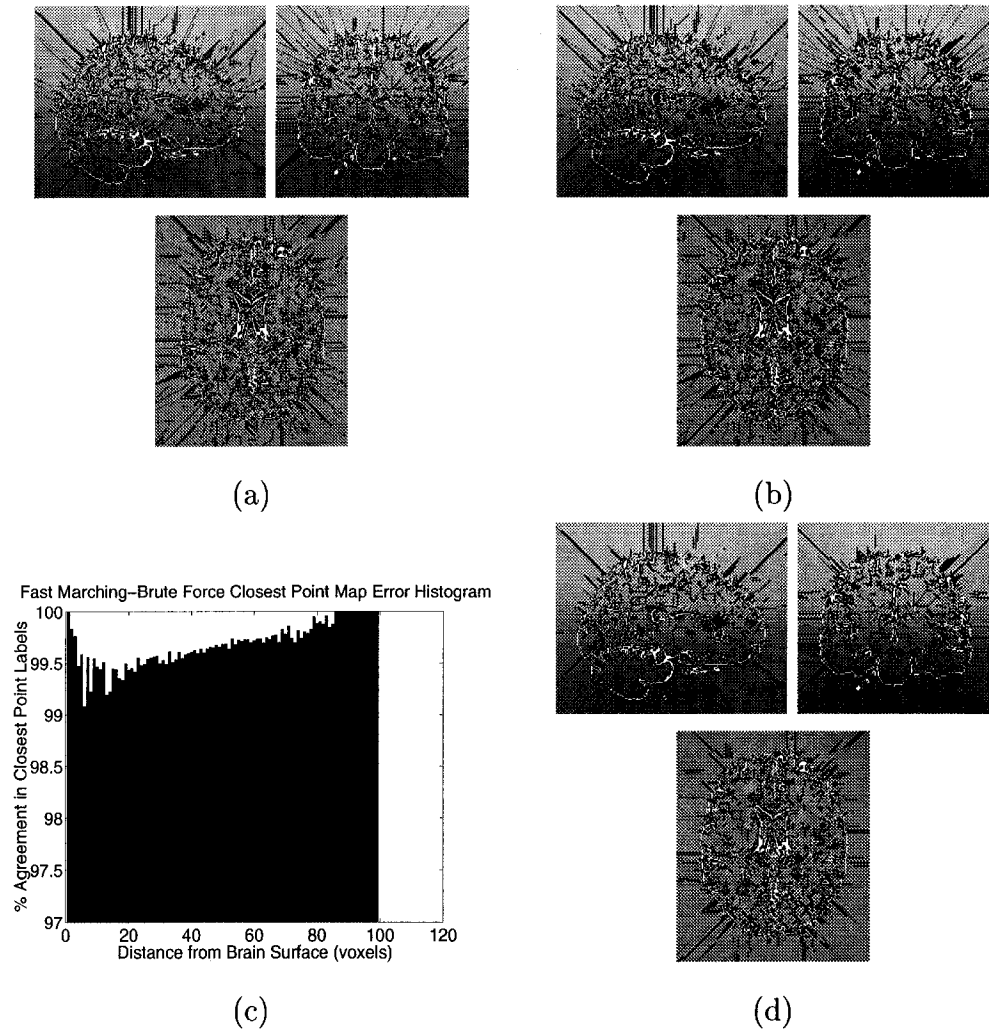


Figure 7.11 : *Closest Point Maps computed by 3D (a) FM and (b) BF techniques, for the brain boundary, labelled white, with surface labels 1-207459 mapped to grey levels 64-192, and with ambiguously labelled voxels shown as black (0); in (c), a bar graph illustrates the percentage of labels in agreement, as a function of distance to surface, based on the same voxel populations presented in figure 7.10 (d); (d) FM Closest Point Map with ambiguities due to digitization artefacts eliminated with a distance threshold of 2 voxels.*

but the registration can dispense with ambiguous voxel-surface point matches (see chapter 4). The voxels which possess two or more closest points are shown as black.

In theory (with an analytical surface, or with voxels of infinite resolution), voxels with two or more closest points should tend to the *medial surface* [351, 44] of the volume circumscribed by the brain boundary (white), but other streaks appear, due to the digitization of the boundary. For example, a flat oblique section of the boundary can lead to oblique streaks perpendicular to this section, where each streak precisely bisects the segment bounded by two boundary points obliquely adjacent to each other (this does not occur for locally flat horizontal or vertical patches, since no such bisection between contiguous voxels is possible). Voxels whose ambiguity is an artefact of the digitization process could be treated differently, in the subsequent registration, than those which truly possess 2 or more closest surface points. Such artefacts can be detected with a simple check on the distance between the positions of the 2 or more surface points matched with an ambiguous voxel.

7.5 Surface Model Validation

The object of the final validation of the model is to quantitatively estimate the accuracy of the outer brain surface identification with synthetic data and to qualitatively demonstrate its performance on real data, with both healthy ICBM subjects and MNI patient data involving pathologies. Validation with real data is conducted with constant parameters values throughout. Each optimal value of each parameter was established by trial and error with real subject data: the sensitivity of the model to parameter values is examined in the quantitative part of the validation.

7.5.1 Healthy Subject Validation: Synthetic Digital Phantom-based Data

The quantitative estimation uses as a reference the distance map computed from the known brain boundary by the brute force method discussed in the preceding

section. The surface model was applied to simulated MRI data featuring 3% noise, considered typical, and the distance of the final brain surface to the underlying brain boundary was ascertained everywhere and characterized by statistical measurements. The sensitivity of this performance to different noise levels was characterized. Also, the sensitivity with respect to each of the 4 parameters which determine the WM/GM Distance Threshold was ascertained, by varying each one about its “optimal” value. For each parameter or noise variation study illustrated here, the results for the usual parameter or noise value are indicated with an asterisk (*). The optimal threshold values were determined iteratively on the digital phantom data and on 8 ICBM data sets, for the T1 and T1/T2 cases. Each value selected was sufficiently high that no GM data were excluded in any of the data sets, but small enough to be useful for coupling to the WM surface.

The results in figure 7.13 indicate that the classification doesn’t gain much from the consideration of T2 data, and in fact the results actually get slightly worse. However, in the multimodal case, as can be seen, the classifier can be run effectively with 6 classes, allowing a better discrimination between muscle and GM tissues. Some error occurs in the sagittal and transverse sinus areas, which in the absence of contrast agents are easily misclassified as GM (despite the added consideration of T2 data). The further consideration of PD data reduces this misclassification. The reason that misclassified voxels can have a greater impact on the brain surface localization for multimodal data sets is that the thresholds used here are compatible with the sole use of T1 data, which leads to slightly inaccurate WM and GM classification, in particular in the cerebellar area and frontal lower brain, and which manifests itself as relatively large expanses of GM. The coupling function is made tolerant of this relatively thick GM layer by large thresholds in these areas, and these values turn out to be slightly too large for the more accurate multimodal voxel classification. However, the inclusion of the sinus areas within the brain surface should not have an adverse clinical impact, as the brain surface likely to be imaged by the range-sensor would also include the sagittal and transverse sinuses. The standard deviation of the distance-to-boundary data does seem to improve for the T1/T2/PD case, namely 0.65

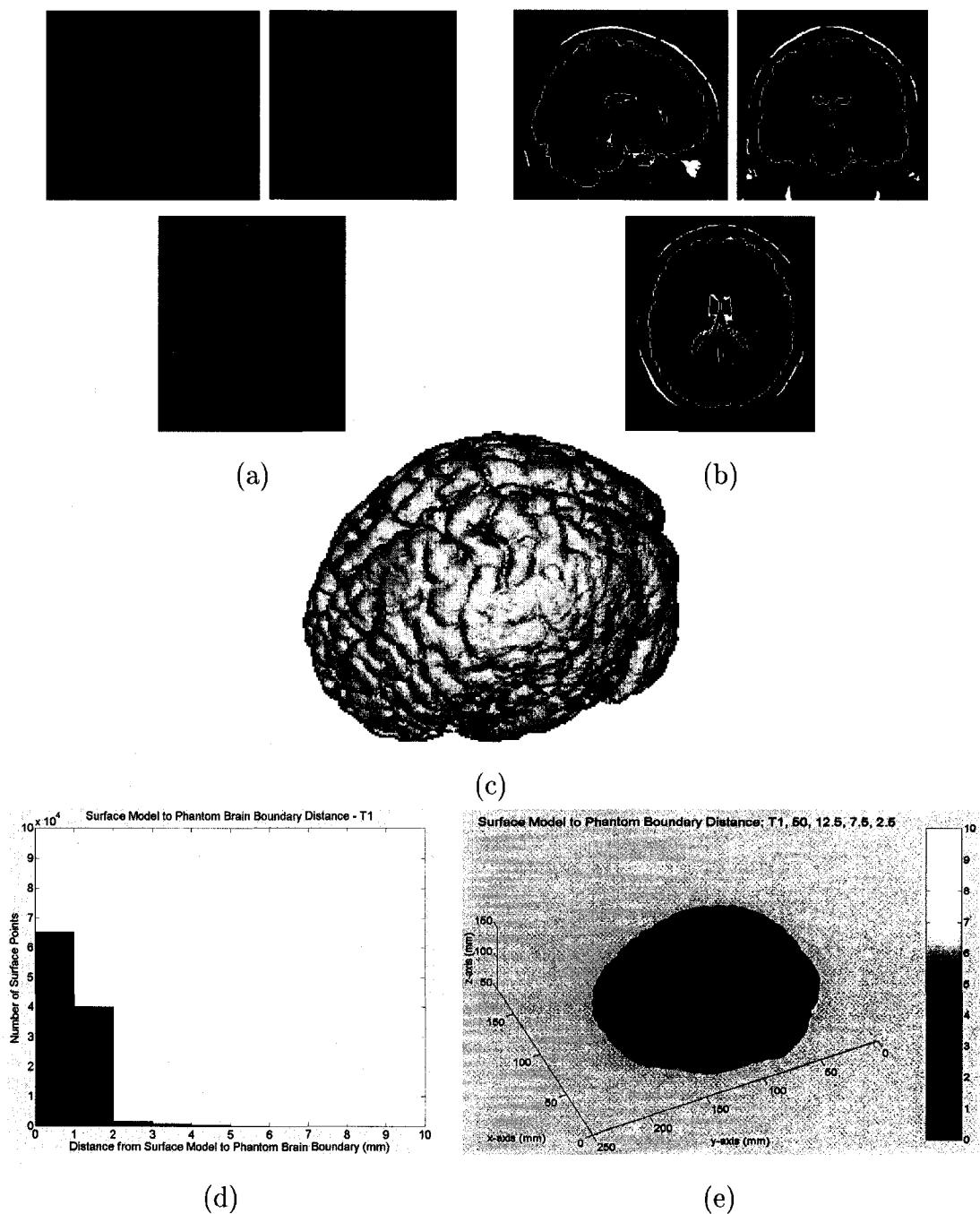


Figure 7.12 : (a) Tissue classes produced by MD classifier and (b)-(e) brain surface identified with surface model, from simulated T1-weighted scan: (b) tri-planar view; (c) rendered view; (d) distribution of distance between brain surface in scan and phantom brain boundary; (e) error between brain surface identified in scan and underlying brain boundary in phantom, shown as colour plot: $\text{avg.} = 0.48 \pm 0.71 \text{ mm}$.

mm instead of 0.71 mm for T1 alone, and the average distance for T1/T2/PD is also quite comparable to T1 alone. Improvements to the classification method, such as the consideration of spatial information, would probably allow the model to perform better with multimodal feature vectors. This is the thrust of ongoing research [13].

The next study illustrates the sensitivity to the four parameters which implement the WM/GM Distance Function. Recall expression (6.6), whereby $\varepsilon_{WM/GM} = \varepsilon_{in}$ at the WM centroid, and $\varepsilon_{WM/GM}$ tends to $\varepsilon_{out}(\phi, \theta)$ at the outer boundary of the brain. The four parameters which determine the coupling between WM and GM are the inner threshold ε_{in} and the three values for the outer threshold ε_{out} , coinciding with the outer cortex $\{\phi \leq \pi/2\}$, the cerebellum $\{\phi > \pi/2, -\pi < \theta \leq 0.0\}$ and the frontal lower brain $\{\phi > \pi/2, 0.0 < \theta \leq \pi\}$. The results in figures 7.15 to 7.16 indicate that the model is most sensitive to the frontal lower brain threshold, and relatively insensitive statistically to the other thresholds. This is hardly surprising, as the inner threshold variation incurs no further bleeding effect, because at the inner boundary lies CSF which the classifier will not interpret as WM or GM, and because the variation of the thresholds other than that of the frontal lower brain, results in bleeding which is restricted to the sagittal or transverse sinus. Variation of the frontal lower brain threshold, on the other hand, can result in substantial surface model bleeding into the eyeballs and muscle tissues, some of which have magnetization properties similar to WM and GM respectively.

Finally, the results in figure 7.14 indicates that the model is relatively insensitive to noise up to and including 10%, but in keeping with the classification statistics, cannot deal with a noise level of 20%. Ironically, the results seem to be slightly worse in the absence of noise, which is consistent with the slightly worse performance of the model with T1/T2 feature vectors. Basically, the classification improves enough in the cerebellum and frontal lower brain to make the usual thresholds slightly too loose, which results in the inclusion of sinus tissues in the final surface. As emphasized before, the inclusion of these tissues in the final surface should have no adverse clinical impact.

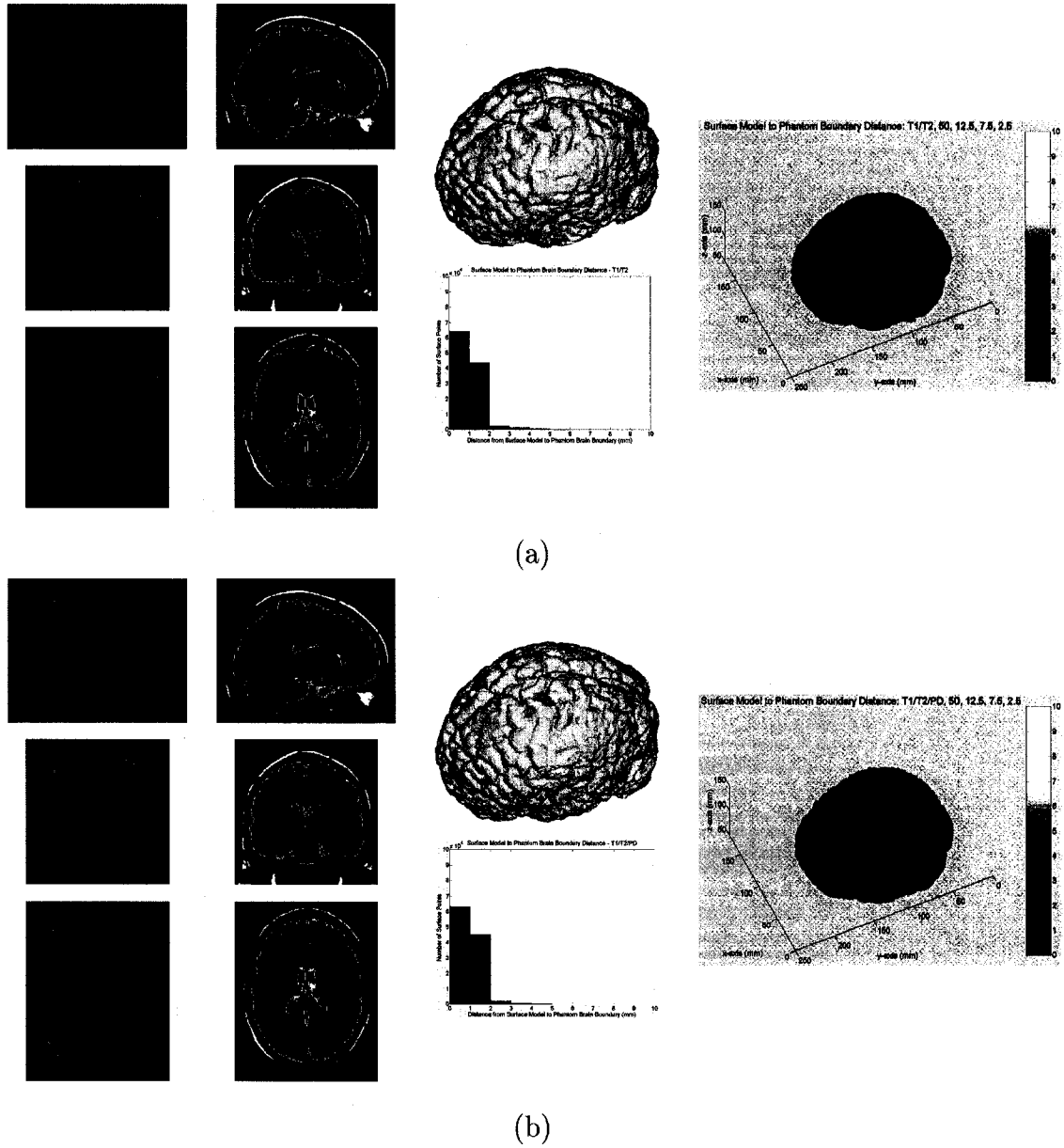


Figure 7.13 : *Tissue classes produced by MD classifier, as in figure 7.12, for (a) T1- and T2-weighted data (avg. = 0.53 ± 0.74 mm.), and (b) T1-, T2- and PD-weighted data (avg. = 0.50 ± 0.65 mm.).*

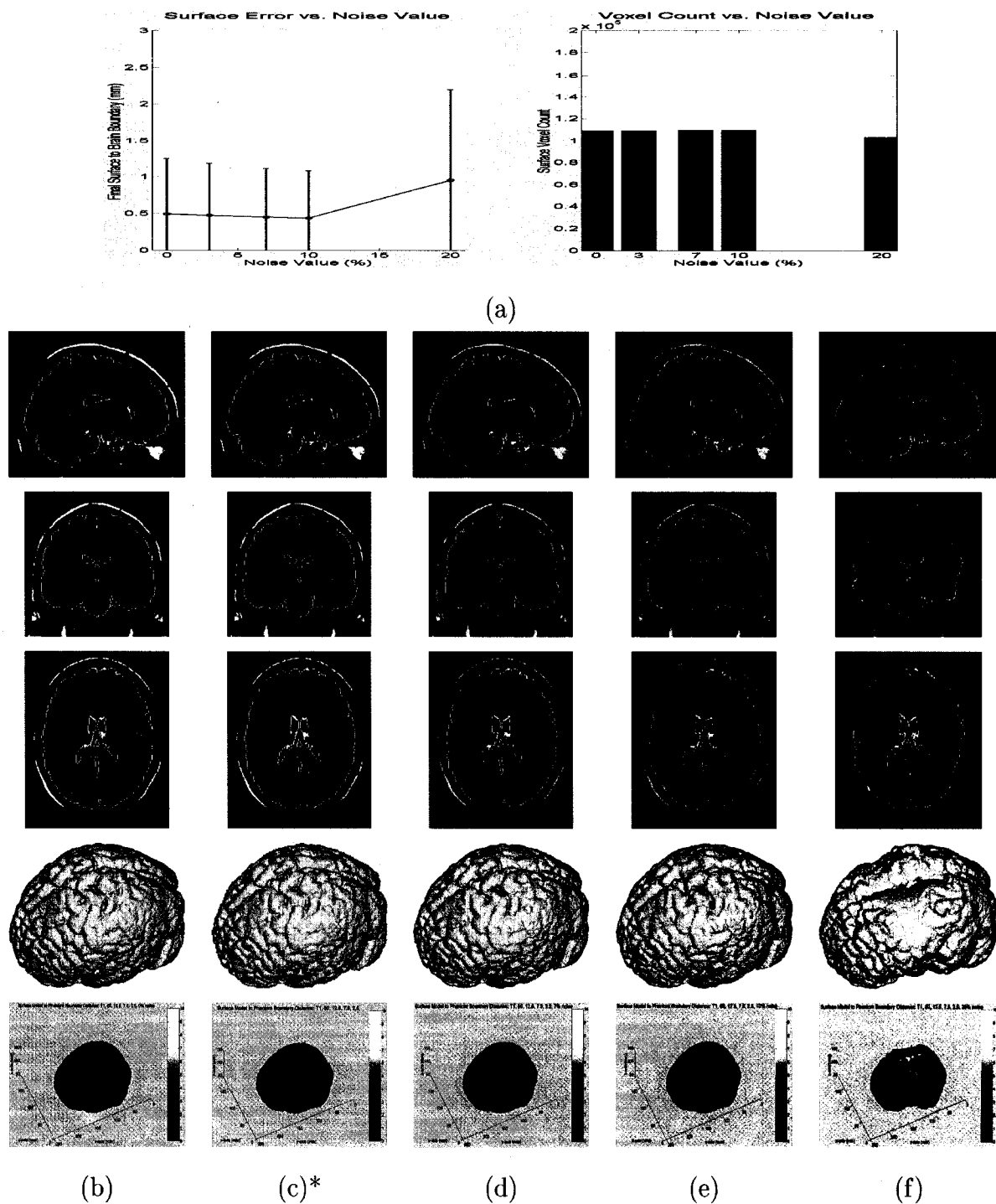


Figure 7.14 : *Sensitivity to noise level: plots of statistics of final surface to brain boundary error (left) and number of surface voxels (right) against noise level; (b)-(f) tri-planar and rendered views of final surface, plus 3D colour plot of error, for (b) noise= 0%, (c) noise=3%, (d) noise=7%, (e) noise=10%, and (f) noise=20%.*

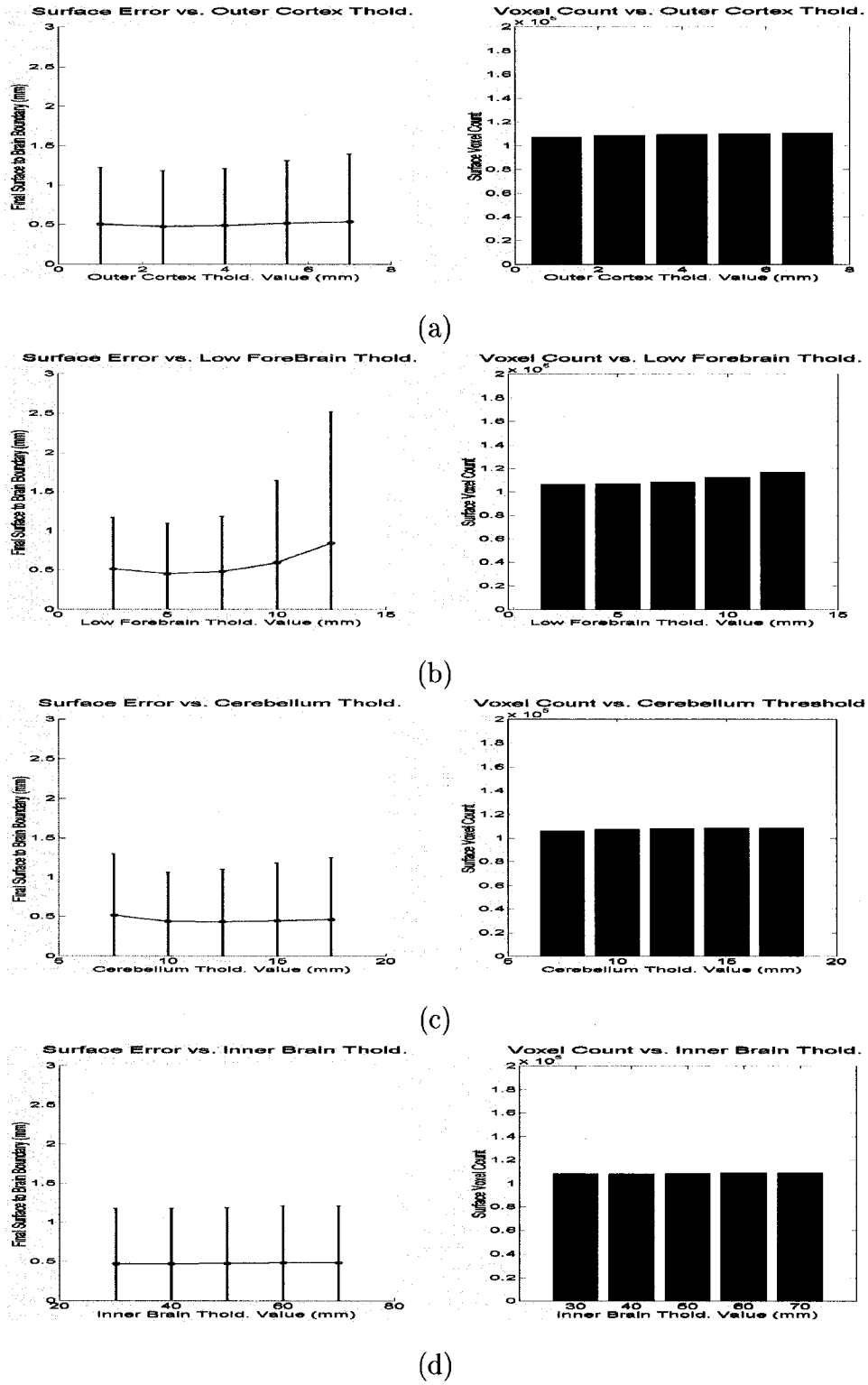


Figure 7.15 : *Sensitivity to model parameters (T1): plots of statistics of final surface to brain boundary error (left) and number of surface voxels (right) versus a variation of the following WDF thresholds: (a) outer cortex; (b) frontal lower brain; (c) cerebellum; and (d) inner brain.*

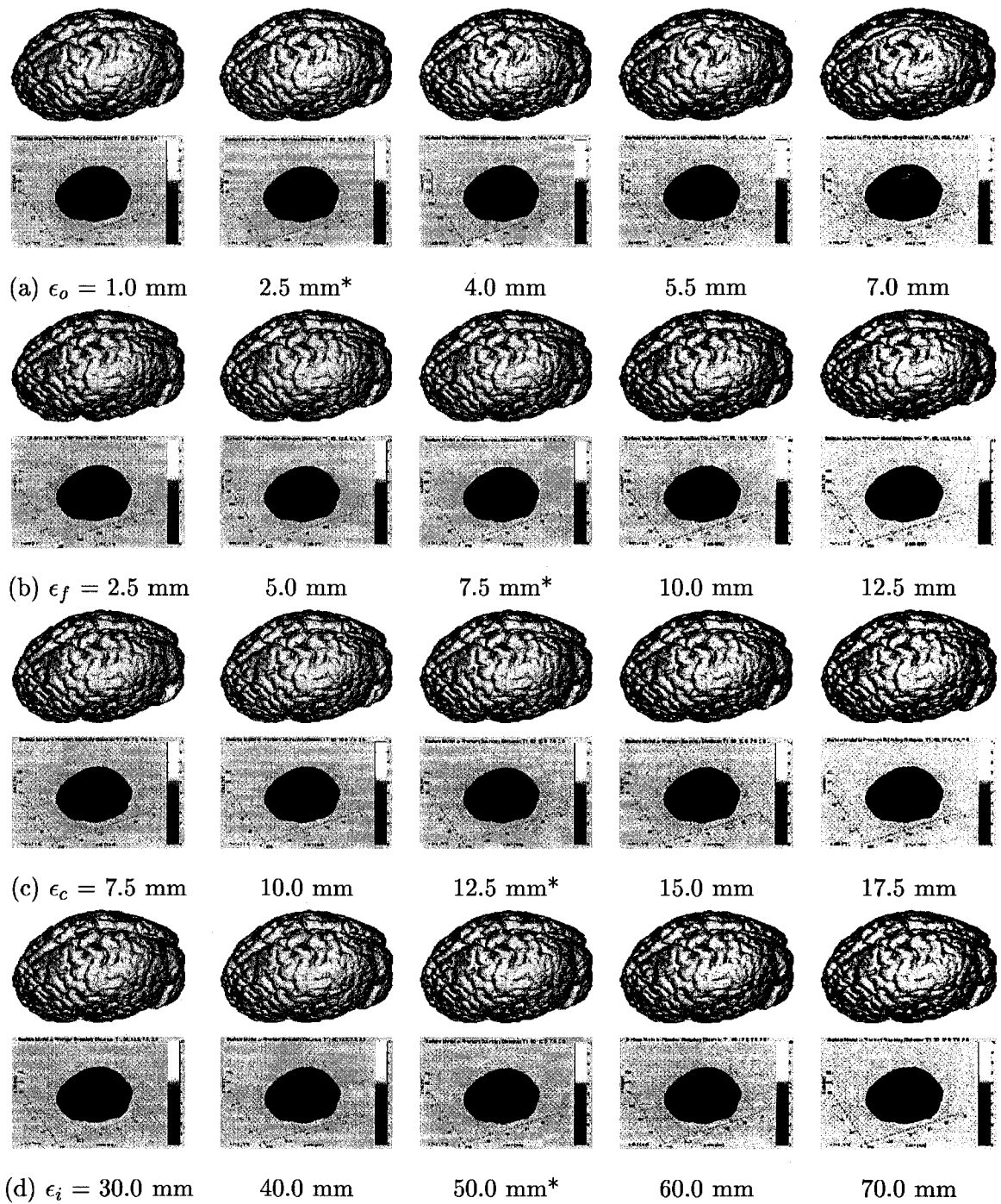


Figure 7.16 : *Sensitivity to model parameters (T1): rendered views of final surface, plus 3D colour plot of error, for variations of the following WDF thresholds, about a standard value indicated by an asterisk (*): (a) outer cortex; (b) frontal lower brain; (c) cerebellum; and (d) inner brain.*

7.5.2 Healthy Subject Validation: ICBM Real Subject Data

The evolution of the zero-level isosurface, from spheres imbedded in white matter, through the white matter boundary, to the outer brain surface, is first illustrated for healthy subjects 00100, 00101 and 00102 from the ICBM database, figure 7.17. These examples represent the worse case, from a classification standpoint: T1 data considered alone, as one would expect the tissue discrimination ability of the classifier to improve with the consideration of T2 and PD data. Indeed, from T1 intensity alone it is almost impossible to distinguish between grey matter and some other tissue classes, such as muscle as well as the sagittal sinus, so that in practice the classifier is trained on only 5 classes under these conditions: GM/muscle, WM, CSF, fat, and background. It is precisely when only T1 data are available that the coupling to white matter surface is most important, which justifies the decision to adopt an image force based on a formal tissue classification and a flexible constraint linking GM to WM.

Final results are also shown for subjects 00103 and 00104 in figure 7.18, demonstrating the capacity of the model to exploit T2 and PD data when available. Once again, slight bleeding effects are seen below the brain, particularly in the multimodal case, but these would likely have little clinical impact.

All brain surfaces were extracted under identical conditions, with regard to threshold and halting conditions, the latter of which is set to 50 “seconds” for the white matter surface, and 75 “seconds” for the final brain surface. The surface model was demonstrated here on data warped to Talairach space, because it is in this form that co-registered T2 and PD data are available. However, the model was developed on “native space” T1 data. Moreover, it is later demonstrated on “native space” MNI patient data, so that warping to Talairach space is not a requirement of the model.

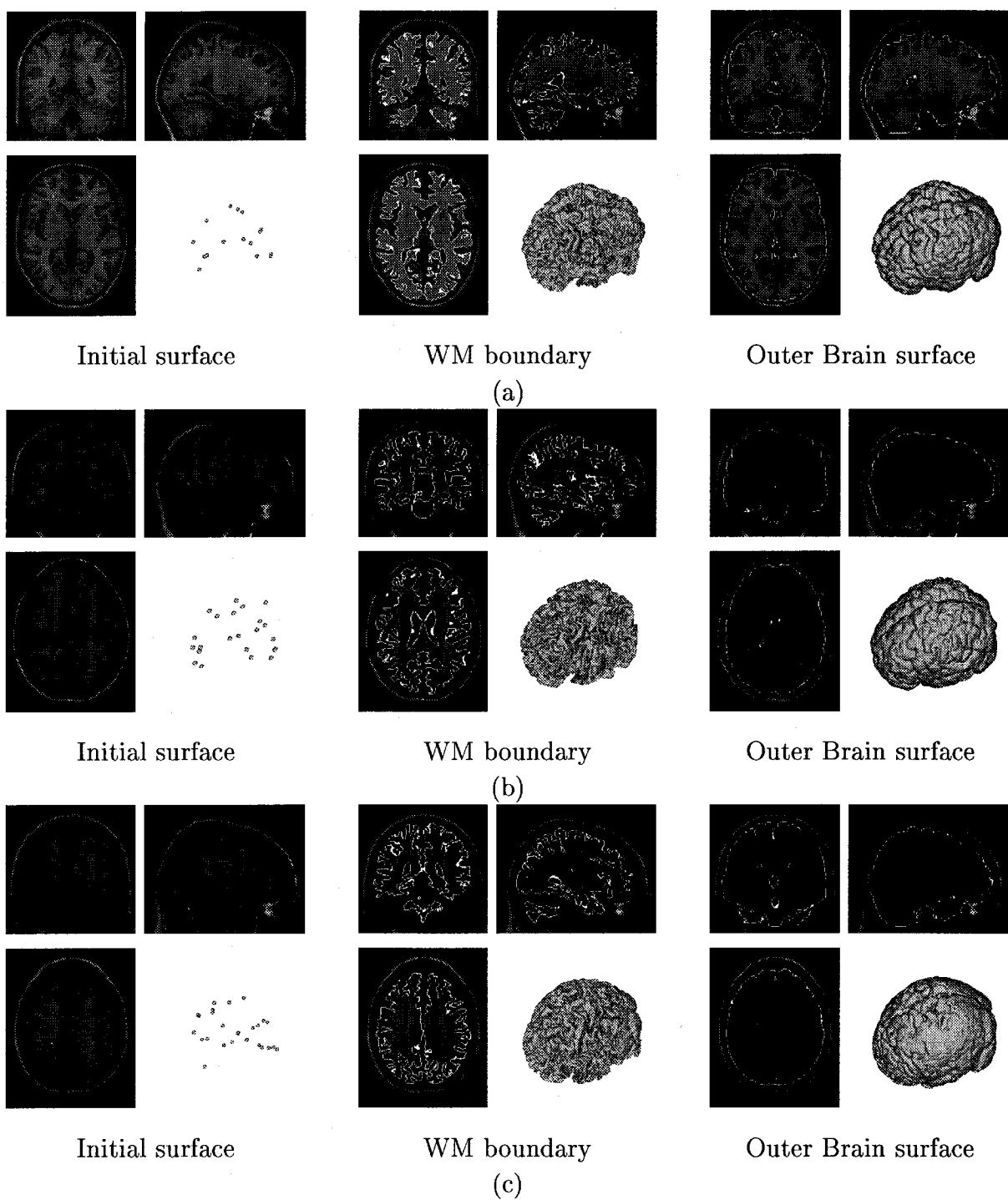


Figure 7.17 : *Evolution of the surface model, with real T1-weighted data: (a) ICBM subject 00100; (b) ICBM subject 00101; (c) ICBM subject 00102.*

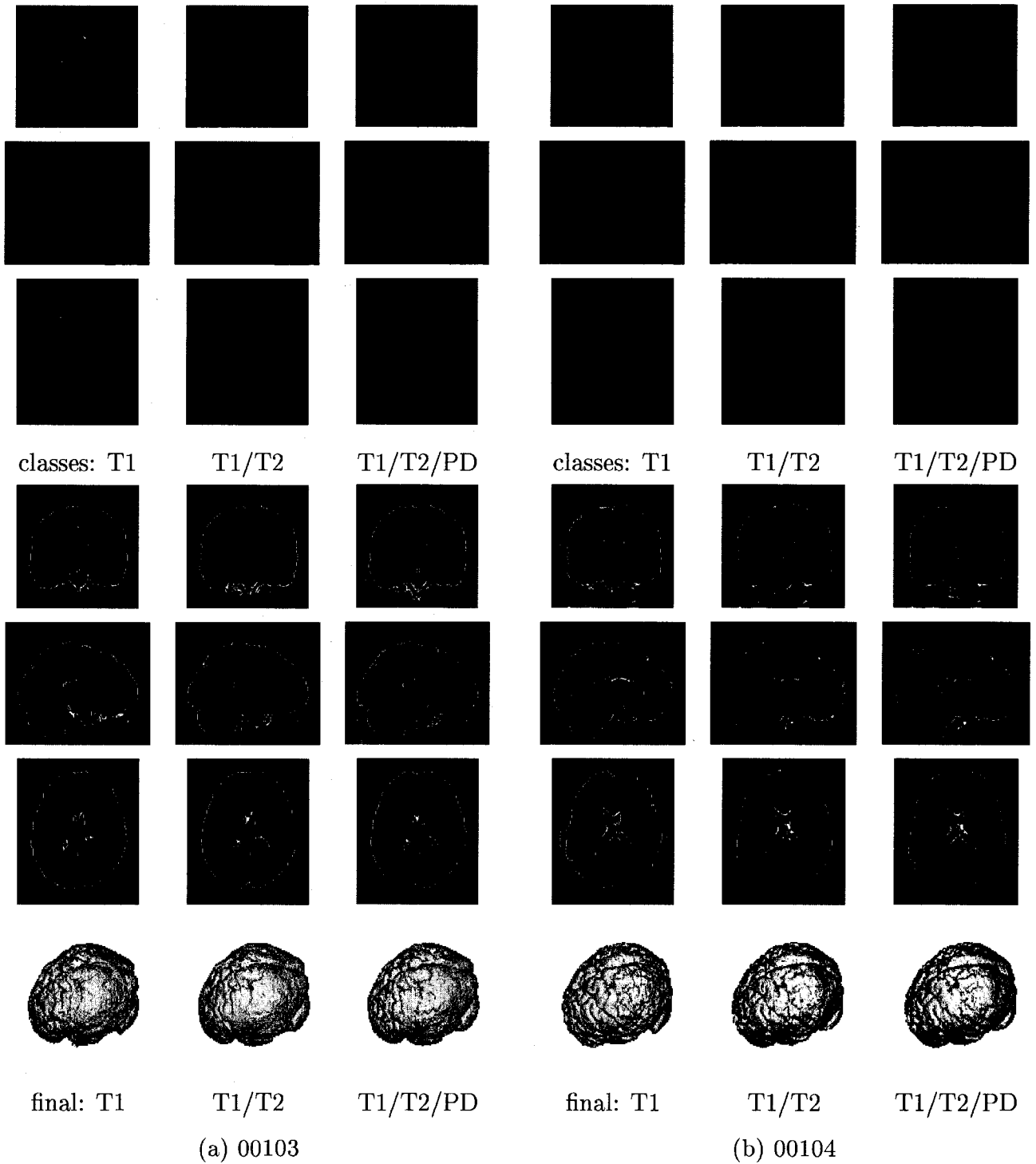


Figure 7.18 : *Application to ICBM subjects (a) 00103 and (b) 00104, based on the model acting on T1 data alone, on T1 and T2, and on T1, T2, and PD. Top: tissue classes based on each feature combination; bottom: tri-planar and rendered views of the corresponding final surface.*

7.5.3 Pathological Subject Validation: Synthetic Digital Phantom-based Data

This study uses synthetic T1-weighted data produced with a digital phantom derived from the standard digital head phantom described earlier, but featuring an imbedded astrocytoma, whose magnetization properties are ascribed by Bottomley [43]. The objective of this study was to demonstrate that the surface model can cope with pathology, to the extent that the localization of the outer brain surface is unaffected by it, and that the final surface excludes pathological voxels. This is also a justification to opt for a more flexible coupling between GM and WM, which makes allowances for imbedded pathology, where a rigid assumption of WM surrounded everywhere by a layer of GM cannot.

7.5.4 Pathological Subject Validation: Real Patient Data

Figure 7.20 provides an illustration of the brain surface identification in a manner that excludes real pathologies. Figure 7.20 (a) is an example where the speed function logic of expression (6.4) is used, where a prior classification with two explicit classes restricts the evolution of the model to WM and GM. In figure 7.20 (b) however, there is a vast expanse of edema which dominates one hemisphere, so that GM surface cannot evolve properly around the pathological tissue to identify the outer brain surface. A modified procedure is adopted, where the pathology boundary is first identified from spheres coinciding with pathology training points, followed by that of WM from spheres coinciding with WM training points. Subsequently, the brain boundary is identified from GM voxels whose final position is near both the pathology and WM surfaces. The final brain surface is comprised of the edema boundary and the healthy brain tissue boundary. These examples suggest that for real patient data featuring pathologies whose shapes are inherently unpredictable, a flexible segmentation method benefits from the treatment of image data in a manner that correlates with anatomical information, as is the case with a classifier.

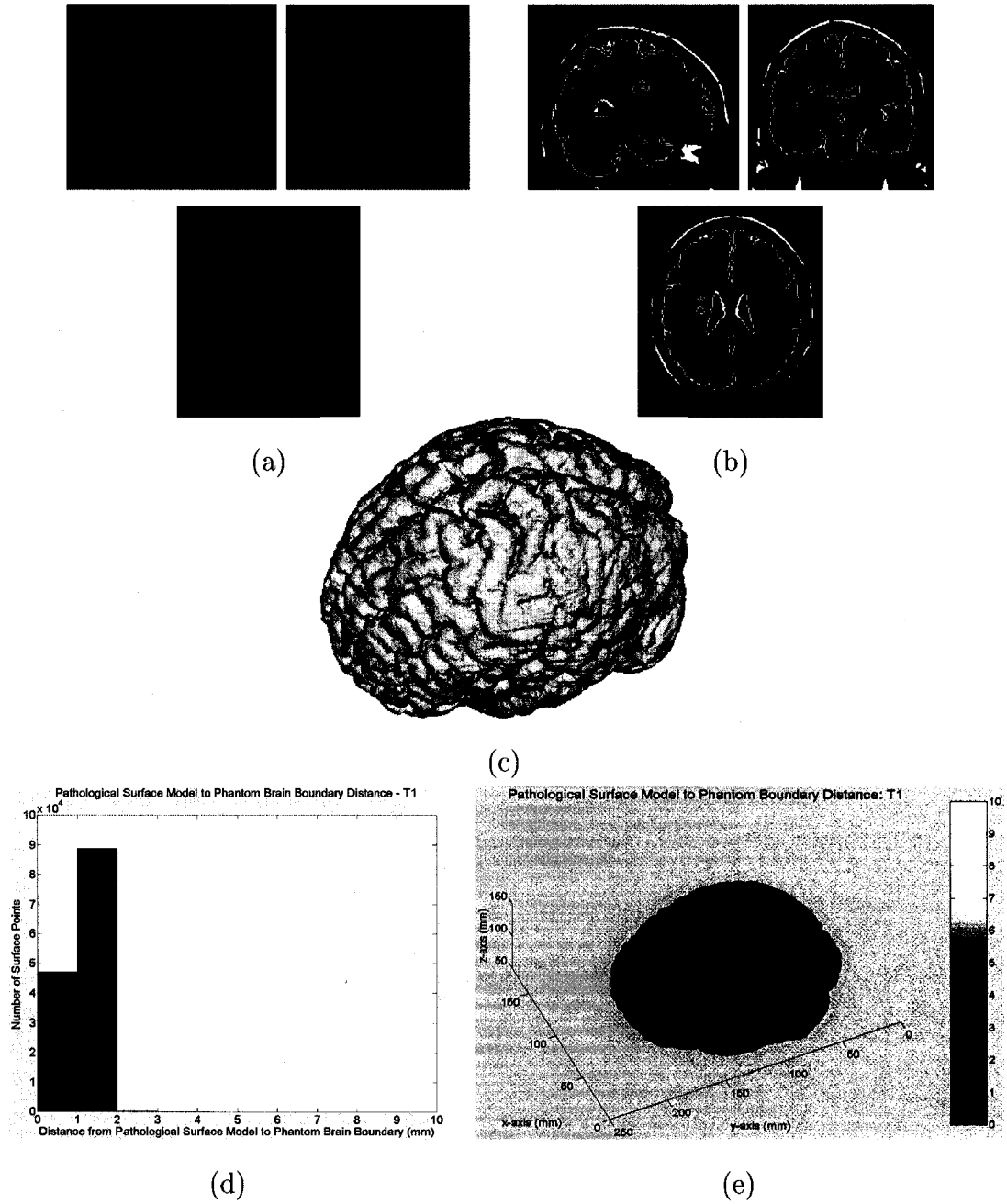


Figure 7.19 : *Surface model behaviour with synthetic pathological data: (a) original pathological phantom and (b)-(e) brain surface identified with surface model, from simulated T1-weighted scan: (b) tri-planar view; (c) rendered view; (d) distribution of distance between brain surface in scan and phantom brain boundary; (e) error between brain surface identified in scan and underlying brain boundary in phantom, shown as colour plot: $avg. = 0.68 \pm 0.51$ mm.*

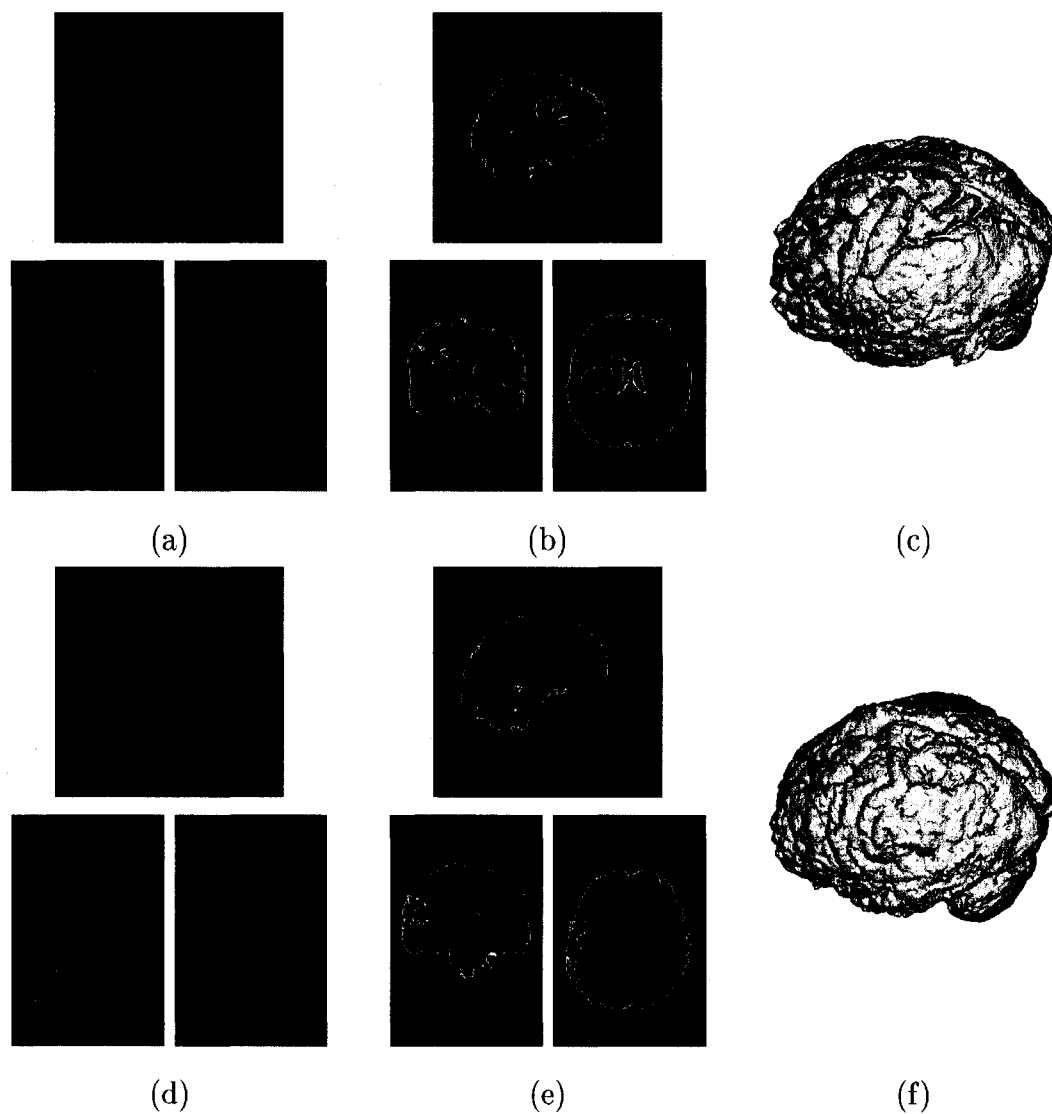


Figure 7.20 : *Surface model behaviour with real pathological T1-weighted data; (a)-(c) patient 1: (a) classification results, featuring 7 classes; (b) tri-planar view; (c) rendered view; (d)-(f) patient 2: (d) classification results, featuring 7 classes; (e) tri-planar view; (f) rendered view.*

7.6 Discussion and Conclusion

This chapter presented a validation study of the various computational stages of the surface evolution model detailed in chapter 6. The validation study of each stage made use of a digital anthropomorphic head phantom, and in many cases a MR simulator that used this phantom as input and produced synthetic T1, T2 or PD data as output. The validation began with a comparison of Gaussian and anisotropic diffusion, which concluded that the two approaches were characterized by comparable localization error. However, anisotropic diffusion was shown to be slightly superior in terms of the uncertainty of the edge position, the preservation of the gradient magnitude over increasing scale and the performance of the subsequent Minimum Distance classifier. A comparison of Minimum Distance and semi-supervised Fuzzy C-Means classifiers found the latter's convergence to be undermined by border effects which worsen with increasing filter scale. The Fast Marching technique for distance and closest point computation was validated on the basis of the similarity of its results with a brute-force closest point search. Finally, the surface evolution model itself was validated with both synthetic and real MR data, featuring both healthy subjects and patients with lesions.

One point not discussed so far is that the model as currently implemented assumes a voxel-based boundary, in particular because the FM technique employed at the time used a voxel-based surface as a starting point. However, FM techniques for triangulated surfaces have since been published [343], and it should be feasible to adapt the Closest Point Computation to triangulated surfaces as well. This improvement would be significant, as it would produce a Distance and Closest Point Map of a surface established with sub-voxel accuracy. The only difficulty then would be to devise a strategy which would refine each voxel a Closest Triangle Map to a Closest Point Map which would lead to homologous pairs usable for registration, on the basis of a projection of that voxel onto the closest triangle.

Finally, issues worth exploring include the limitations of this surface model, and whether a physically-based model could have been used as effectively. One weakness

of the surface evolution model is the limited regularization property of the smoothing term, compared to the elastic surface tension achievable in physical models. Because of this trait, the model presented here is comparatively susceptible to bleeding effects, and requires an explicit coupling mechanism between a stable WM core and the final brain surface. However, physical models which are topologically adaptable must also compromise on the strength of their regularizing term, so that they too may be afflicted by unintended bleeding in low-contrast areas. Moreover, while topologically-adaptable physical models have been recently published for identifying 2D contours, a 3D surface formulation is not as easily derived as contour evolution models are extended to surfaces. Nevertheless, an equivalent implementation based on a physical model with stronger tensile properties could be achieved, for example by extending Montagnat's topologically adaptable simplex contours to surfaces in 3D [85]. The integration of a classifier with such a model would once again be beneficial, proceeding from spheres coinciding with training points, and possibly exploiting tissue-specific refinements to the model.

Chapter 8

Conclusions and Future Directions

This thesis research addressed the problem of intra-surgical brain deformation, and suggested a non-rigid surface registration framework for characterizing this deformation. The objective of this framework was to produce a smoothed, dense vector displacement function from the undeformed preoperative brain surface, semi-automatically identified in MR data, to the exposed intraoperative brain surface characterized with a laser range-sensor and referred to the MR coordinate system. This vector displacement function could then be interpolated everywhere in the brain volume, on the basis of an ulterior volumetric, constitutively realistic FE model.

Furthermore, the methods proposed here were conceived from assumptions based on practical clinical requirements. First, the semi-automatic method for identifying the brain surface in MR should be able to cope with the presence of imdedded pathology, even very large tumours. Next, the 3D surface capture and surface registration methods should be fast, producing results virtually at the instant they are required in the OR, with no user intervention except to initiate the processes. And finally, the method for relating the MR and range spaces must function despite the presence of large swaths of surgical draping, possibly occluding one or more intrasurgical fiducials, which suggests a calibration-based solution. Just how sensitive these methods are to myriad clinical factors such as the administration of drugs and the size of the tumour and of the craniotomy, and how accurate is volumetric motion estimation based on topmost surface tracking and constitutive modelling, should be addressed

in future work. The nature of the boundary conditions on the constitutive modelling was not discussed in detail. The framework proposed here would best function by initially identifying rigid boundaries constraining a volumetric finite element model, by explicitly finding bone tissue voxels with the classification algorithm described in this thesis, possibly with the aid of co-registered and resampled preoperative CT data.

Chapter 3 proposed a method based on a laser range-sensor for determining the 3D shape of the cortical surface in the operating room, as well as a means for transforming 3D range coordinates to MR space. This procedure was described as the first requirement for determining a baseline for ascertaining the displacement of the brain during surgery. The sensor, providing measurements in the $x - z$ plane, was factory-calibrated, while the linearity of the measurements along the positioner, or y , axis was established with a positioner calibration plate described in this thesis. The rigid transformation between the range and MR coordinate systems was obtained by combining the range-probe transformation at the time of the operation, computed by calibration and sensor base tracking stages, with the IGS probe-range transformation. The validation of this procedure involved a cranial skin phantom featuring adhesive surgical fiducials, where accuracy was assessed on the basis of fiducial registration error statistics between range and MR coordinates. Future directions for research include the mechanical adaptation of the range-sensor for use in the OR, enabling the repeated positioning of the sensor in scanning and retracted configurations, and refinements to the range-MR transformation procedure, possibly involving implantable fiducials whose detection in range and MR data could be automated.

Another requirement for establishing the baseline for intra-surgical motion is the identification of the brain surface in MR, and a surface model designed to achieve this was presented in the chapters 5 to 7. Refinements to surface models contributed by this research include the use of a classifier to determine the image-based speed function of the model, as well as a new, more flexible approach to coupling the white matter boundary to the outer grey matter surface. Validation studies for this model comprised qualitative studies, featuring healthy subject and patient data, and quan-

titative studies which involved a digital anthropomorphic head phantom as well as data produced by a numerical MRI simulator. These validation studies constitute an important contribution to the application of surface models to brain surface identification. The brain surface identified in MR data by this surface model was post-processed to yield a Closest Point Map defined over the MR domain. A subsequent Iterative Closest Point registration with a range surface, taking place in the OR, would then be accelerated by the explicit determination of homologous point-pairs by table look-up. Future directions for improving the model include a triangulated surface approach to estimating the zero-level isosurface, from which distances and closest points could then be computed, whereby sub-voxel registration would then be achievable. A topologically adaptable surface model with more control over tensile properties was also seen as desirable.

Details of a new, highly efficient ICP registration method, exploiting homologous pairs obtained by look-up as well as 2D recursive splines, were presented in chapter 4. Special considerations for the non-rigid registration of range data with MR data were emphasized. These considerations include the greater density of range data in comparison with the brain surface points in MR, the use of continuity to disambiguate and validate homologous pairs, and the computation of an inverse MR-range (or rather, preoperative to intraoperative) transformation from a forward range-MR transformation. Validation of the registration made use of a new, constitutively realistic, deformable brain phantom, whose deformation could be triggered from within, as well a study involving an analytical displacement applied to synthetic cortical range images derived from a digital anthropomorphic head phantom. The results of the registration were comparable to other ICP techniques, in terms of the proximity between surfaces produced by the final transformation. Furthermore, the estimation of the non-rigid motion was on average within 1mm, but the slightly less accurate characterization achievable in sulci suggests that a more shape-sensitive ICP registration should be considered.

For this surface framework registration to be clinically useful, a feasible objective of each stage should be sub-millimeter accuracy. The overall fiducial registration error

was shown to be in the 2mm range, and based on an analysis of the accumulation of target and fiducial registration errors, these results were in a range consistent with the localization technology, which suggests that sub-mm accuracy can be achieved either by substituting an optical tracker for the FARO arm or by adopting implantable fiducials optimally designed for automatic detection. The surface model validation showed that accuracy to within roughly 0.5 mm is achievable, although this could be improved with triangulated surfaces. However, the discrete phantom used for validation does not incorporate partial volume effects, and as such this accuracy estimate could deteriorate as a result. However, there exist promising voxel classification methods which explicitly account for partial volume [66, 345]. The ICP registration on average is characterized by an accuracy within 1mm, but the spatial distribution of the error suggests that further improvement, resulting from the consideration of local surface shape, is desirable.

The underlying assumption of this thesis is that an accurate volumetric displacement characterization is possible with a displacement map at the exposed cortical surface and with a suitable finite-element model. The documented use of volumetric FE models, with sparse surface displacement data and with surface models fit to intra-operative MR data, suggests that at the very least, the accurate volumetric estimation of intra-surgical brain displacement with this framework is highly feasible. A future validation study which would support this hypothesis is the comparison of the position of beads imbedded in an elastic brain-shaped phantom, as detected within a MR volume, against the position predicted on the basis of a surface registration and volumetric interpolation approach. Furthermore, the method presented here is entirely compatible with the integration of probe-based information about the volume of resected tissue [356], whose interaction with the finite-element brain model could be inspired from cutting models used in surgical simulation [234]. Finally, one clinical application of this surface registration framework which is less reliant on volumetric interpolation is the warping of a patient-specific cortical function map to that patient's intraoperatively exposed cortex in the OR, in the context of stimulating a particular functional area [395]. Such an application could make direct use of the

brain surface displacement field computed with the methodology presented in this thesis.

Appendix A

A Survey of Biomechanical Modeling of the Brain for Intra-surgical Displacement Estimation and Medical Simulation

A.1 Introduction

Biomechanical modeling of human and animal brain tissue is a growing field of research, whose applications currently include simulating, with a view to minimizing, head injuries in car impacts [425], generically modeling dynamic behaviour, such as brain shift [358], in the surgical theater, and increasingly, providing medical experts with clinical tools such as surgical simulators [139] and predictive models for tumour growth [201, 412]. Another clinical application currently being investigated is the compensation of a 3D patient-specific graphical model, used in image guidance, for intrasurgical brain shift [258, 110], to account for quantitative displacement information provided in the OR by a range sensor or by a hand-held locating device.

The requirements of surgical simulation and intrasurgical deformation estimation for image guidance are a trade-off between computational efficiency and realism, due

to the need of the former to provide a response to a virtual surgical intervention that is representative of human tissue in terms of continuity and motion amplitude. Realism is required to give the surgeon precise volumetric displacement information on demand, within a short time frame deemed tolerable in a surgical context. However, in the case of simulation, this trade-off favours efficient computation, i.e.: update rates in excess of 100 Hz[52], particularly if haptic feedback is involved. In contrast, image guidance will tolerate somewhat longer computation times if a high degree of realism, such as quantitatively predicting material behaviour to sub-mm resolution, can be achieved.

The focus of this appendix, which serves as a foundation for a survey paper [17], is to review the important contributions to biomechanical modeling of healthy and pathological brain tissue, as well as general techniques applicable to simulation and accurate image guidance of brain surgery, such as numerical efficiencies ¹.

A.2 Preliminaries

A.2.1 Biomechanics

Biomechanics [118] is the study of the mechanics of living tissue, particularly from a continuum mechanics [239] perspective. The latter is the branch of mechanics concerned with external loading forces in solids, liquids and gases, with the resulting deformation or flow of these materials, and the state of internal traction, or *stress*, inherent in these materials. The deformation of a solid is referred to as *strain* (see Malvern [239] for a review).

The dynamic behaviour of an individual material or tissue is characterized by its *constitutive equations* relating stress to strain. Constitutive equations are relevant to characterizing the dynamics of the brain but cannot be solved without expressing

¹The co-authors contributed as follows. M. Miga suggested some readings on the physiological basis for poroelastic models. J. Nemes proof-read this manuscript and pointed out inconsistencies in the notation. K. Chinzei provided insight on rheological experiments. T.M. Peters proof-read the final manuscript.

the unknown stress tensor field as a function of the strain tensor field, which can be estimated from known displacements. While the complex relation between stress and strain defies a closed-form description, it can be analyzed by considering an *idealized material response*.

Biomechanical modeling of the brain is often approached by finding an numerical solution for the displacements, deformations, stresses and forces, as well as possibly other states, such as hydrostatic pressure, in relation to a history of “loading”. There are two basic approaches for numerically estimating or simulating biomechanical deformations: *finite element modeling* and *mass-spring systems*. Finite element modeling [26, 428] has become the standard method for quantitatively analyzing a wide variety of engineering problems, typically of a mechanical or electromagnetic nature, and in particular for material deformation. The analysis of material deformation is based on expressing equations which characterize the mechanical equilibrium and which must be satisfied everywhere in the system under investigation. In a few simplified cases, an analytical solution to these equations can be determined, but in general this solution must be numerically estimated. An exact solution would require force and momentum equilibrium at all times everywhere in the body, but the finite element method replaces this requirement with the weaker one that equilibrium must be maintained in an average sense over a finite number of divisions of the volume of the body. The actual division of complex geometries into simple elements, such as tetrahedra and hexahedra, corresponds to the *meshing* problem [298], as illustrated in figure A.1, which remains an active research area.

A mass-spring system [415, 127, 81] is an approximation of a biomechanical system as a collection of point masses connected by elastic springs, and is derived from the field of computer animation. The parameters available to determine the biomechanical behaviour are the mass values and the visco-elastic spring characteristics.

Linearly Elastic Solid

The simplest idealized solid is the *Hookean elastic solid*, which is characterized by a *linear* elastic response. For a cylindrical bar subject to a tensile or compressive

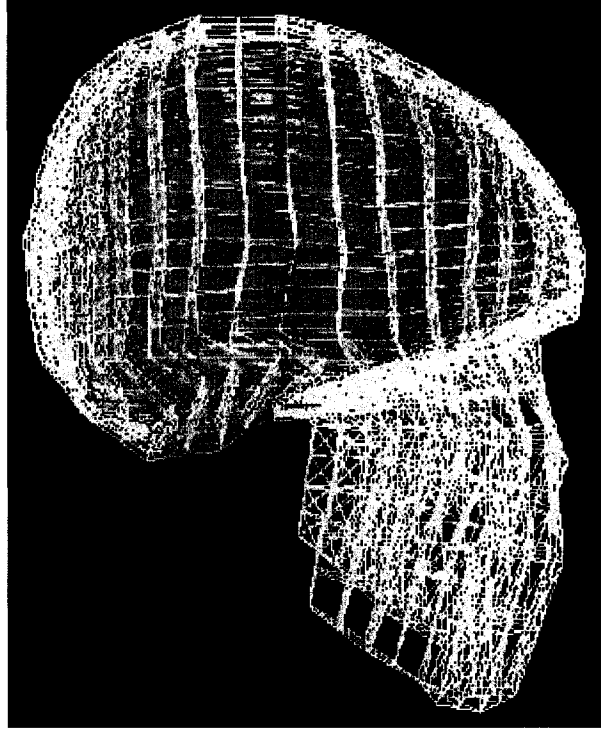


Figure A.1 : *Meshing used at Wayne State University, courtesy of C. Zhou [425].*

stress σ in its axial direction, the resulting strain ϵ is given by $\sigma = E\epsilon$, where E is Young's modulus and is a characteristic of the material. This relationship can also be stated in terms of a *compliance* J : $\epsilon = J\sigma$. For an incompressible material, this cylinder undergoes a decrease or increase in its diameter respectively. The ratio of radial strain ϵ_d to axial strain ϵ_a is given by Poisson's ratio: $\nu = -\frac{\epsilon_d}{\epsilon_a}$ and is also a characteristic of the material. In 3D, an elastic material is characterized by the stress tensor σ which is linearly proportional to the strain tensor ϵ :

$$\sigma = \mathcal{C}\epsilon \text{ or } \sigma_{ij} = C_{ijkl}\epsilon_{kl} \quad \text{and} \quad \epsilon = \mathcal{S}\sigma \text{ or } \epsilon_{ij} = S_{ijkl}\sigma_{kl}, \quad (\text{A.1})$$

where $\mathcal{C} = [C_{ijkl}]$ and $\mathcal{S} = [S_{ijkl}]$ are 4th order tensors ($3^4 = 81$ components) of elastic moduli and compliance respectively, and where the *Einstein summation convention*² is used. This expression simplifies considerably under assumptions of elastic symmetry

²Indices which are repeated on either side of the equal sign, in this case k and l , indicate summations over these indices. This convention is maintained throughout the text.

and isotropy, namely:

$$C_{ijkl} = \lambda \delta_{ij} \delta_{kl} + \mu (\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}) , \quad (\text{A.2})$$

where λ and μ are Lamé's elastic constants and where δ is the Kronecker delta³ [239]. These are related to Young's and shear moduli E and G , and to Poisson's ratio ν as follows:

$$\mu = G = \frac{E}{2(1+\nu)} \quad \text{and} \quad \lambda = \frac{\nu E}{(1+\nu)(1-2\nu)} , \quad (\text{A.3})$$

whereby the isotropic Hooke's law is a system of 6 equations of the stress components $\sigma_1 \sigma_2 \sigma_3 \tau_{12} \tau_{23} \tau_{13}$ expressed as follows:

$$\begin{aligned} \epsilon_1 &= \frac{1}{E} [\sigma_1 - \nu(\sigma_2 + \sigma_3)] & \gamma_{23} &= \frac{1}{G_{23}} \tau_{23} \\ \epsilon_2 &= \frac{1}{E} [\sigma_2 - \nu(\sigma_3 + \sigma_1)] & \gamma_{13} &= \frac{1}{G_{13}} \tau_{13} \\ \epsilon_3 &= \frac{1}{E} [\sigma_3 - \nu(\sigma_1 + \sigma_2)] & \gamma_{12} &= \frac{1}{G_{12}} \tau_{12} \end{aligned} \quad (\text{A.4})$$

Newtonian Fluid

A simple fluid idealization which is relevant to modeling the fluid constituent of brain tissue is the *Newtonian fluid*. Fluid *at rest* or *in uniform flow* cannot sustain a shear stress, so that the shear (off-diagonal) components of stress are null. Moreover, stress in this case is assumed *hydrostatic* (its principal stresses are equal): $\sigma_{ij} = -p\delta_{ij}$. In a *deforming* fluid, the total stress includes a *viscosity* component which is a function of the rate-of-deformation tensor $\mathbf{D}$: $\boldsymbol{\sigma} = -p\mathbf{I} + \mathcal{F}(\mathbf{D})$. If $\mathcal{F}$ is assumed linear, i.e.:

$$\sigma_{ij} = -p\delta_{ij} + C_{ijkl} D_{kl} , \quad (\text{A.5})$$

the fluid is called Newtonian. Under assumptions of symmetry and isotropy, the viscosity tensor $[C_{ijkl}]$ also simplifies in the same manner as equation (A.2), where this time λ and μ are two independent parameters of viscosity. Also, if the shear terms are deemed negligible, the deforming fluid is called *inviscid*. Other material idealizations relevant to modeling brain tissue are presented in §A.3.

³ $\delta_{pq} = 1$ if $p = q$, and 0 if $p \neq q$.

A.2.2 Numerical Estimation

Finite Element Modeling

The displacement finite element method numerically solves for unknown displacements, deformations, stresses, forces and possibly other variables of a solid body. An exact solution would require force and momentum equilibrium at all times everywhere in the body,

$$\int_S \mathbf{t} dS + \int_V \mathbf{f} dV = 0 \quad \int_S (\mathbf{x} \times \mathbf{t}) dS + \int_V (\mathbf{x} \times \mathbf{f}) dV = 0. \quad (\text{A.6})$$

but the finite element method replaces this requirement with the weaker one that equilibrium must be maintained in an average sense over a finite number of divisions of the volume of the body. The application of the equilibrium equations to numerical analysis is based on using *Gauss' theorem* to restate the equilibrium conditions as a single integral, called the *virtual work principle*.

The volume which is modelled is defined as Ω and is subject to boundary conditions. Assuming Cartesian coordinates, and adopting the nomenclature of [428, 10], the displacement at the node i of a given element is labelled $\mathbf{a}_i = [u_i \ v_i \ w_i]^T$, while the displacement at any point in Ω is expressed $\mathbf{u} = [u(x, y, z) \ v(x, y, z) \ w(x, y, z)]^T$. The latter is fully determined by the nodal displacements and by the *shape functions* that govern the interpolation between them. For a tetrahedral element, we have:

$$\mathbf{u} = [(\mathbf{I}N_i) \ (\mathbf{I}N_j) \ (\mathbf{I}N_m) \ (\mathbf{I}N_p)] \mathbf{a}^e \equiv \mathbf{N}\mathbf{a}^e, \quad (\text{A.7})$$

where $N_i = 1$ at vertex (x_i, y_i, z_i) but zero elsewhere, and so on, and where $\mathbf{a}_{12 \times 1}^e = [\mathbf{a}_i \ \mathbf{a}_j \ \mathbf{a}_p \ \mathbf{a}_p]^T$ is comprised of all nodal displacements within a given tetrahedral element. For a small strain assumption, the relationship between strain and nodal

displacement is a simple one:

$$\boldsymbol{\epsilon} = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \\ \gamma_{12} \\ \gamma_{23} \\ \gamma_{13} \end{bmatrix} \equiv \begin{bmatrix} \frac{\partial u}{\partial x} \\ \frac{\partial v}{\partial y} \\ \frac{\partial w}{\partial z} \\ \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} \\ \frac{\partial v}{\partial z} + \frac{\partial w}{\partial y} \\ \frac{\partial w}{\partial x} + \frac{\partial u}{\partial z} \end{bmatrix} = \mathbf{B}\mathbf{a}^e \equiv [\mathbf{B}_i \ \mathbf{B}_j \ \mathbf{B}_m \ \mathbf{B}_p] \mathbf{a}^e \quad , \quad (\text{A.8})$$

where for example $\mathbf{B}_i$ is obtained by deriving $\mathbf{I}N_i$ appropriately. The virtual work principle states that for a virtual displacement $\delta\mathbf{a}^e$ applied to the system, static equilibrium requires that the external virtual work must equal the internal work done within the element. Defining *nodal forces* $\mathbf{q}^e$ which are statically equivalent to boundary stresses and body forces comprising boundary conditions, and $\mathbf{b}$ the concentrated loads acting on the body, the virtual work principle is expressed for an infinitesimal volume:

$$\delta\mathbf{a}^{eT} \mathbf{q}^e = \delta\boldsymbol{\epsilon} : \boldsymbol{\sigma} - \delta\mathbf{u}^T \mathbf{b} \quad . \quad (\text{A.9})$$

This expression is integrated with respect to volume, while also substituting for $\delta\boldsymbol{\epsilon}$ and $\delta\mathbf{u}$:

$$\delta\mathbf{a}^{eT} \mathbf{q}^e = \delta\mathbf{a}^{eT} \left(\int_{V^e} \mathbf{B}^T \boldsymbol{\sigma} - \mathbf{N}^T \mathbf{b} \right) dV \quad . \quad (\text{A.10})$$

Finally, $\boldsymbol{\sigma}(\mathbf{a}^e, \mathbf{B}, \boldsymbol{\sigma}_0)$ is estimated according to the constitutive properties of the assumed continuum. For a linearly and isotropically elastic solid [239], whose constitutive properties $\mathcal{C}$ simplify to a matrix $\mathbf{D}(\lambda, \mu)$, and after some manipulation [428], we have

$$\begin{aligned} \mathbf{q}^e &= \mathbf{K}^e \mathbf{a}^e + \mathbf{f}^e \quad \text{where} \quad \mathbf{K}^e = \int_{V^e} \mathbf{B}^T \mathbf{D} \mathbf{B} dV \quad \text{and} \\ \mathbf{f}^e &= - \int_{V^e} \mathbf{N}^T \mathbf{b} dV - \int_{V^e} \mathbf{B}^T \mathbf{D} \boldsymbol{\epsilon}_0 dV + \int_{V^e} \mathbf{B}^T \boldsymbol{\sigma}_0 dV \quad . \end{aligned} \quad (\text{A.11})$$

Summing the elemental stiffness matrices and forces we obtain:

$$\mathbf{K}\mathbf{a} = \mathbf{f} \quad . \quad (\text{A.12})$$

The matrix $\mathbf{K}$ is called the *stiffness matrix*, and has a sparse structure. The unique solution of expression A.12 requires one or more boundary conditions, which modify the stiffness matrix and make it nonsingular. For some dynamic systems, this equation may be modified to further include *mass* ($\mathbf{M}$) and *damping* ($\mathbf{C}$) effects:

$$\mathbf{M}\ddot{\mathbf{a}} + \mathbf{C}\dot{\mathbf{a}} + \mathbf{K}\mathbf{a} = \mathbf{f} . \quad (\text{A.13})$$

Computation Efficiencies: towards FE-based Surgical Simulation

In the context of brain shift estimation, we are interested in the collection of nodal displacements a_i that solve expression (A.12). The application of FE methods to intra-surgical deformation estimation is discussed in § A.3, within a broader survey of FE modelling of the brain. In a surgical simulation context, the concentrated loads term also accounts for user-controlled virtual cutting forces, and the corresponding set of nodal displacements is then found. Haptic feedback to the user can be computed from the surface tractions and body forces on the elements in contact with the surgical tool.

A volumetric, dynamically deformable FE approach was long thought to be too slow for implementing haptic feedback in the context of surgical simulation⁴. Recently however, some researchers [9, 25] have demonstrated the viability of efficient FE numerical schemes, designed with haptic rate force feedback in mind. As illustrated in figure A.2, Astley [9, 10] has demonstrated a FE software architecture that allows a hierarchy of meshes, featuring a *parent* and one or more *child* meshes, to be defined and computed independently. This decoupling is accomplished by representing each system as a simple equivalent, inspired from Norton and Thevenin equivalents of electronic circuit analysis, within each other's stiffness matrix. Each child mesh, particularly relevant to haptic feedback, can be dense, non-linearly elastic and restricted to a small volume near the pathology of interest, while maintaining the parent mesh linearly elastic and relatively sparse. Basdogan [25] has proposed two efficiencies for the dynamic FE equilibrium equations contained in expression (A.13): first, a

⁴In contrast to FE methods reliant on extensive precomputation[47, 30], which may preclude changes in volumetric topology.

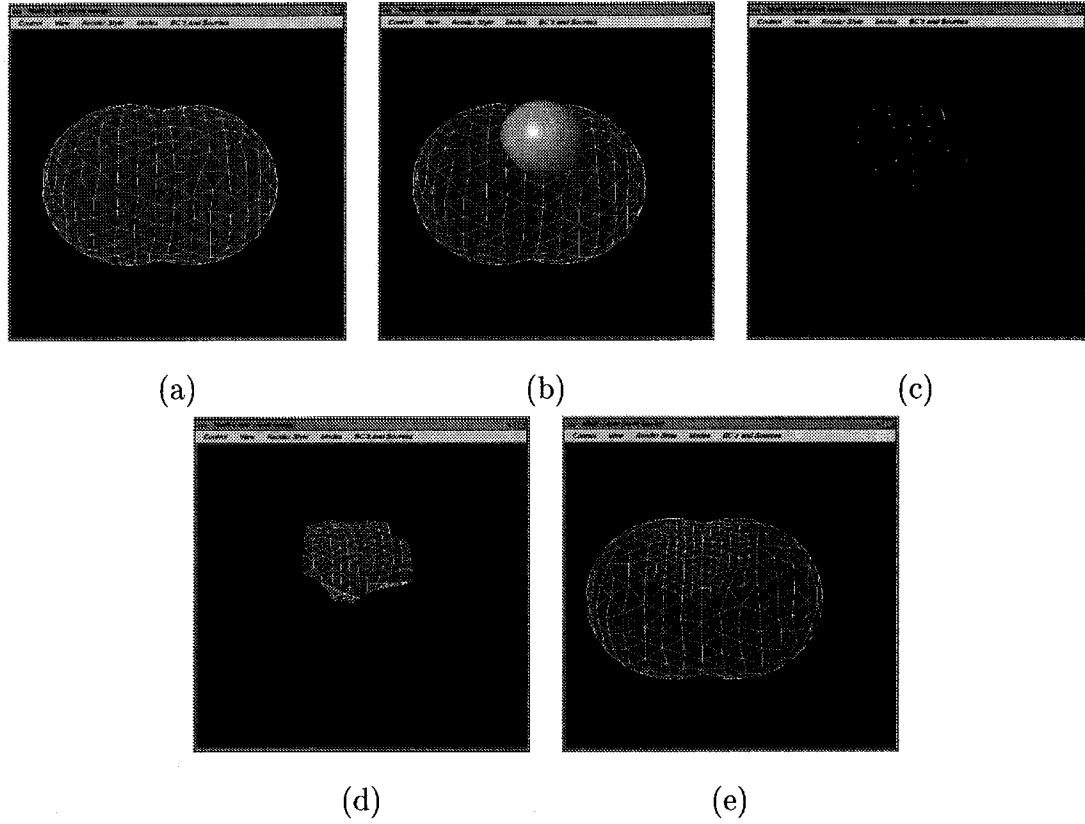


Figure A.2 : *A hierarchical FE mesh, courtesy of O. Astley [10]: (a) initial parent mesh; (b) definition of volume of interest for child mesh (c); (d) refinement of child mesh, then integrated with parent mesh.*

modal transformation and reduction, as suggested by Pentland [306] in the context of active surface models, and last, a new technique called the *Spectral Lanczos Decomposition* method, based on the re-arrangement and the Laplace transformation of expression (A.13). Berkley [30] has investigated the effects of permuting the stiffness matrix to make it narrowly banded and prioritizing the rows of expression (A.12), according to the importance of the node: boundary condition, visible, interior, or contact node. Bro-Nielsen and Cotin [47] have also considered a new partition of the FE system equation, on the basis of surface and interior nodes, and have proposed a method that inverts the stiffness matrix in a pre-computed manner. However, the complexity of stiffness matrix processing may limit the applicability of the Berkley and Bro-Nielsen techniques, given that surgically induced changes in topology would impose a sequence of new stiffness matrices over time.

Mass-Spring Systems

A mass-spring system is characterized by each node i , having a mass m_i and position $\mathbf{x}_i$, and being imbedded in a mesh where each edge coincides with a spring k . Each node is subject to an equation of the form [415]:

$$m_i \frac{d^2 \mathbf{x}_i}{dt^2} + \gamma_i \frac{d\mathbf{x}_i}{dt} + \mathbf{g}_i = \mathbf{f}_i \quad i = 1, \dots, N \text{ where} \quad (\text{A.14})$$

$$\mathbf{g}_i(t) = \sum_{j \in \mathcal{N}_i} \mathbf{s}_k \text{ and where } \mathbf{s}_k = \frac{c_k e_k}{\|\mathbf{r}_k\|} \mathbf{r}_k . \quad (\text{A.15})$$

In this equation, $\mathbf{s}_k$ represents the force on the k th spring linking the node i to a neighbouring node j . This force is a function of the vector separation of the nodes $\mathbf{r}_k = \mathbf{x}_j - \mathbf{x}_i$, of the deformation of the spring $e_k = \|\mathbf{r}_k\| - l_k$, and of the characteristics of the spring: its natural length l_k , its stiffness c_k and its velocity-dependent damping γ_i . The quantity $\mathbf{f}_i$ is the net external force acting on node i , which may include a surgical tool or the effect of gravity.

Applications of Mass-Spring Systems

In general, mass-spring systems are used mainly for surgical simulation [281, 81], and somewhat less for estimating brain shift, with the possible exceptions of Edwards [99] and of Škrinjar [359], because of the difficulty in making spring constants conform to the measured properties of elastic anatomical tissues and the criticality of accurate constitutive modelling in the OR. Their use in surgical simulation involves modeling the effect of surgical forces $\mathbf{f}_i$ in expression (A.14) by integrating the equations of motion forward through simulated time. The sum of spring forces $\mathbf{g}_i$ on the nodes in contact with the virtual tool, can then provide the user with a sense of tissue resistance. Škrinjar [359] adopted a method based on the Kelvin visco-elastic spring-dashpot model (see the section on visco-elasticity in § A.3.1), where the spring force expression (A.15) also features a term dependent on the relative velocity between nodes i and j . Edwards [99] further incorporated terms that promote material incompressibility and inhibit surface folding in the deformation computation.

A.3 Finite Element Modeling of the Brain

There are three categories of brain finite element models, expressed in terms of the types of loads simulated:

- those which view the brain under *impacts*, typically caused by auto collisions [411, 186, 335, 425];
- those which model the effect of *pathologies* [277, 201, 412];
- and recently, those which model *surgical loads* [358, 258, 266].

Brain models can also be classified according to the nature of their underlying idealized material or *continuum*, which may consider brain tissue either as

- a strictly *visco- or hyper-elastic solid*, or
- as a *hybrid of elastic solid and inviscid fluid constituents* (poroelastic or biphasic).

There is a correlation between these categorizations: most impact collision models view brain matter as a simple elastic solid, whereas tumour growth models account not only for solid and fluid constituents, but possibly for biological and biochemical factors as well, modelled as pseudo-forces [412], and finally surgical models of both solid [358, 266] and solid-liquid hybrid [258, 372, 4] types exist.

A.3.1 Solid Brain Models

This section provides an overview of solid continua and FE models of the brain, tracing their history from early impact response research, through *rheological* studies characterizing constitutive properties by experiments featuring the compression and stretching of animal brain tissue, to recent physical models better adapted for resolving surgically induced displacements.

NonLinear Solid Continua: Hyper-elastic and Visco-elastic Solids

Beyond the *linearly elastic* solid, two other idealized solids are commonly found in the literature: *hyper-elastic* and *visco-elastic* solids. Elasticity theory posits that a deformation is thermodynamically *reversible* provided that it occurs at infinitesimal speed, where thermodynamic equilibrium is maintained at every instant [204]. At finite velocities, where the body is not always in equilibrium, processes take place to return it to equilibrium. Such processes imply that the motion is irreversible and that mechanical energy is dissipated into heat. Hyper-elasticity ignores these thermal effects: work done in hyper-elastic deformation is assumed stored and available for reversing the deformation, while visco-elasticity makes no such assumption.

A hyper-elastic material is also characterized by an *elastic potential function* U (or *strain-energy* function, also denoted W), which is a scalar function of one of the strain or deformation tensors, whose derivative with respect to deformation determines the corresponding stress component:

$$\sigma_{ij} = \frac{\partial U(\epsilon)}{\partial \epsilon_{ij}}, \quad (\text{A.16})$$

where $[\sigma_{ij}]$ and $[\epsilon_{ij}]$ are work-conjugate stress and strain measures. Hyper-elasticity is based on the assumption that the elastic potential always exists as a function of strains alone [239]. One special case of this strain energy function is the Mooney-Rivlin expression for isotropic incompressible material [134]:

$$\begin{aligned} U &= \sum_{i=0}^N \sum_{j=0}^N C_{ij} (I_1 - 3)^i (I_2 - 3)^j \text{ which, taking } N = 1, \text{ reduces to} \\ U &= C_{10}(I_1 - 3) + C_{01}(I_2 - 3), \end{aligned} \quad (\text{A.17})$$

where I_1 and I_2 are the first two of three *invariants* of the strain tensor:

$$I_1 = \lambda_1^2 + \lambda_2^2 + \lambda_3^2, \quad I_2 = \lambda_1^2 \lambda_2^2 + \lambda_2^2 \lambda_3^2 + \lambda_1^2 \lambda_3^2 \text{ and } I_3 = \lambda_1^2 \lambda_2^2 \lambda_3^2, \quad \text{and where } (\text{A.18})$$

the λ_i represent the principal stretch ratios of a deformed material (note that $I_3 = 1$ for an incompressible material).

Visco-elasticity is characterized by a relationship between stress and strain which depends on time, and constitutive relations are typically expressed as an integral, i.e.:

$$\boldsymbol{\sigma}(t) = \int_0^t \mathcal{C}(t - \tau) \frac{d\boldsymbol{\epsilon}(\tau)}{d\tau} d\tau \quad \text{or} \quad \boldsymbol{\epsilon}(t) = \int_0^t \mathcal{S}(t - \tau) \frac{d\boldsymbol{\sigma}(\tau)}{d\tau} d\tau \quad (\text{A.19})$$

Some phenomena associated with visco-elastic materials include *creep*, whereby strain increases with time under constant stress; *relaxation*, where stress decreases with time under constant strain; and finally, a dependency of the effective stiffness on the rate of application of the load [202].

Finally, transient creep and relaxation responses can be modelled exponentially; for example, $J(t) = J_0 (1 - e^{-t/\tau_c})$ and $E(t) = E_0 e^{-t/\tau_r}$ respectively. Exponential response functions arise in simple discrete models composed of *springs*, which are perfectly elastic ($\sigma_s = E\epsilon_s$), and *dashpots*, which are perfectly viscous ($\sigma_d = \eta d\epsilon_d/dt$). Consequently, spring-dashpot models are considered useful idealizations for visco-elastic behaviour. The two simplest such models are the *Maxwell* model, consisting of a spring and a dashpot in *series*, and the *Voigt/Kelvin* model, with the spring and dashpot arranged in *parallel* [202].

Impact Response Models

Early (1960's and early 1970's) studies of impact response were formulated as *analytical* continuum models based on spherical, elliptical and cylindrical idealizations (see e.g.: Ljung's model for brain deformation due to skull rotation [223] and the review appearing in [193]), but the complex shape of the brain limited the applicability of this approach [411]. With the advent of the finite element method in the 1970's, the skull, brain and CSF could be divided into small elements, typically hexahedra, tetrahedra, and shells, and complex geometries could be modelled as the sum of simple shapes.

A comparison of early impact finite element models in terms of geometry, material characterization and boundary conditions is featured in Khalil & Viano [186]. Early finite element models relied on simple material and kinetic idealizations, viewing the head as an elastic shell, approximating the skull, filled with fluid. These models

assume homogeneous, isotropic and linearly (visco)-elastic material subject to small deformations. Moreover, many of the earliest models were 2D approximations of coronal [185] and mid-sagittal [347] sections, by virtue of the resolution achievable in comparison with a 3D model. Progressively more descriptive 3D models, most assuming some form of symmetry, appeared in the late 70's [410, 348, 156]. Impact models were characterized by a dynamic equation, typically neglecting damping, i.e.: $\mathbf{M}\ddot{\mathbf{a}} + \mathbf{K}\mathbf{a} = \mathbf{f}$ which was numerically integrated.

Research in the 80's and early 90's was reviewed by Sauren & Claessens [335]. Material properties were still assumed homogeneous and isotropic; linearly elastic constitutive models were used in general, in combination with small-deformation theory. Notable exceptions include Ueno and Mendis[253], who employed large-deformation theory. Mendis' characterization based on a large deformation assumption first appeared in his Ph.D. thesis, and was later published in [253]. Subsequently, King and his collaborators, in particular Ruan et al. [332] and Zhou et al. [425], described a comprehensive 3D approach which was highly detailed anatomically. The Zhou model, illustrated in figure A.1, emphasized details of gray and white matter and ventricles to match regions of high shear stress to locations of diffuse axonal injury.

Rheological Studies and Strain Models

Early investigations into the constitutive properties of brain tissue are attributed to McElhaney and his collaborators [101, 120]. Advani [300, 253] introduced more descriptive physical models based on a Mooney-Rivlin strain energy function. More recently, Miller and Chinzei have published studies characterizing brain constitutive properties under conditions approximating surgical loads.

Early reviews of rheological studies of animal and human brain tissue appear in Ommaya [293] and in Galford and McElhaney [120]. McElhaney and his collaborators [101, 120] performed extensive analyses of the stress-strain relation of human and monkey brain tissue, assuming a visco-elastic model. In [101], Estes and McElhaney noted that under compressive loading at rates v varying between 0.02 and 10 inches-per-second, the stress-strain curves were concave upward, suggesting that there

was no linear portion where a meaningful Young's modulus might be determined. A model of the form

$$\ln(\sigma/\dot{\epsilon}) = a + b \ln(t) \quad \text{where} \quad t = \frac{h-h_0}{v} , \quad (\text{A.20})$$

where h and h_0 are the instantaneous and original heights respectively of a given cylindrical sample, better accounted for the strain rate dependency. Galford and McElhaney [120] performed creep compliance tests with human and monkey brains, in order to characterize a 4-parameter model featuring Maxwell and Kelvin idealizations in series. The authors also performed tensile creep studies on scalp and dura samples.

Pamidi & Advani [300] modeled the visco-elastic behaviour of human brain tissue under a *large deformation* assumption, by viewing the constitutive properties in terms of a power function H encompassing inertial, restoring and dissipative forces:

$$\sigma_{ij} = \partial H / \partial \dot{\epsilon}_{ij} \quad \text{where} \quad H = \dot{U} + D \quad (\text{A.21})$$

where U is the familiar Mooney-Rivlin strain energy function in expression (A.17), D is the Rayleigh dissipation function of the material, and $\dot{\epsilon}_{ij}$ is a component of the *strain rate* tensor $\dot{\epsilon}$. This formulation led to two discrete spring-and-dashpot nonlinear characterizations, as well as a continuum model for an *isochoric* (volume-preserving) deformation.

Mendis, Stalnaker and Advani [253] first adopted a purely hyper-elastic model, again characterized by a first order Mooney-Rivlin strain energy function, and proposed a procedure for estimating the coefficients C_{01} and C_{10} in expression (A.17) for brain tissue, based on the uniaxial compression data of Estes and McElhaney [101]. Mendis then described a large-deformation finite element representation of the uniaxial soft tissue specimens used by Estes, and showed a comparison of the empirical stress values in the latter's compression experiment with the stress predicted by the hyper-elastic finite element model. Mendis also proposed a visco-elastic characterization of Estes' brain tissue samples based on a strain energy function *dependent on the time history* of the strain invariants, provided in expression (A.18):

$$U(t) = \int_0^t C_{10}(t-\zeta) \frac{d}{d\zeta} I_1(\zeta) + C_{01}(t-\zeta) \frac{d}{d\zeta} I_2(\zeta) d\zeta , \quad (\text{A.22})$$

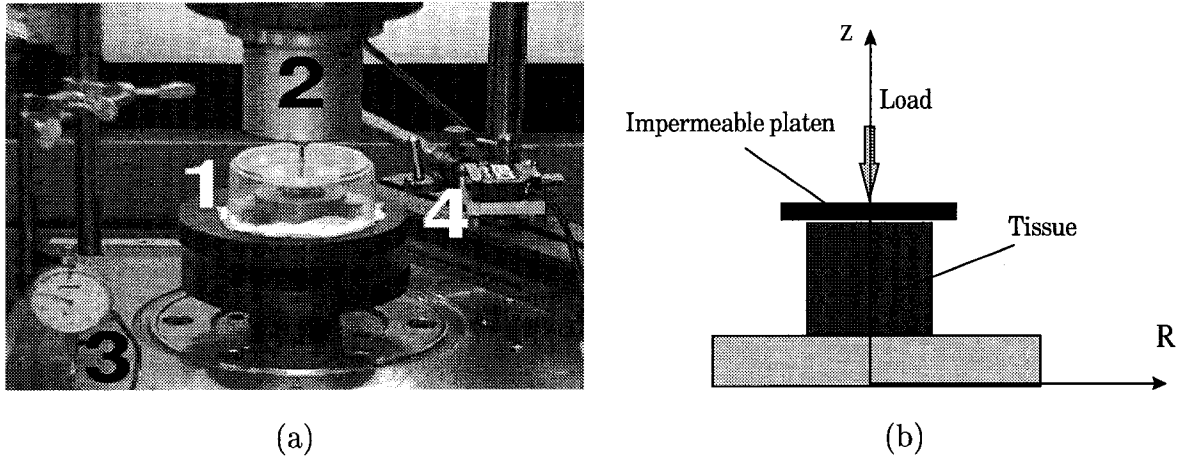


Figure A.3 : *Brain tissue rheological studies: (a) illustration of uniaxial compression apparatus; (b) layout with coordinate axes. Components: 1- specimen and loading platens, 2- load cell to measure axial force, 3- micrometer to measure axial displacement, and 4- laser to measure radial displacement. Courtesy of Karol Miller and Kiyoyuki Chinzei [262].*

in order to simulate experimental stress responses at four different strain rates.

Miller and Chinzei conducted rheological studies similar to those performed by Estes and McElhaney, but with much reduced loading velocities [262]. The former group pointed out that the strain rates investigated by McElhaney and his collaborators are relevant to injury modeling, but not as appropriate for characterizing the effects of surgical loads, particularly in view of their own finding that brain constitutive properties are strongly rate-dependent. They presented results of unconfined uniaxial compression tests of cylindrical brain tissue samples, based on the apparatus illustrated in figure A.3. This test was carried out under three different loading velocities: 500 mm/min, 5 mm/min, and 0.005 mm/min, corresponding to strain rates of about $0.64s^{-1}$, $0.64 \times 10^{-2}s^{-1}$, and $0.64 \times 10^{-5}s^{-1}$ respectively, and figure A.4 illustrates the rate dependency of brain constitutive properties.

Solid FE Models for Medical Applications

Both Ferrant [110] and Hagemann [137] have proposed small-strain linearly elastic FE models in research which dealt specifically with non-rigid registration of the human brain. Ferrant developed automatic image-based meshing algorithms for tomographic data, and applied his model to registering pre- and intra-operative MR volumes. He indicated that his registration method does not preclude the use of nonlinear constitutive properties. Hagemann validated his model by registering 2D pre- and post-operative images of the head of a patient.

Miller and Chinzei proposed a “hyper-viscoelastic” model, based on a generalization of the Mooney-Rivlin strain energy function expressed as an integral, in the same vein as Mendis’,

$$U(t) = \int_0^t \left\{ \sum_{i+j=1}^N C_{ij}(t-\zeta) \frac{d}{d\zeta} [(I_1 - 3)^i (I_2 - 3)^j] \right\} d\zeta, \quad (\text{A.23})$$

but they emphasized that a second-order characterization was necessary to fully capture the rate-dependent behaviour (i.e.: $N = 2$). Moreover, their strain-energy function is based on invariants of the left Cauchy-Green deformation tensor⁵. This continuum was subsequently used in a finite element implementation [266] using ABAQUS commercial software [153]. Lastly, in other publications [263, 265] they argued in favour of a purely solid continuum for modeling brain tissue, rather than a hybrid of solid and liquid, on the grounds that the latter does not account for stress-strain rate dependence as well as solid models can.

Finally, linear elastic models for tumour growth have been proposed [412, 201]. Kyriacou and Davatzikos [201] simulated the uniform contraction and expansion of a

⁵Note that in general for the strain-energy function to be *frame-indifferent*, or invariant under a change of reference, it must be a function of the *right stretch tensor* $\mathbf{U}$ or of the *right Cauchy-Green deformation tensor* $\mathbf{C} = \mathbf{F}^t \mathbf{F} = \mathbf{U}^2$, where $\mathbf{F}$ is the *deformation gradient tensor* [239]. If the material is further assumed isotropic, strain-energy can also be a function of the *left stretch tensor* $\mathbf{V}$ or of the *left Cauchy-Green deformation tensor* $\mathbf{B} = \mathbf{F} \mathbf{F}^t = \mathbf{V}^2$ in order to be frame-indifferent (see [398], pp. 307-314). The tensors $\mathbf{B}$ and $\mathbf{C}$ are not to be confused with the symbols used in the FE equilibrium equations.

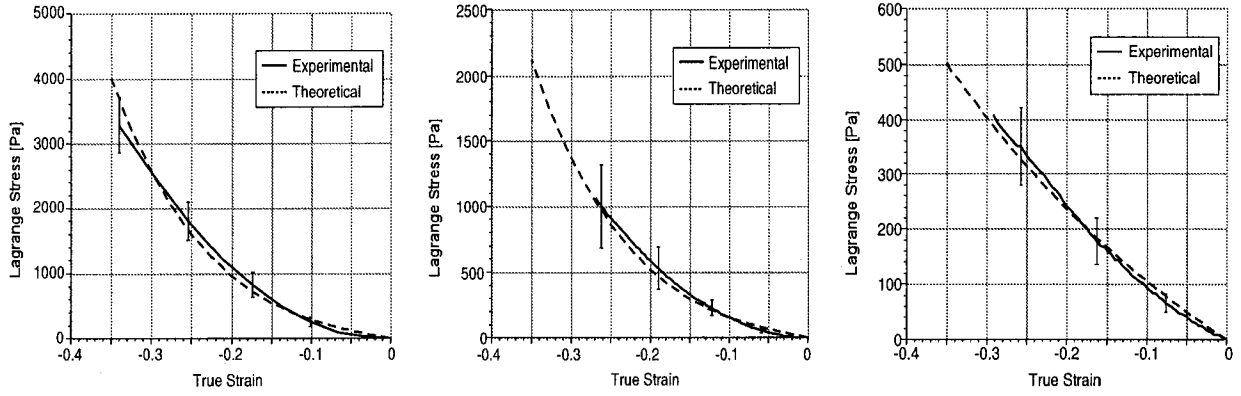


Figure A.4 : *Brain tissue rheological studies: illustration of uniaxial compression results vs. , under three different loading velocities: (a) 500 mm/min; (b) 5 mm/min; (c) 0.005 mm/min, courtesy of Karol Miller and Kiyoyuki Chinzei [262].*

tumour model obtained from MR image data, with a hyperelastic idealization. This application facilitates the application of a brain atlas to a subject with an imbedded lesion. Wasserman et al. incorporated a variety of pseudo-forces to account for biological and chemical, as well as mechanical, processes, contributing to tumour growth, in the context of a predictive clinical model [412].

A.3.2 Hybrid Solid-Liquid Brain Models

This section provides an overview of brain models consisting of both solid and liquid components. I first review literature that characterizes the physiology of the cranial cavity in a manner that accounts for its fluid component[138, 94]. This is followed by an overview of publications that integrate both components in a hybrid continuum, namely the *poro-elastic* [34, 258] and *mixture* models [45, 273, 4].

Biomechanics of the Cranial Cavity Featuring Solid and Fluid Components

Hakim et al. [138] proposed a detailed mechanical interpretation of intracranial anatomy, that accounted for both solid and fluid components, with an emphasis on describing the phenomenon of hydrocephalus. In particular, the brain parenchyma was described

as “*submicroscopic sponge of visco-elastic material*”. They completed the mechanical picture with a description of the linkage between brain and skull:

The brain does not rest directly on the inner surfaces of the skull, but is floating within the CSF (which is approximately the same density) and moored in position by the arachnoidal strands which tether the arachnoid membrane to the pia mater.

Hakim also evoked two parallel fluid compartments consisting of the CSF and extracellular spaces of the parenchyma, supplied by separate sources of blood (to the choroid plexuses and directly to the parenchymal tissue), and drained by the intracranial venous system. The CSF is secreted by the the choroid plexuses, flows from the lateral ventricles, through the foramen, aqueduct and subarachnoid spaces, and discharges into the venous system by way of the arachnoidal villi of the superior sagittal sinus. The interaction between the open venous system and the closed CSF system, in particular as it relates to CSF pressure, subdural stress, and ventricular size, was described by rectilinear and spherical models.

Dóczi [94] provided a recent survey of medical literature on volume regulation of brain tissue, that emphasized fluid distribution as well. In particular, starting from the assumption of incompressibility of the constituents of the skull (blood, CSF and brain tissue), which implied that their total volume must remain constant, he asked the question “*which components of the intracranial cavity could decrease in size, and to what extent compensate for a large extracerebral mass or for significantly enlarged (4- to 8-fold) hydrocephalic ventricles*”? Given that the cerebral blood volume and CSF correspond to 50ml and 100ml of the available space, he concluded that the brain itself must change in size (assuming no tissue is lost in the process), and described the factors involved in this process.

Solid-Liquid Continua: Poro-elastic and Mixture Continua, with Related FE Models

Biot [34] was the first to describe a three-dimensional continuum consisting of a porous solid, assumed linearly and isotropically elastic and under small strain, and containing water, assumed incompressible, in its pores. He suggested that such a *consolidation* model, whereby a poro-elastic medium containing an incompressible fluid gradually settling under load, could describe a wet sponge or water-saturated soil. The water in the pores is characterized by an instantaneous *variation in content* θ and an increment in pressure p . Biot modified the 3D Hookean solid model, as appears in the expression (A.4), to account for the fluid pressure term p , which after some manipulation [34, 258] can be stated simply as a function of the displacement $\mathbf{u} = [u_1 \ u_2 \ u_3]^T$:

$$G\nabla^2 \mathbf{u} + \frac{G}{1-2\nu} \nabla \epsilon - \alpha \nabla p = 0 \quad \text{where} \quad (A.24)$$

$$\nabla \equiv \left[\frac{\partial}{\partial x_1} \ \frac{\partial}{\partial x_2} \ \frac{\partial}{\partial x_3} \right]^T \quad \text{and} \quad \nabla^2 \equiv \sum_{i=1}^3 \frac{\partial^2}{\partial x_i^2} ,$$

and $\epsilon = \epsilon_1 + \epsilon_2 + \epsilon_3$ represents the volume increase of the continuum per unit initial volume. This expression represents a system of 3 equations in 4 unknowns, u_1 , u_2 , u_3 and p , which requires a fourth equation for its solution. The last equation is derived from *Darcy's law*, governing the flow of water in a porous medium, which states that the rate of flow is proportional to the gradient in water pressure:

$$\mathbf{v} = k \nabla p , \quad (A.25)$$

where k is the coefficient of permeability of the solid. This expression, together with the assumption of fluid incompressibility, leads to the fourth equation [34]:

$$k \nabla^2 p - \alpha \frac{\partial \epsilon}{\partial t} - \frac{1}{Q} \frac{\partial p}{\partial t} = 0 , \quad (A.26)$$

where Q is a measure of the amount of water that can be forced into the porous solid under pressure while the solid is kept constant.

Miga and his collaborators [261, 258] solved the weak forms of expressions (A.24) and (A.26), summed on elements comprising both animal and human brains, with

the Weighted Residuals FE method. Miga also rigourously investigated the stability of his FE numerical implementation of the consolidation model [260], and he has also applied his FE model to characterizing brain shift, on the basis of sparse displacement information [259]. The consolidation model has also been applied to modeling the pressure increase to brain edema and the biomechanics of hydrocephalus [383, 304].

A related but more general continuum, the mixture model, has been developed by Mow et al. [273] and by Bowen et al. [45], and applied to hydrated soft tissues by Spilker [372]. This model is characterized by each spatial point being simultaneously occupied to some degree by all the constituents comprising the mixture, where the a th body is assigned a *reference configuration* $\mathbf{x} = \chi_a(\mathbf{X}_a, t)$. The a th constituent is characterized by its *bulk density* $\rho_a(\mathbf{x}, t)$, where $\rho(\mathbf{x}, t) = \sum_{a=1}^N \rho_a(\mathbf{x}, t)$, representing the mass of a per unit volume of mixture, and by its *true density* $\gamma_a(\mathbf{x}, t)$ representing the mass of a per unit volume of a . The *volume fraction* of the a th constituent is given by:

$$\phi_a(\mathbf{x}, t) = \rho_a(\mathbf{x}, t) / \gamma_a(\mathbf{x}, t) \quad \text{where} \quad \sum_{a=1}^N \phi_a(\mathbf{x}, t) = 1. \quad (\text{A.27})$$

Bowen derived equations of balance of linear momentum, moment of momentum and energy for the mixture, in a manner which accounts for their diffusion and on the basis of equations characterizing individual constituents. He then described the special case of an incompressible elastic solid and $N - 1$ incompressible fluids. At the same time, Mow [273] developed the governing equations for a mixture consisting of a solid and a liquid phase, and applied them to characterizing cartilage tissue in accordance with creep and stress relaxation tests. For a system idealized as quasi-static, these equations are expressed as follows, where s and f indicate solid and fluid phases respectively:

$$\begin{aligned} \text{momentum:} & \quad \nabla \cdot \boldsymbol{\sigma}^\alpha + \Pi^\alpha = 0, & \alpha = s, f \\ \text{constitutive (s):} & \quad \boldsymbol{\sigma}^s = -\phi^s p \mathbf{I} + \lambda_s e^s \mathbf{I} + 2\mu^s \boldsymbol{\epsilon}^s \\ \text{constitutive (f):} & \quad \boldsymbol{\sigma}^f = -\phi^f p \mathbf{I} \\ \text{diffusive drag:} & \quad \Pi^s = -\Pi^f = K(\mathbf{v}^f - \mathbf{v}^s) \\ \text{continuity:} & \quad \nabla \cdot (\phi^f \mathbf{v}^f + \phi^s \mathbf{v}^s) = 0. \end{aligned} \quad (\text{A.28})$$

Here, ∇ is the gradient, σ is the stress tensor of either phase, Π represents the diffusive momentum between the two phases, ϕ^s and ϕ^f represent volume fractions or *solidity* and *porosity* respectively, $\mathbf{I}$ is the identity tensor, ϵ^s is the strain tensor of the solid phase, and $\mathbf{v}$ is the velocity vector. The scalar p is the apparent pressure, K is the diffusive drag coefficient, while the following scalars characterize the solid phase: e^s is the dilatation, while λ_s and μ_s are elastic constants. Needless to say, the simultaneous satisfaction of the system of equations (A.28) constitutes a formidable challenge, in terms of the expression of their corresponding weak form and their FE-based numerical solution [372, 4]. Zhu and Suh [427] have recently formulated a dynamic variant of this model for the subsequent application to brain impact studies.

A.4 Summary

This section proposed a literature review of the physical modeling of the brain, in particular as these publications relate to estimating its volumetric displacement field during surgery and simulating biomechanical response to virtual surgical tools. I reviewed relevant biomechanical concepts, in particular solid and liquid continua which are common in the literature, as well as the two leading approaches for numerical simulation, namely finite-element modeling and mass-spring networks. Finite element models of the brain were categorized foremost on the basis of the underlying continuum: solid and solid-liquid hybrid. The history of solid brain modeling was traced from impact models to models simulating surgical loads. The anatomical basis for a model accounting for solid and liquid components was presented, along with a discussion of the consolidation and mixture models.

Appendix B

Derivation of Surface Evolution Model

The geodesic model was originally derived by Caselles[57] and of Tannenbaum[187, 419] and their respective collaborators, and is again best explained for the lesser dimensional problem of identifying an anatomical contour within a 2D image. The foundation of the contour evolution model is the *Geometric Heat Equation*,

$$\frac{\partial \mathcal{C}}{\partial t} = \kappa \mathbf{n}, \quad (\text{B.1})$$

which describes the evolution of a family of curves so that their Euclidean perimeter shrinks as quickly as possible[119, 133]. More formally, let us define a family of curves $\mathcal{C} = \mathcal{C}(p, t)$, where t parametrizes the family and p the given curve ($0 \leq p \leq 1$), as well as the Euclidean length functional

$$L(t) \equiv \int_0^1 \left\| \frac{\partial \mathcal{C}}{\partial p} \right\| dp \equiv \int_0^1 ds \text{ where } ds = \|\mathcal{C}_p\| dp = (x_p^2 + y_p^2)^{1/2} dp. \quad (\text{B.2})$$

It can be shown[419], by taking the first variation in the previous expression,

$L'(t) \equiv - \int_0^{L(t)} \left\langle \frac{\partial \mathcal{C}}{\partial t}, \kappa \mathbf{n} \right\rangle ds$, that the direction in which L decreases most rapidly is given by expression (B.1). Furthermore, it can be shown that the equivalent expression for a level-set representation, $\{X(t) \in \mathbb{R}^2 : \Psi(X, t) = 0\}$ such that $\mathcal{C} \equiv X(t)$, is given by

$$\frac{\partial \Psi}{\partial t} = \|\nabla \Psi\| \text{div} \left(\frac{\nabla \Psi}{\|\nabla \Psi\|} \right). \quad (\text{B.3})$$

This curve evolution equation is the cornerstone of the classical level-set formulation proposed by Caselles et al. and Malladi et al., and featured in expression (5.10).

The geodesic model of expression (6.2) is obtained via a similar development, by replacing the metric based on Euclidean arc-length ds by a *conformal* metric which factors in image information,

$$ds_F = Fds = (x_p^2 + y_p^2)^{1/2} F dp. \quad (\text{B.4})$$

The justification for adopting a new metric is a unification of the curve evolution framework with the physical model framework, and which therefore replaces simple Euclidean curve shortening with a *gradient evolution equation*[187], whereby a contour evolves in the direction of the gradient of a potential energy function[57], $\mathcal{U}(C) \equiv -\lambda g(\|\nabla I(C)\|)^2$, itself a function of the image gradient at the points along the contour. Finally, setting $L_F \equiv \int_0^1 ds_F$ and taking its first variation leads to the fastest-shrinking direction being given by the expression:

$$\frac{\partial \mathcal{C}}{\partial t} = (F\kappa - \nabla F \cdot \mathbf{n}) \mathbf{n}, \quad (\text{B.5})$$

whose level-set formulation is:

$$\frac{\partial \Psi}{\partial t} = F \|\nabla \Psi\| \left[\text{div} \left(\frac{\nabla \Psi}{\|\nabla \Psi\|} \right) \right] + \nabla F \cdot \nabla \Psi. \quad (\text{B.6})$$

The surface evolution model in 3D proceeds from a similar analysis. The classical level-set surface model is based on *mean curvature surface evolution*, whereby the Euclidean area functional is defined[419]:

$$A(t) \equiv \int_0^1 \int_0^1 dS \text{ where } dS \equiv (\|\mathcal{S}_u\|^2 \|\mathcal{S}_v\|^2 - \langle \mathcal{S}_u, \mathcal{S}_v \rangle^2)^{1/2} dudv. \quad (\text{B.7})$$

By taking the first variation, the direction in which the *surface area* shrinks the most rapidly can be likewise obtained [164]:

$$\frac{\partial \mathcal{S}}{\partial t} = H \mathbf{n}. \quad (\text{B.8})$$

Just as in the 2D case, the Euclidean area functional can be replaced by a conformal area functional which depends on $F(x, y, z)$: $dS_F = FdS$, which leads to the

corresponding gradient evolution equation:

$$\frac{\partial \mathcal{S}}{\partial t} = (FH - \nabla F \cdot \mathbf{n}) \mathbf{n} , \quad (\text{B.9})$$

whose level-set formulation is the same as in expression (B.6). Here it is understood that $\{X(t) \in \mathbb{R}^3 : \Psi(X, t) = 0\}$ such that $\mathcal{S} \equiv X(t)$, and that $H = \text{div}(\nabla \Psi / \|\nabla \Psi\|)$ is the mean curvature of an isosurface of Ψ . Adding a constant inflation term ν to H , within the square brackets of expression (B.6), yields our surface identification model.

Bibliography

- [1] N. Abdelmalek, Algebraic error analysis for surface curvatures of 3-D range images obtained by different methods, *IEEE Proc. ICPR*, pp. 529-534, 1990.
- [2] D. Adelsteinsson and J.A. Sethian, A Fast Level Set Method for Propagating Interfaces, *J. Computational Physics*, No. 118, pp. 269-277, 1995.
- [3] M.S. Alp, M. Dujovny, M. Misra, F.T. Charbel, and J.I. Ausman, Head Registration Techniques for Image-guided Surgery, *Neurol. Res.*, Vol. 20, pp. 31-37, 1998.
- [4] E.S Almeida and R.L. Spilker, Mixed and Penalty Finite Element Models for the Nonlinear Behaviour of Biphasic Soft Tissues in Finite Deformations: Part I - Alternate Formulations, *Int. J. Comp. Methods in Biomech. and Biomed. Engrg*, Vol. 1, pp. 25-46, 1997.
- [5] L. Alvarez, P.L. Lions and J.-M. Morel, Image Selective Smoothing and Edge Detection by Nonlinear Diffusion. II, *SIAM J. Numerical Analysis*, Vol. 29, No. 3, pp. 845-866, June 1992.
- [6] A.A. Amini and J.S. Duncan, Differential Geometry for Characterizing 3D Shape Change, *Proc. SPIE 1768: Math. Methods in Med. Imag.*, pp. 170-181, 1992.
- [7] L.K. Arata and A.P. Dhawan, Iterative Principal Axes Registration: a New Algorithm for Retrospective Correlation of MR-PET Brain Images, *Proc. Ann. Conf. IEE Eng. Med. Biol. Soc.*, pp. 2776-2778, 1992.

- [8] K.S. Arun, T.S. Huang and S.D. Blostein, Least-squares Fitting of Two 3-D Point Sets, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 9, No. 5, pp. 698-700, May 1987.
- [9] O. Astley and V. Hayward, Real-Time Finite-Elements Simulation of General Visco-elastic Materials for Haptic Presentation, *IROS '97, IEEE/RJS Int. Conference on Intelligent Robots and Systems*, Sept. 1997.
- [10] O. Astley, *A Software Architecture for Surgical Simulation Using Haptics*, Ph.D. thesis, McGill University, 1999.
- [11] M.A. Audette, Correspondance stéréoscopique basée sur la structure de l'ombrage, or Shading-Based Stereoscopic Correspondence, *Master's Thesis*, Ecole Polytechnique, Aug. 1992.
- [12] M.A. Audette, P. Cohen and J. Weng, *Shading-Based Two-View Matching*, *Int. Joint Conf. Artificial Intelligence*, 1992.
- [13] M.A. Audette and K. Chinzei, Global Structure-preserving Voxel Classification with Limited Supervision for Highly Realistic, Patient-specific Surgical Simulation, *manuscript in preparation*.
- [14] M.A. Audette, F.P. Ferrie and T.M. Peters, An Algorithmic Overview of Surface Registration Techniques for Medical Imaging, *Medical Image Analysis*, Vol. 4, No. 3, pp. 201-217, Sept. 2000, selected for re-publication in IMIA Yearbook of Medical Informatics, Schattauer Publ. Co., 2002.
- [15] M.A. Audette and T.M. Peters, Level-set Surface Segmentation and Registration for Computing Intrasurgical Deformations, *SPIE Medical Imaging 1999*, San Diego, Feb. 20-26, 1999.
- [16] M.A. Audette, K. Siddiqi, and T.M. Peters, Level-set Surface Segmentation and Fast Cortical Range Image Tracking for Computing Intrasurgical Deformations, *Medical Image Computing and Computer-Assisted Intervention (MICCAI99)*, Cambridge, England, pp. 788-797, 1999.

- [17] M.A. Audette, M. Miga, J. Nemes, K. Chinzei, and T. Peters, A Survey of Biomechanical Modeling of the Brain for Intra-surgical Displacement Estimation and Medical Simulation, *manuscript in preparation*.
- [18] M.A. Audette, K. Siddiqi, F.P. Ferrie, and T. Peters, An Integrated Range-sensing, Segmentation and Registration Framework for the Characterization of Intra-surgical Brain Deformations in Image-Guided Surgery, submitted to *Comp. Vis. and Image Understand.- Special Issue on Non-rigid Registration*, 2002.
- [19] N. Ayache, *Artificial Vision for Mobile Robots: Stereo Vision and Multisensory Perception*, MIT Press, Cambridge, Mass., 1991.
- [20] R. Bajcsy and S. Kovacic, Multiresolution Elastic Matching, *Computer Vision, Graphics and Image Processing*, Vol. 46, pp. 1-21, 1989.
- [21] R. Bajcsy and F. Solina, Three Dimensional Object Representation Revisited, *Proc. IEEE Int. Conf. Comp. Vision*, pp. 231-240, 1987.
- [22] www.barebonesstore.com/products/default.asp?Prod=86.
- [23] A.H. Barr, Superquadrics and Angle-Preserving Transformations, *IEEE Computer Graphics Applications*, Vol. 1, pp. 11-23, 1981.
- [24] A.H. Barr, SGlobal and Local Deformations of Solid Primitives, *Comp. Graph.*, 18 (3), pp. 21-30, 1984.
- [25] C. Basdogan, Real-Time Simulation of Dynamically Deformable Finite Element Models Using Modal Analysis and Spectral Lanczos Decomposition Methods, *Proc. Medicine Meets Virtual Reality*, 2001.
- [26] K.-J. Bathe, *Finite Element Procedures in Engineering Analysis*, Prentice-Hall, 1982.
- [27] Beekley Corp., www.beekley.com.
- [28] *Belle Fontaine wig shop*, Montreal, Canada.

- [29] A.M. Bensaid, L.O. Hall, J.C. Bezdek, and L.P. Clarke, Partially Supervised Clustering for Image Segmentation, *Pattern Recognition*, Vol. 29, No. 5, pp. 859-871, 1996.
- [30] J. Berkley, S. Weghorst, H. Gladstone, G. Raugi, D. Berg, and M. Ganter, Banded Matrix Approach to Finite Element Modeling for Soft Tissue Simulation, *Virtual Reality: Research, Development, and Application*, Vol. 4, pp. 203-212, 1999.
- [31] J. Besag, Spatial Interaction and the Statistical Analysis of Lattice Systems, *J. Roy. Stat. Soc. B*, Vol. 36, 192-236, 1974
- [32] P.J. Besl and R.C. Jain, Invariant Surface Characteristics for 3D Object Recognition in Range Images, *Comp. Vis., Graph., and Image Proc.*, 33, pp. 33-80, 1986 .
- [33] P.J. Besl and N.D. McKay, A Method for Registration of 3-D Shapes, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.14, No.2, pp. 239-256, 1992.
- [34] M.A. Biot, General Theory of Three-Dimensional Consolidation, *J. Appl. Phys.*, Vol. 12, pp. 155-164, Feb. 1941.
- [35] F. Blais and J.A. Beraldin, Calibration of an Anamorphic Laser Based 3-D Range Sensor, *SPIE Proc. Videometrics V- Vol. 3174*, pp. 113-122, 1997.
- [36] P.M. Black, T. Moriarty, E. Alexander, P. Stieg, E.J. Woodard, P.L. Gleason, C.H. Martin, R. Kikinis, R.B. Schwartz, F.A. Jolesz, The Development and Implementation of Intraoperative Magnetic Resonance Imaging and its Neurosurgical Applications, *Neurosurg.*, Vol. 41, pp. 831-845, 1997.
- [37] F. Bloch, Nuclear Induction, *Phys. Rev.*, Vol. 70, pp. 460-474, Oct. 1946.
- [38] R.C. Bolles and P. Horaud, 3DPO: A Three-Dimensional Part Orientation System (1986), *Int. J. Rob. Res.* 5 (3), pp. 3-26, 1986.

- [39] F.L. Bookstein, Principal Warps: Thin-Plate Splines and the Decomposition of Deformations, *IEEE Trans. PAMI* 11 (6), pp. 567-585, 1989.
- [40] G. Borgefors, Distance Transformations in Arbitrary Dimensions, *CVGIP* 27, pp. 321-345, 1984.
- [41] G. Borgefors, Distance Transformations in Digital Images, *CVGIP* 34, pp. 344-371, 1986.
- [42] P.A. Bottomley, T. Foster, R. Argersinger, and L. Pfeifer, A Review of Normal Tissue Hydrogen NMR Relaxation Times and Relaxation Mechanisms from 1-100 MHz: Dependence on Tissue Type, NMR Frequency, Temperature, Species, Excision, and Age., *Med. Phys.*, Vol. 11, No. 4, pp. 425-448, July/Aug. 1984.
- [43] P.A. Bottomley, C. Hardy, R. Argersinger, and G. Allen-Moore , A Review of ^1H Nuclear Magnetic Resonance Relaxation in Pathology: Are T_1 and T_2 Diagnostic?, *Med. Phys.*, Vol. 14, No. 1, pp. 1-37, Jan/Feb. 1987.
- [44] S. Bouix and K. Siddiqi, Divergence-Based Medial Surfaces, *Proc. Euro. Conf. Comp. Vision*, 2000.
- [45] R.M. Bowen, Incompressible Porous Media Models by Use of the Theory of Mixtures, *Int. J. Engng. Sci*, Vol. 18, pp. 1129-1148, 1980.
- [46] BrainLab, www.brainlab.com.
- [47] M. Bro-Nielsen and S. Cotin, Real-time Volumetric Deformable Models for Surgery Simulation using Finite Elements and Condensation, *EUROGRAPHICS'96*, Vol. 15, No. 3, pp. 57-66.
- [48] R.A. Bronen and G. Sze, Magnetic Resonance Imaging Contrast Agents: Theory and Application to the Central Nervous System, *J. Neurosurg.*, Vol. 73, No. 6, pp.820-839, Dec. 1990.
- [49] L.G. Brown, A Survey of Image Registration Techniques, *ACM Comp. Surveys* 24 (4), pp. 326-376, 1992.

- [50] R.A. Brown, A Stereotactic Head Frame for Use with CT Body Scanners, *Invest. Radiol.*, No. 14, pp. 300-304.
- [51] G.A. Brown, M.C. Willis, K. Firoozbakhsh, A. Bamada, C.L. Tessman, and A. Montgomery, Computed Tomography Image-guided Surgery in Complex Acetabular Fractures, *Clin. Orthopaed. and Related Res.*, No. 370, pp. 219-236, Jan. 2000.
- [52] G.C. Burdea, *Force and Touch Feedback for Virtual Reality*, John Wiley and Sons, 1996.
- [53] R.J. Campbell and P.J. Flynn, Eigenshapes for 3D Object Recognition in Range Data, *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 1999.
- [54] J. Canny, *A Computational Approach to Edge Detection*, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. PAMI-8, No. 6, pp. 679-698, Nov. 1986.
- [55] M.B. Carpenter, *Core Text of Neuroanatomy*, 4th ed., Williams and Wilkins, Hong Kong, 1996.
- [56] V. Caselles, F. Catte, T. Coll, and F. Dibos, A Geometric Model for Active Contours in Image Processing, *Numerische Mathematik*, No. 66, 1993.
- [57] V. Caselles, R. Kimmel and G. Sapiro, Geodesic Active Contours, *Proc. IEEE Int. Conf. Computer Vision*, pp. 694-699, 1995.
- [58] V. Caselles, R. Kimmel, G. Sapiro, and C. Sbert, Minimal Surfaces: a Three Dimensional Segmentation Approach, *Technion EE Pub. No. 973*, 1995.
- [59] V. Caselles, JM Morel, G. Sapiro, and A. Tannenbaum, ed., Special Issue on Partial Differential Equations and Geometry-Driven Diffusion in Image Processing and Analysis, *IEEE Trans. Image Processing*, Vol. 7, No. 3, March 1998.
- [60] , F. Catté, PL Lions, JM Morel, and T. Coll, Image Selective Smoothing and Edge Detection by Nonlinear Diffusion, *SIAM J. Numerical Analysis*, Vol. 29, pp. 182-193, 1992.

- [61] H.H. Chen and T.S. Huang, Maximal Matching of 3-D Points for Multiple-Object Motion Estimation, *Patt. Rec.* 21 (2), pp. 75-90, 1988.
- [62] C.W. Chen, T.S. Huang, and M. Arrott, Modeling, Analysis, and Visualization of Left Ventricle Shape and Motion by Hierarchical Decomposition, *IEEE Trans. PAMI*, Vol. 16 (4), 342-356, 1994.
- [63] Y. Chen and G. Medioni, Object Modelling by Registration of Multiple Range Images, *IEEE Proc. Conf. Rob. and Auto.*, pp. 2724-2729, 1991.
- [64] F.A. Cheikh, R. Hamila, M. Gabbouj and J. Astola, Impulse Noise Removal in Highly Corrupted Color Images, *IEEE Int. Conf. Image Proc.*, 1996.
- [65] K.C. Chu and B.K. Rutt, Polyvinyl Alcohol Cryogel: An Ideal Phantom Material for MR Studies of Arterial Flow and Elasticity, *Magnetic Resonance in Medicine*, No. 37, pp. 314-319, 1997.
- [66] H.S. Choi, D.R. Haynor, and Y. Kim, Partial Volume Tissue Classification of Multi-channel Magnetic Resonance Images- A Mixel Model, *IEEE Trans. Med. Imag.*, Vol. 10, pp. 395-407, 1991.
- [67] L.P. Clarke , R.P. Velthuisen, M.A. Camacho, J.J. Heine, M. Vaidyanathan, L.O. Hall, R.W. Thatcher, and M.L. Silbiger, MRI Segmentation: Methods and Applications, *Magnetic Resonance Imaging*, Vol. 13, No. 3, pp. 343-368, 1995.
- [68] M.J. Clarkson, D. Rueckert, D.L.G. Hill, and D.J. Hawkes, Registration of Video Images to Tomographic Images by Optimising Mutual information., *Med. Image Comput. and Comput. Assis. Interv. - MICCAI*, pp. 579-588, 1999.
- [69] I. Cohen, N. Ayache, and P. Sulger, Tracking Points on Deformable Objects Using Curvature Information, *INRIA Tech. Report No. 1595*, 1992.
- [70] L.D. Cohen, On Active Contour Models and Balloons, *CVGIP: Image Under.*, Vol. 53, No. 2, pp. 211-218, 1991.

- [71] A. Collignon, D. Vandermeulen, P. Suetens, G. Marchal, Surface Based Registration of 3D Medical Images, *Proc. SPIE 1898: Med. Imag.*, pp. 32-42, 1993.
- [72] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Marchal, Automated Multimodality Image Registration Using Information Theory, *Inform. Proc. in Med. Imag.*, Kluwer, pp. 263-274, 1995.
- [73] D.L. Collins, P. Neelin, T.M. Peters and A.C. Evans, Automatic 3D Intersubject Registration of MR Volumetric Data in Standardized Talairach Space, *Journal of Computer Assisted Tomography*, Vol. 18, No. 2, pp. 192-205, 1994.
- [74] D.L. Collins, A.P. Zijdenbos, V. Kollokian, J.G. Sled, N.J. Kabani, C.J. Holmes and A.C. Evans, Design and Construction of a Realistic Digital Brain Phantom, *IEEE Trans. Medical Imaging*, Vol. 17, No. 3, pp. 463-468, June 1998.
- [75] D.L. Collins, personal communication.
- [76] R.M. Comeau, A.F. Sadikot, A. Fenster, T.M. Peters, Intraoperative Ultrasound for Guidance and Tissue Shift Correction in Image-guided Neurosurgery, *Medical Physics*, Vol. 27, No. 4, April 2000.
- [77] R.M. Comeau, *Intraoperative Ultrasound Imaging for the Detection and Correction of Tissue Movement in Image-guided Surgery*, Ph.D. Thesis, Dept. Biomedical Engineering, McGill University, Montreal, 2000.
- [78] T.F. Cootes, A. Hill, C.J. Taylor, and J. Haslam, The Use of Active Shape Models for Locating Structures in Medical Images, *Image and Vis. Comput.*, Vol. 12, No. 6, pp. 355-366, July 1994.
- [79] G. Coppini, Tensor Description of 3D Time Varying Surfaces Using Scattered Landmarks: an Application to Heart Motion, in *Time-Varying Image Processing and Moving Object Recognition*, ed. Cappellini, V., North-Holland, Amsterdam, pp. 158-163, 1987.

- [80] L. da F. Costa, Robust Skeletonization through Exact Euclidean Distance Transform and its Application to Neuromorphometry, *Real-Time Imaging*, Vol. 6, pp. 415-431, 2000.
- [81] S. Cotin, H. Delingette and N. Ayache, Efficient Linear Elastic Models of Soft Tissues for Real-time Surgery Simulation, *IEEE Trans. Visual. and Comput. Graph.*, Vol. 5, No. 1, pp. 62-73, Jan-Mar. 1999.
- [82] J.J. Craig, *Introduction to Robotics: Mechanics and Control*, Addison-Wesley, 1989.
- [83] H. Delingette, Simplex Meshes: a General Representation for 3D Shape Reconstruction, *Proc. IEEE Conf. Comp. Vis. and Patt. Rec.*, pp. 856-859, 1994.
- [84] H. Delingette, General Object Reconstruction Based on Simplex Meshes, *Int. J. Comp. Vis.*, Vol. 32, No. 2, pp. 111-146, 1999.
- [85] H. Delingette and J. Montagnat, New Algorithms for Controlling Active Contours Shape and Topology, European Conference on Computer Vision (ECCV'00), 2000.
- [86] H. Delingette and J. Montagnat, Shape and Topology Constraints on Parametric Active Contours *Comp. Vis. and Image Under.*, Vol. 83, pp. 140-171, 2001.
- [87] R. Deriche, Using Canny's Criteria to Derive a Recursively Implemented Optimal Edge Detector, *Int'l J. Computer Vision*, pp. 167-187, 1987.
- [88] R. Deriche, Fast Algorithms for Low-Level Vision, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 12, No. 1, pp. 78-87, Jan. 1990.
- [89] R. Deriche, Recursively Implementing the Gaussian and its Derivatives, *INRIA tech. report No. 1893*, 1993.
- [90] U.R. Dhond and J.K. Aggarwal, Structure from Stereo- A Review, *IEEE Trans. on Systems, Man, and Cybernetics*, Vol. 19, No. 6, Nov./Dec. 1989.

- [91] P. Dierckx, *Curve and Surface Fitting with Splines*, Clarendon, 1995, Oxford.
- [92] M.P. do Carmo, *Differential Geometry of Curves and Surfaces*, Prentice-Hall, Englewood Cliffs, N.J., 1976.
- [93] P.K. Doshi, L. Lemieux, D. Fish, S.D. Shorvon, W. Harkness, D.G.T. Thomas, Frameless Stereotaxy and Interactive Neurosurgery with the ISG Viewing Wand, *Acta Neurochir. Suppl.*, Vol. 64, pp. 49-53, 1995.
- [94] T. Dócsi, Volume Regulation of the Brain Tissue- a Survey, *Acta Neurochir.*, Vol. 121, pp. 1-8, 1993.
- [95] J.M. Drake, J. Prudencio and S. Holowka, A Comparison of a PUMA Robotic System and the ISG Viewing Wand for Neurosurgery, *Interactive Image-guided Neurosurgery*, ed. R.J. Maciunas, American Association of Neurological Surgeons, pp. 121-133, 1993.
- [96] J. Duchon, Interpolation des fonctions de deux variables suivant le principe de la flexion des plaques minces, *R.A.I.R.O Analyse Numérique* 10, pp. 5-12, 1976.
- [97] R.O. Duda and P.E. Hart, *Pattern Analysis and Scene Analysis*, Wiley Interscience, 1973. T
- [98] G. Edner, Stereotactic Biopsy of Intracranial Space-occupying Lesions, *Acta Neurochir.*, TNo. 33, pp. 213-214, 1981.
- [99] P.J. Edwards, D.L.G. Hill, J.A. Little, D.J. Hawkes, Deformation for Image Guided Interventions Using a Three Component Tissue Model, *Proc. Inform. Proc. Med. Imag.- IPMI*, pp. 218-231, 1997.
- [100] I. Essa, S. Sclaroff and A. Pentland, Physically Based Modelling for Graphics and Vision, *Directions in Geometric Computing*, Information Geometers, pp. 161-195, 1993.
- [101] M.S. Estes and J.H. McElhaney, Response of Brain Tissue of Compressive Loading, *ASME report 70-BHF-13*, 1970.

- [102] A.C. Evans, W. Dai, D.L. Collins, P. Neelin, and T.S. Marrett, Warping of a Computerized 3-D Atlas to Match Brain Image Volumes for Quantitative Neuroanatomical and Functional Analysis, *Proc. SPIE 1445 Med. Imag. V: Image Proc.*, pp. 236-246, 1991.
- [103] H.R. Everett, *Sensors for Mobile Robots : Theory and Application*, A.K. Peters, 1995.
- [104] T.L. Faber and E.M. Stokely, Orientation of 3D Structures in Medical Images, *IEEE Trans. Patt. Anal. and Mach. Intell.*, Vol. 10, No. 5, pp. 626-633, Sept. 1988.
- [105] T.J. Fan, G. Medioni, and R. Nevatia, Recognizing 3-D Objects Using Surface Descriptions, *IEEE Trans. PAMI* 11 (11), pp. 1140-1157, 1989.
- [106] O.D. Faugeras and M. Hebert, The Representation, Recognition, and Locating of 3-D Objects, *Int. J. Robotics Research*, Vol.5, No.3, pp. 27-52, Fall 1986.
- [107] J. Feldmar and N. Ayache, Rigid and Affine Registration of Free-Form Surfaces, using Differential Properties, *Proc. Euro. Conf. Comp. Vision*, pp. 397-406, 1994.
- [108] J. Feldmar and N. Ayache, Locally Affine Registration of Free-form Surfaces, *IEEE Proc. Conf. Comp. Vision and Patt. Rec.*, pp. 496-501, 1994.
- [109] J. Feldmar and N. Ayache, Rigid, Affine and Locally Affine Registration of Smooth Surfaces, *INRIA Technical Report No. 2220*, 1994.
- [110] M. Ferrant, S.K. Warfield, A. Nabavi, F.A. Jolesz, and R. Kikinis, Registration of 3D Intraoperative MR Images of the Brain Using a Finite Element Biomechanical Model, *Med. Image Comput. and Comput.-Assist. Interv.- MICCAI*, pp. 19-27, 2000.
- [111] F.P. Ferrie, S. Mathur, and G. Soucy, Feature Extraction for 3-D Model Building and Object Recognition, *Three-Dimensional Object Recognition Systems*, ed. Jain, A.K., Flynn, P.J., Elsevier, pp. 57-88, Amsterdam, 1993.

- [112] K.W. Finnis, Y.P. Starreveld, A.G. Parrent, and T.M. Peters, A 3-Dimensional Database of Deep Brain Functional Anatomy, and Its Application to Image-Guided Neurosurgery, *Med. Image Comput. and Comput.-Assist. Interv.- MIC-CAI*, pp. 1-8, 2000.
- [113] J.M. Fitzpatrick, J.B. West and C.R. Maurer, Predicting Error in Rigid-body, Point-based Registration, *Trans. IEEE Med. Imag.*, pp. 694-702, Vol. 17, No. 5, Oct. 1998.
- [114] P.J. Flynn and A.K. Jain, On Reliable Curvature Estimation, *IEEE Proc. Conf. Comp. Vision and Patt. Rec.*, pp. 110-116, 1989.
- [115] J.D. Foley, A. van Dam, S.K. Feiner, and J.F. Hughes, *Computer Graphics-Principles and Practice*, Addison-Wesley, Reading, Mass., 1990. T
- [116] K.T. Foley and M.M. Smith, Image-guided Spine Surgery, *Neurosurg. Clin. North Amer.*, TVol. 7, No. 2, pp.171-186, Apr. 1996.
- [117] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, 2nd ed., Academic Press, 1990.
- [118] Y.C. Fung, *Biomechanics : Mechanical Properties of Living Tissues*, 2nd ed., Springer-Verlag, 1993.
- [119] M. Gage and R.S. Hamilton, The Heat Equation shrinking Convex Plane Curves, *J. Differential Geometry*, Vol. 23, pp. 69-96, 1986.
- [120] J.E. Galford and J.H. McElhaney, A Viscoelastic Study of Scalp, Brain, and Dura, *J. Biomechanics*, Vol. 3, pp. 211-221, 1970.
- [121] R.L. Galloway, R.J. Maciunas and J.W. Latimer, The Accuracies of Four Stereotactic Frame Systems: An Independent Assessment, *Biomed. Instr. and Tech*, pp. 457-460, Nov./Dec. 1991.

- [122] R.L. Galloway, R.J. Maciunas and C.A. Edwards II, Interactive Image-Guided Surgery, *IEEE Trans. Biomedical Imaging*, Vol. 39, No. 12, pp. 1226-1231, Dec. 1992.
- [123] Y. Ge, C.R. Maurer Jr. and J.M. Fitzpatrick, Surface-based 3-D Image Registration Using the Iterative Closest Point Algorithm with a Closest Point Transform, *Medical Imaging 1996: Image Processing, Proc. SPIE 2710*, pp. 358-367, 1996.
- [124] GE Medical Systems Signa SP and IntraOperative MR Imaging, <http://www.gemedicalsystems.com/rad/mri/products/spi/index.html>.
- [125] G. Gerig G. Gerig, O. Kbler, R. Kikinis and F.A. Jolesz, Nonlinear Anisotropic Filtering of MRI Data, *IEEE Trans. Med. Imaging*, Vol. 11, No. 2, pp. 221-232, June 1992.
- [126] D.T. Gering, A. Nabavi, R. Kikinis, N. Hata, L. O'Donnell, W. Grimson, F. Jolesz, P. Black, and W. Wells III, An Integrated Visualization System for Surgical Planning and Guidance Using Image Fusion and Interventional Imaging, *Medical Image Computing and Computer-Assisted Intervention- MICCAI'99*, pp. 809-819, 1999.
- [127] S.F.F. Gibson, 3D ChainMail: a Fast Algorithm for Deforming Volumetric Objects, *Proc. Symp. Interactive 3D Graphics- ACM SIGGRAPH*, pp. 149-154, 1997.
- [128] P.L. Gildenberg, The History of Stereotactic Neurosurgery, *Neurosurg. Clin. of North Amer.*, Vol. 1, No. 4, Oct. 1990.
- [129] C. Giorgi, Intraoperative Fusion of Field Images with CT/MRI Data by Means of a Stereotactic Mechanical Arm, *Minim. Invas. Neurosurg.*, Vol. 37, pp. 53-55, 1994.
- [130] D.C. Gobbi, B.C.K. Lee, and T.M. Peters, Real-time 3D Ultrasound for Intraoperative Surgical Guidance, *Proc. SPIE- Medical Imaging, Vol. 4319*, pp. 264-271, 2001.

- [131] D.B. Goldgof, H. Lee, and T.S. Huang, Motion Analysis of Nonrigid Surfaces, *Proc. Conf. Comp. Vision and Patt. Rec.*, pp. 375-380, 1988.
- [132] D.B. Goldgof, H. Lee, and T.S. Huang, Feature Extraction and Terrain Matching, *Proc. Conf. Comp. Vision and Patt. Rec.*, pp. 899-904, 1988.
- [133] M. Grayson, The Heat Equation Shrinks Embedded Plane Curves to Round Point, *J. Diff. Geom.*, Vol. 26, pp. 285-314, 1987.
- [134] A.E. Green and W. Zerna, *Theoretical Elasticity*, 2nd ed., Clarendon Press, 1968.
- [135] H. Gudbjartsson and S. Patz, The Rician Distribution of Noise MRI Data, *Mag. Res. in Med.*, Vol. 34, p. 910-914, 1995.
- [136] A. Guézic and N. Ayache, Smoothing and Matching of 3-D Space Curves, *Int. J. Comp. Vision* 12 (1), 79-104, 1994
- [137] A. Hagemann, K. Rohr, H.S. Stiehl, U. Spetzger, and J.M. Gilsbach., Nonrigid Matching of Tomographic Images Based on a Biomechanical Model of the Human Head, *Proc. SPIE - Med. Imag.: Image Proc.*, 1999.
- [138] S. Hakim, J.G. Venegas and J.D. Burton, The Physics of the Cranial Cavity, Hydrocephalus and Normal Pressure Hydrocephalus: Mechanical Interpretation and Mathematical Model, *Surg. Neurol.*, Vol. 5, March 1976.
- [139] K.V. Hansen and O.V. Larsen, Using Region-of-Interest Based Finite Element Modelling for Brain Surgery Simulation, *Medical Image Computing and Computer-Assisted Intervention- MICCAI'98*, pp. 305-316, 1998.
- [140] R. Hartley, A Gaussian-Weighted Multiresolution Edge Detector, *Computer Vision, Graphics, Image Processing*, Vol. 30, pp. 70-83, 1985.
- [141] V. Hayward, P. Gregorio, O. Astley, S. Greenish, M. Doyon, L. Lessard, J. McDougall, I. Sinclair, S. Boelen, X. Chen, J.-P. Demers, J. Poulin, I. Benguigui, N. Almey, B. Makuc, and X. Zhang, Freedom-7: A High Fidelity Seven Axis Haptic

Device With Application To Surgical Training, *ISER-6 International Symposium on Experimental Robotics*, June 1997.

- [142] M. Hebert, C. Caillas, E. Krotkov, I. Kweon, and T. Kanade, Terrain Mapping for a Roving Planetary Explorer, *IEEE Proc. Int. Conf. Rob. and Auto.*, pp. 997-1002, 1989.
- [143] M.P. Heilbrun, P. McDonald, C. Wiker, S. Koehler, W. Peters Y. Sieminov, and C. Wiker, Implementation of a Machine Vision Method for Stereotactic Localization and Guidance, *Interactive Image-guided Neurosurgery*, ed. R.J. Maciunas, American Association of Neurological Surgeons, pp. 169-177, 1993.
- [144] K. Held, E. Rota Kops, B.J. Krause, W.M. Wells III, R. Kikinis, and H.W. Miller-Grtner, Markov Random Field Segmentation of Brain MR Images, *IEEE Trans. Med. Imag.*, Vol. 16, pp. 878-887, 1998.
- [145] R.M. Henkelman, Measurement of Signal Intensities in the Presence of Noise in MR Images, *Med. Phys.*, Vol. 12, No. 2, pp. 232-233, Mar/Apr 1985.
- [146] R.M. Henkelman, Erratum: Measurement of Signal Intensities in the Presence of Noise in MR Images, *Med. Phys.*, Vol. 13, No. 4, p. 544, Jul./Aug. 1986.
- [147] R.M. Henkelman and M.J. Bronskill, General Artifacts and Artifact Reduction Techniques, *The Physics of MRI- 1992 AAPM Summer School Proceedings*, ed. M.J. Bronskill and P. Sprawls, pub. American Inst. Physics, pp. 288-305, 1992.
- [148] C.J. Henri, D.L. Collins, and T.M. Peters, Multimodality Image Integration for Stereotactic Surgical Planning, *Medical Physics* 18 (2), 167-177, 1991.
- [149] C.J. Henri, G.B. Pike, D.L. Collins, and T.M. Peters, Three-dimensional Display of Cortical Anatomy and Vasculature: Magnetic Resonance Angiography Versus Multimodality Integration, *J. Digital Imaging*, Vol. 4, No. 1, Feb. 1991.
- [150] A.J. Herline, J.D. Stefansic, J.P. Debelak, S. Hartman, C.W. Pinson, R.L. Galloway, and W.C. Chapman, Image-guided Surgery: Preliminary Feasibility Stud-

- ies of Frameless Stereotactic Liver Surgery, *Archives Surg.*, Vol. 134, no. 6, pp. 644-649, June 1999.
- [151] G.T. Herman and H.K. Liu, Three-dimensional Display of Human Organs from Computed Tomograms, *Comp. Graph. and Image Proc.* 9 (1), Jan., pp. 1-21, 1979.
 - [152] M.F. Heuckel, An Operator which Locates Edges in Digitized Pictures, *J. Assoc. Computing Machinery*, vol. 20, pp. 113-125, 1971.
 - [153] Hibbitt, Karlsson and Sorensen, Inc., *ABAQUS Theory Manual*, 1995.
 - [154] D.L.G. Hill, CR Maurer, Jr., MY Wang, RJ Maciunas, J. A Barwise, and JM Fitzpatrick, Estimation of Intraoperative Brain Surface Movement, *Proc. CVRMed-MRCAS*, pp. 449-458, 1997.
 - [155] C.J. Holmes, R. Hoge, DL Collins, R. Woods, A. Toga, and AC Evans, Enhancement of MR images using Registration for Signal Averaging., *J. Comp. Asst. Tomog.*, Vol. 22, pp. 324-333, Mar 1998.
 - [156] R.R. Hosey and Y.K. Liu, A Homeomorphic Finite Element Model of Impact Head and Neck Injury, *Int. Conf. Proc. Finite Elements in Biomechanics*, Vol. 2, B.R. Simon ed., pp. 851-871, 1980.
 - [157] V. Horsley and R.H. Clarke, The Structure and Functions of the Cerebellum Examined by a New Method, *Brain*, No. 31, pp. 45-124.
 - [158] B.K.P. Horn, Closed-form Solution of Absolute Orientation Using Unit Quaternions, *J. Optical Society of America A*, Vol. 4, No. 4, pp. 629-642, April 1987.
 - [159] B.K.P. Horn and B.G. Schunck, B.G., Determining Optical Flow, *Art. Intel.* 17, pp. 185-203, 1981.
 - [160] D.I. Hoult and P.C. Lauterbaaur, The Sensitivity of the Zeugmatographic Experiment Involving Human Samples, *J. Magnetic Resonance*, No. 34, pp. 425-433, 1979.

- [161] T.S. Huang, S.D. Blostein, and E.A. Margerum, Least-Squares Estimation of Motion Parameters from 3-D Point Correspondences, *IEEE Proc. Conf. Comp. Vision and Patt. Rec.*, pp. 198-201, 1986.
- [162] M. Hueckel, An Operator Which Locates Edges in Digitized Pictures, *J. Assoc. Computing. Machinery*, Vol. 18, pp. 113-125, 1971.
- [163] R.A. Hummel, Feature Detection Using Basis Functions, *Computer Graphics and Image Proc.*, Vol. 9, pp. 40-55, 1979.
- [164] G. Huisken, Flow by Mean Curvature of Convex Surfaces into Spheres, *J. Diff. Geom.*, Vol. 20, pp. 236-266, 1984.
- [165] R.A. Hummel, Representations Based on Zero-Crossings in Scale-Space, *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 204-209, 1986.
- [166] Image Guided Technologies Inc., www.imageguided.com.
- [167] Intelligent Actuator (IAI), www.intelligentactuator.com.
- [168] ISG Technologies Inc., *Viewing Wand Image-Guided Surgical System- Operator's Guide*, 1997.
- [169] IZI Medical Products, www.izimed.com.
- [170] J.J. Jacq and C. Roux, Automatic Registration of 3D Images Using a Simple Genetic Algorithm with a Stochastic Performance Function, *IEEE Proc. Eng. Med. and Biol. Soc.*, pp. 126-127, 1993.
- [171] F. John, *Partial Differential Equations*, 4th edition, Springer-Verlag, 1982.
- [172] A.E. Johnson and M. Hebert, Surface Registration by Matching Oriented Points, *Int. Conf. Recent Advances in 3-D Digital Imaging and Modeling*, 1997.
- [173] A.E. Johnson and M. Hebert, Surface Matching for Object Recognition in Complex Three-dimensional Scenes, *Image and Vision Computing* 16, pp. 635-651, 1998.

- [174] L. Junck, J.G. Moen, G.D. Hutchins, M.B. Brown, and D.E. Kuhl, Correlation Methods for the Centering, Rotation, and Alignment of Functional Brain Images, *J. Nucl. Med.*, Vol. 31, pp. 1220-1276.
- [175] R.E. Kalman, A New Approach to Linear Filtering and Prediction Problems, *Trans. AMSE, J. Basic Eng.* 82D, pp. 35-45, 1960.
- [176] J. Kaminsky, T. Brinker, A. Samii, G. Arango, P. Vorkapic, and M. Samii, Technical Considerations Regarding Accuracy of the MKM Navigation System. An Experimental Study of Impact Factors, *Neurol. Res.*, Vol. 21, June 1999.
- [177] B. Kamgar-Parsi, J.L. Jones, and A. Rosenfeld, Registration of Multiple Overlapping Range Images: Scenes without Distinctive Features, *IEEE Trans. PAMI* 13 (9), pp. 857-871, 1991.
- [178] E.J. Kandel and Y.V. Scavinsky, Stereotaxic Apparatus and Operations in Russia in the 19th Century, *J. Neurosurg.*, No. 37, pp. 407-411.
- [179] M. Kass, A. Witkin and D. Terzopoulos, Snakes: Active Contour Models, *Int'l. J. Computer Vision*, pp. 321-331, 1987.
- [180] N. Kehtarnavaz and S. Mohan, A Framework for Estimation of Motion Parameters from Range Images, *CVGIP* 45, pp. 88-105, 1989.
- [181] P.J. Kelly, B.A. Kall, S. Goerss, and F. Earnest, Computer-assisted Stereotaxic Laser Resection of Intra-axial Brain Neoplasms, *J. Neurosurg.*, No. 64, pp. 427-439.
- [182] P.J. Kelly, B.A. Kall and S.J. Goerss, Results of Computed Tomography-based Computer-assisted Stereotactic Resection of Metastatic Intracranial Tumors, *Neurosurgery*, No. 22, pp. 7-17.
- [183] P.J. Kelly, Stereotactic Excision of Brain Tumours, *Stereotactic and Directed Surgery of Brain Tumours*, ed. D.G.T. Thomas, Churchill Livingstone, pp. 89-109, 1993.

- [184] R. Khadem, C.C. Yeh, M. Sadeghi-Tehrani, M.R. Bax, J.A. Johnson, J.N. Welch, E.P. Wilkinson, and R. Shahidi, Comparative Tracking Error Analysis of Five Different Optical Tracking Systems, *Comput. Aided Surg.*, Vol. 5, pp. 98-107.
- [185] T.T. Khalil and R.P. Hubbard, Parametric Study of Head Response by Finite Element Modelling, *J. Biomechanics*, Vol. 10, pp. 119-132, 1977.
- [186] T.B. Khalil and D.C. Viano, Critical Issues in Finite Element Modelling of Head Impact, *Proc. 26th Stapp Car Crash Conf.*, SAE Paper 821150, pp. 87-101, 1982.
- [187] S. Kichenassamy, A. Kumar, P.J. Olver, A. Tannenbaum, and A. Yezzi, Gradient Flows and Geometric Active Contour Models, *Proc. IEEE Int'l. Conf. Computer Vision*, pp. 810-815, 1995.
- [188] R. Kikinis, P.L. Gleason, W. Lorensen, W.M. Wells, W.E.L. Grimson, T. Lozano-Perez, G.J. Ettinger, S.J. White, and F. Jolesz, Image Guidance Techniques for Neurosurgery, *Proc. SPIE 2359: Visualization in Biomedical Computing*, pp. 537-540, 1994.
- [189] B.B. Kimia and K. Siddiqi, Geometric Heat Equation and Nonlinear Diffusion of Shapes and Images, *Computer Vision and Image Understanding*, Vol. 64, No. 3, pp. 305-322, Nov. 1996.
- [190] B.B. Kimia, A.R. Tannenbaum and S.W. Zucker, Shapes, Shocks, and Deformations I: The Components of Two-Dimensional Shape and the Reaction-Diffusion Space, *Int'l J. Computer Vision*, Vol. 15, No. 3, pp. 189-224, 1995.
- [191] R. Kimmel, N. Kiryati, and A.M. Bruckstein. Sub-pixel Distance Maps and Weighted Distance Transforms, *J. Mathematical Imaging and Vision*, No. 6, pp. 223-233, 1996.
- [192] R. Kimmel, N. Kiryati, and A.M. Bruckstein. Analyzing and Synthesizing Images by Evolving Curves with the Osher-Sethian Method, *Int. J. Comp. Vision* 24 (1), pp. 37-56, 1997.

- [193] A.I. King and C.C. Chou, Mathematical Modelling, Simulation and Experimental Testing of Biomechanical System Crash Response, *J. Biomechanics*, no. 9, pp. 301-317, 1976.
- [194] E. Kishon, T. Hastie, and H. Wolfson, 3-D Curve Matching Using Splines, *Proc. Euro. Conf. Comp. Vision*, pp. 589-591, 1990.
- [195] J.J. Koenderink, The Structure of Images, *Biological Cybernetics*, pp. 363-370, 1984.
- [196] V. Kollokian, *Performance Analysis of Automatic Techniques for Tissue Classification in Magnetic Resonance Images of the Human Brain*, Master's Thesis, Dept. Computer Science, Concordia University, Nov. 1996.
- [197] D. Kondziolka and L.D. Lunsford, Intraoperative Computed Tomographic Localization, R.J. Maciunas, ed., *Interactive Image-Guided Neurosurgery*, American Association of Neurological Surgeons, pp. 203-216, 1993.
- [198] R.K.-S. Kwan, A.C. Evans and G.B. Pike, An Extensible MRI Simulator for Post-Processing Evaluation, *Proc. Visualization in Biomedical Computing, 4th Int. Conf. VBC'96*, pp. 135-140, Springer-Verlag, 1996.
- [199] R.K.-S. Kwan, *An Extensible MRI Simulator for Quantitative Evaluation of Image-Processing and Classification Methods*, Master's Thesis, Dept. Biomedical Engineering, McGill University, 1998.
- [200] R.K.-S. Kwan, A.C. Evans and G.B. Pike, MRI Simulation-Based Evaluation of Image-Processing and Classification Methods, *IEEE Trans. Medical Image Analysis*, Vol. 18, No. 11, Nov. 1999.
- [201] S.K. Kyriacou and C. Davatzikos, A Biomechanical Model of Soft Tissue Deformation with Applications to Non-rigid Registration of Brain Images with Tumor Pathology, *Medical Image Computing and Computer-Assisted Intervention-MICCAI'98*, pp. 531-538, 1998.

- [202] R.S. Lakes, *Viscoelastic Solids*, CRC Press, 1999.
- [203] Y. Lamdan and H.J. Wolfson, Geometric Hashing: A General and Efficient Model-Based Recognition Scheme, *IEEE Proc. Int. Conf. Comp. Vision*, pp. 238-249, 1988.
- [204] L.D. Landau and E.M. Lifshitz, *Theory of Elasticity*, 3rd ed., Course of Theoretical Physics Vol. 7, Pergamon Press, 1986.
- [205] S.P. Langron and A.J. Collins, Pertubation Theory for Generalized Procrustes Analysis, *J. Royal Stat. Soc. B*, Vol. 47, pp. 277-284, 1985.
- [206] L. Lapidus, *Numerical Solution of Partial Differential Equations in Science and Engineering*, Wiley, 1982.
- [207] Laser Institute of America, www.laserinstitute.org/corporate/industsys.html, 1998.
- [208] S. Lavallée, Registration for Computer-Integrated Surgery: Methodology, State of the Art, *Computer-Integrated Surgery - Technology and Clinical Applications*, ed. R.H. Taylor, MIT Press, Cambridge, Mass., pp. 77-97, 1996.
- [209] S. Lavallée and R. Szeliski, Recovering the Position and Orientation of Free-form Objects from Image Contours Using 3D Distance Maps, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 17, no. 4, pp. 378-390, 1995.
- [210] S. Lavallée, R. Szeliski, and L. Brunie, Anatomy-Based Registration of Three-Dimensional Medical Images, Range Images, X-Ray Projections, and Three-Dimensional Models Using Octree-Splines, *Computer-Integrated Surgery - Technology and Clinical Applications*, ed. R.H. Taylor, MIT Press, Cambridge, Mass., pp. 115-143, 1996.
- [211] Stryker Leibinger, www.leibinger.net.

- [212] F. Leitner and P. Cinquin, From splines and snakes to snake splines. *Geometric Reasoning for Perception and Action - Workshop*, Springer-Verlag, pp. 264-281, 1993.
- [213] L. Lemieux, C.J. Henri, R. Winston, D.L. Collins and T.M. Peters, The Mathematics of Stereotactic Localization, *Stereotactic and Image Directed Surgery of Brain Tumours*, ed. by D.G.T. Thomas, Churchill Livingstone, pp. 193-216, 1993.
- [214] L. Leksell, A Stereotaxic Apparatus for Intracerebral Surgery, *Acta Chir. Scand.*, Vol. 99, pp. 229-233.
- [215] L. Leksell, The Stereotaxic Method and Radiosurgery of the Brain, *Acta Chir. Scand.*, Vol. 102, pp. 316-319.
- [216] L. Leksell and B. Jernberg, Stereotaxis and Tomography- A Technical Note, *Acta Neurochirurg.*, No. 52, pp. 1-7, 1980.
- [217] D. Lemoine, C. Barillot, B. Gibaud, and E. Pasqualini, An Anatomical-Based 3D Registration System of Multimodality and Atlas Data in Neurosurgery, *Proc. Conf. Inform. Proc. in Med. Imag.*, pp. 154-164, 1991.
- [218] R.J. Leveque, *Numerical Methods for Conservation Laws*, 2nd ed., Birkhäuser Verlag, 1992.
- [219] S.Z. Li, *Markov Random Field Modeling in Computer Vision*, Springer-Verlag, 1995.
- [220] Z.C. Lin, H. Lee, and T.S. Huang, A Frequency-Domain Algorithm for Determining Motion of a Rigid Object from Range Data Without Correspondences, *IEEE Proc. Conf. Comp. Vision and Patt. Rec.*, pp. 194-198, 1986.
- [221] C. Lindquist and L. Steiner, Stereotactic Radiosurgical Treatment of Arteriovenous Malformations of the Brain, *Modern Stereotactic Neurosurgery*, ed. L.D. Lunsford, Boston, Martinus Nijhoff, pp. 491-505, 1988.

- [222] P. Lipson, A. Yuille, DO Keefe, J. Cavanaugh, J. Taffe, and D. Rosenthal, Deformable Templates for Feature Extraction from Medical Images, *Proc. Europ. Conf. Comput. Vis.- ECCV*, pp. 413-417, 1990.
- [223] C. Ljung, A Model for Brain Deformation Due to Rotation of the Skull, *J. Biomechanics*, No. 8, pp. 263-274, 1975.
- [224] W.E. Lorensen and H.E. Cline, Marching Cubes: A High Resolution 3D Surface Construction Algorithm, *Computer Graphics*, Vol. 21, No. 3, pp. 163-169, July 1987.
- [225] D.G. Luenberger, *Linear and Nonlinear Programming*, 2nd ed., Addison-Wesley, Reading, Mass., 1984.
- [226] L.D. Lunsford, R.E. Latchaw and J.K. Vries, Stereotactic Implantation of Deep Brain Electrodes Using Computed Tomography, *Neurosurgery*, No. 13, pp. 280-286, 1983.
- [227] L.D. Lunsford, R. Parrish, and L. Albright, Intraoperative Imaging with a Therapeutic Computed Tomographic Scanner, *Neurosurg.*, Vol. 15, No. 4, pp. 561.
- [228] L.D. Lunsford and D. Kondziolka, Gamma Knife Radiosurgery for Brain Tumours, *Stereotactic and Directed Surgery of Brain Tumours*, ed. D.G.T. Thomas, Churchill Livingstone, pp. 169-192, 1993.
- [229] D. MacDonald, *Display - Program for Display and Segmentation of Surfaces and Volumes*, McConnell Brain Imaging Center document, 1995.
- [230] D. MacDonald, D. Avis and A.C. Evans, Proximity Constraints in Deformable Models for Cortical Surface Identification, *Medical Image Computing and Computer-Assisted Intervention- MICCAI'98*, pp. 650-659, 1998.
- [231] R.J. Maciunas, ed., *Interactive Image-Guided Neurosurgery*, American Association of Neurological Surgeons, 1993.

- [232] R.J. Maciunas, ed., *Clinical Frontiers of Interactive Image-Guided Neurosurgery*, *Neurosurgery Clinics of North America*, Vol. 7, No. 2, April 1996.
- [233] R.J. Maciunas, M.S. Berger, B. Copeland, M. Mayberg, R. Selker, and G.S. Allen, A Technique for Interactive Image-Guided Neurosurgical Intervention in Primary Brain Tumors, *Clinical Frontiers of Interactive Image-Guided Neurosurgery*, R.J. Maciunas, ed., *Neurosurgery Clinics of North America*, Vol. 7, No. 2, pp. 245-266, April 1996.
- [234] M. Mahvash and V. Hayward, Haptic Rendering of Cutting: A Fracture Mechanics Approach, *www.Haptics-e.org*, Vol. 2, No. 3.
- [235] J.B.A. Maintz, P.A. v.d. Elsen, and M.A. Viergever (1996), Evaluation of Ridge Seeking Operators for Multimodality Medical Image Matching, *IEEE Trans. PAMI* 18 (4), pp. 353-365, 1996.
- [236] J.B.A. Maintz, and M.A. Viergever, A Survey of Medical Image Registration, *Med. Image Anal.* 2 (1), pp. 1-36, 1998.
- [237] R. Malladi, J.A. Sethian and B.C Vemuri, Shape Modeling with Front Propagation: A Level Set Approach, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 17, No. 2, Feb. 1995.
- [238] R. Malladi and J.A. Sethian, A Real-Time Algorithm for Medical Shape Recovery, *Proc. IEEE Int'l. Conf. Computer Vision*, pp. 304-310, 1998.
- [239] L.E. Malvern, *Introduction to the Mechanics of a Continuous Medium*, Prentice-Hall, Englewood Cliffs, N.J., 1969.
- [240] D. Marr, *Vision*, W.H. Freeman and Co., 1982.
- [241] D. Marr and E.C. Hildreth, Theory of Edge Detection, *Proc. R. Soc. Lond. B*, No. 207, pp. 187-217, 1980.

- [242] T.S. Marrett, L. Collins, G.B. Pike, S. Ku, T.M. Peters, and A.C. Evans, Volumetric MRI/PET Image Correlation Display and Analysis, *SMRM 8th Ann. Meet. Soc. Magn. Res. Med.*, Aug. 1989.
- [243] C.R. Maurer and J.M. Fitzpatrick, *A Review of Medical Image Registration, Interactive Image-Guided Surgery*, ed. R.J. Maciunas, American Association of Neurological Surgeons, pp. 17-44, 1993.
- [244] C.R. Maurer Jr., G.B. Aboutanos, B.M. Dawant, R.J. Maciunas, and J.M. Fitzpatrick, Registration of 3-D Images Using Weighted Geometrical Features, *IEEE Trans. Medical Imaging*, Vol. 15, No. 6, pp. 836-849, Dec. 1996.
- [245] C.R. Maurer Jr., D.L.G. Hill, A.J. Martin, H. Liu, M. McCue, D. Rueckert, D. Lloret, W.A. Hall, R.E. Maxwell, D.J. Hawkes, and C.L. Truwit, Measurement of Intraoperative Brain Deformation Using a 1.5 Tesla Interventional MR System: Preliminary Results, *IEEE Trans. Medical Imaging*, Vol. 17, No. 5, pp. 817-825, Oct. 1998.
- [246] C.R. Maurer Jr., D.L.G. Hill, R.J. Maciunas, J.A. Barwise, J.M. Fitzpatrick, and M.Y. Wang, Measurement of Intraoperative Brain Surface Deformation Under a Craniotomy, *Medical Image Computing and Computer-Assisted Intervention - MICCAI '98*, 1998.
- [247] C.R. Maurer Jr., J.M. Fitzpatrick, R.J. Galloway, M.Y. Wang, R.J. Maciunas, and G.S. Allen, Registration of Head Volume Images Using Implantable Fiducial Markers, *IEEE Trans. Med. Imag.*, Vol. 16, No. 4, Aug. 1997.
- [248] C.R. Maurer Jr, R.J. Maciunas, and J.M. Fitzpatrick, Registration of Head CT Images to Physical Space Using a Weighted Combination of Points and Surfaces, *IEEE Trans. Med. Imag.*, Vol. 17, No. 5, Oct. 1998.
- [249] C.R. Maurer, R.P. Gaston, D.L.G. Hill, M.J. Gleeson, M.G. Taylor, M.R. Fenlon, P.J. Edwards, and D.J. Hawkes, AcouStick: A Tracked A-Mode Ultrasonog-

- raphy System for Registration in Image-Guided Surgery, *Med. Image Comput. and Comput. Assis. Interv.- MICCAI*, pp. 953-962, 1999.
- [250] T. McInerney and D. Terzopoulos, A Dynamic Finite Element Surface Model for Segmentation and Tracking in Multidimensional Medical Images with Application to Cardiac 4D Image Analysis, *J. Computerized Medical Imaging and Graphics*, Vol. 19, No. 1, Jan-Feb. 1995.
 - [251] T. McInerney and D. Terzopoulos, Deformable Models in Medical Image Analysis: a Survey, *Med. Image Anal.*, Vol. 1, No. 2, 91-108, 1996.
 - [252] T. McInerney and D. Terzopoulos, Topology Adaptive Deformable Surfaces for Medical Image Volume Segmentation, *IEEE Trans. Medical Image Analysis*, Vol. 18, No. 10, pp. 840-850, Oct. 1999.
 - [253] K.K. Mendis, R.L. Stalnaker and S.H. Advani, A Constitutive Relationship for Large Deformation Finite Element Modeling of Brain Tissue, *J. Biomechanical Engineering*, Vol. 117, pp. 279-285.
 - [254] S. Menet, P. Saint-Marc and G. Medioni, B-Snakes: Implementation and Applications to Stereo, *Artificial Intelligence and Computer Vision*, Elsevier Science, pp. 223-236, 1991.
 - [255] D. Metaxas and I.A. Kakadiaris, Elastically Adaptive Deformable Models, *Proc. Euro. Conf. on Computer Vision*, pp. 550-559, 1996.
 - [256] D. Metaxas and D., Terzopoulos, Shape and Nonrigid Motion Estimation Through Physics-Based Synthesis, *IEEE Trans. PAMI* 15 (6), pp. 580-591, 1993.
 - [257] J. Michiels, H. Bosmans, P. Pelgrims, D. Vandermeulen, J. Gybels, G. Marchal, and P. Suetens, On the Problem of Geometric Distortion in Magnetic Resonance Images for Stereotactic Neurosurgery, *Magn. Reson. Imag.*, Vol. 12, No. 5, pp. 749-765, 1994.

- [258] M.I. Miga, *Development and Quantification of a 3D Brain Deformation Model for Model-Updated Image-Guided Stereotactic Neurosurgery*, Ph.D. Thesis, Dartmouth College, Hanover, NH, 1999.
- [259] M.I. Miga, D.W. Roberts, A. Hartov, S. Eisner, J. Lemery, F.E. Kennedy, K.D. Paulsen, Updated Neuroimaging Using Intraoperative Brain Modeling and Sparse Data, *Stereotac. and Func. Neurosurg.*, Vol. 72, pp. 103-106, 1999.
- [260] M.I. Miga, K.D. Paulsen and F.E. Kennedy, Von Neumann Stability Analysis of Biot's General Two-Dimensional Theory of Consolidation, *Int. J. Numer. Meth. in Engng.*, Vol. 43, pp. 955-974, 1998.
- [261] M.I. Miga, K.D. Paulsen, F.E. Kennedy, P.J. Hoopes, A. Hartov, and D.W. Roberts, A 3D Brain Deformation Model Experiencing Comparable Surgical Loads, *Proc. 19th Int. Conf. IEEE EMBS*, pp. 773-776, 1997.
- [262] K. Miller and K. Chinzei, Constitutive Modelling of Brain Tissue; Experiment and Theory, *J. Biomech.*, Vol. 30, No. 11/12, pp.1115-1121, 1997.
- [263] K. Miller and K. Chinzei, Modelling of Brain Tissue Mechanical Properties; Bi-Phasic Versus Single-Phase Approach, *Comp. Meth. Biomech. and Biomed. Engrg. - 2*, ed. J. Middleton, M. L. Jones, and G.N. Pande, Gordon and Breach Science Publishers, pp. 535-542, 1998.
- [264] K. Miller and K. Chinzei, Simple Validation of Biomechanical Models of Brain Tissue, *J. Biomech.*, Vol. 31, Suppl. 1, May 1998.
- [265] K. Miller, Modelling Soft Tissue Using Biphasic Theory - A Word of Caution, *Comp. Meth. Biomech. Biomed. Eng.*, Vol.1, pp.261-263, 1998.
- [266] K. Miller, Constitutive Model of Brain Tissue Suitable for Finite Element Analysis of Surgical Procedures, *J. Biomechanics*, No. 32, pp. 531-537.

- [267] S.K. Mishra, D.B. Goldgof, T.S. Huang, Motion Analysis and Epicardial Deformation Estimation from Angiography Data, *IEEE Proc. Conf. Comp. Vision and Patt. Rec.*, pp. 331-336, 1991.
- [268] O. Monga, S. Benayoun, and O.D. Faugeras, From Partial Derivatives of 3D Density Images to Ridge Lines, *IEEE Proc. Conf. Comp. Vision and Patt. Rec.*, pp. 354-359, 1992.
- [269] O. Monga, S. Benayoun, Using Partial Derivatives of 3D Images to Extract Typical Surface Features, *Comp. Vision and Image Under.* 61 (2), pp. 171-189, 1995.
- [270] O. Monga, R. Deriche and J.-M. Rocchisiani, 3D Edge Detection Using Recursive Filtering: Application to Scanner Images, *CVGIP: Image Understanding*, Vol. 53, No. 1, pp. 76-87, Jan. 1991.
- [271] J. Montagnat, H. Delingette, N. Scapel, and N. Ayache, Representation, Shape, Topology and Evolution of Deformable Surfaces. Application to 3D Medical Segmentation, *Image and Vis. Comput.*, Vol. 19, No. 14, pp. 1023-1040, Dec. 2001.
- [272] R. Mosges and L. Klimek, Computer-assisted surgery of the paranasal sinuses, *J. Otolaryngol.* 22:2, pp. 69-71., 1993.
- [273] V.C. Mow, S.C. Kuei, W.M. Lai, and C.G. Armstrong, Biphasic Creep and Stress Relaxation of Articular Cartilage in Compression: Theory and Experiments, *Trans. ASME- J. Biomech. Eng.*, Vol. 102, pp. 73-84, 1980.
- [274] P. Munger, GR Crelier, and T.M. Peters., An Inverse Problem Approach to the Correction of Distortion of EPI Images, *IEEE Trans. Med. Imag.*, Vol. 19, No. 7, pp. 681-689, July 2000.
- [275] H. Murase and S.K. Nayar, Visual Learning and Recognition of 3-D Objects from Appearance, *Int. J. Computer Vision* 14, pp. 5-24, 1995.

- [276] B.W. Murphy, P.L. Carson, J.H. Ellis, Y.T. Zhang, R.J. Hyde, and T.L. Chenevert, Signal-to-Noise Measures for Magnetic Resonance Imagers, *Magn. Res. Imag.*, Vol. 11, pp. 425-428, 1993.
- [277] T. Nagashima, The Finite Element Analysis of Brain Oedema Associated with Intracranial Meningiomas, *Acta Neurochirurgica, Suppl.*, 51, pp. 155-157, 1990.
- [278] V.S. Nalwa and T.O. Binford, On Detecting Edges, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. PAMI-8, No. 6, pp. 699-711, Nov. 1986.
- [279] C. Nastar and N. Ayache, Fast Segmentation, Tracking, and Analysis of Deformable Objects, *IEEE Proc. Int. Conf. Comp. Vision*, pp. 275-279, 1993.
- [280] S.K. Nayar and Y. Nakagawa, Shape from Focus, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 16, No. 8, pp. 824-831, Aug. 1994.
- [281] P.F. Neumann, L.L. Sadler and J. Gieser, Virtual Reality Vitrectomy Simulator, *Medical Image Computing and Computer-Assisted Intervention- MICCAI'98*, pp. 910-917, 1998.
- [282] J. Newman, Role of Stereotactic Biopsy in Diagnosing Breast Cancer, *Radiol. Technol.*, Vol. 68, No. 4, pp. 131-148, Mar-Apr. 1997.
- [283] C. Nimsky, O. Ganslandt, S. Cerny, P. Hastreiter, G. Greiner, and R. Fahlbusch, Quantification of, Visualization of, and Compensation for Brain Shift Using Intraoperative Magnetic Resonance Imaging, *Neurosurg.*, Vol. 47, No. 5, pp. 1070-1080, Nov. 2000.
- [284] H.K. Nishihara, *Practical Real-time Imaging Stereo Matcher*, *Optical Engineering*, Vol. 23, No. 5, Oct. 1984.
- [285] D.G. Nishimura, *Principles of Magnetic Resonance Imaging*, Stanford University, 1996.
- [286] L. Nolte, L. Zamorano, E. Arm, H. Visarius, Z. Jiang, U. Berlemann, and O. Schwarzenbach, Image-guided Computer-assisted Spine Surgery: a Pilot Study

- on Pedicle Screw Fixation, *Stereotac. and Func. Neurosurg.*, Vol. 66, No. 1-3, pp.108-117, 1996.
- [287] Northern Digital Inc., *www.ndigital.com*.
- [288] R.D. Nowak, Wavelet-Based Rician Noise Removal for Magnetic Resonance Imaging, *IEEE Trans. Image Proc.*, Vol. 8, No. 10, pp. 1408-1419, 1999.
- [289] M. O'Donnell and W. Edelstein, NMR Imaging in the Presence of Magnetic Field Inhomogeneities and Gradient Field Nonlinearities, *Med. Phys.*, Vol. 12, No. 1, pp. 20-26, 1985.
- [290] B.M. Oliver, Thermal and Quantum Noise, *Proc. IEEE*, Vol. 53, pp. 436-454.
- [291] A. Olivier, T.M. Peters and G. Bertrand, Stereotactic System and Apparatus for Use with MRI, CT and DSA, *Appl. Neurophys.*, No. 48, pp. 94-96.
- [292] A. Olivier, M. Alonzo, R. Comeau, and T.M. Peters, Image-Guided Surgery of Epilepsy, *Clinical Frontiers of Interactive Image-Guided Neurosurgery*, R.J. Maciunas, ed., *Neurosurgery Clinics of North America*, Vol. 7, No. 2, pp. 245-266, April 1996.
- [293] A.K. Ommaya, Mechanical Properties of Tissues of the Nervous System, *J. Biomech.*, Vol. 1, No. 2, 1968, pp. 127-138.
- [294] T.C. Origitano, O. Al-Mefty, J.P. Leonetti, F. DeMonte, and O. Reichman, Vascular Considerations and Complications in Cranial Base Surgery, *Neurosurg.*, Vol. 35, No. 3, pp.351-362, Sept. 1994.
- [295] S. Osher and J.A. Sethian, Fronts Propagating with Curvature-Dependent Speed: Algorithms Based On Hamilton-Jacobi Formulations, *J. Computational Physics*, No. 79, pp. 12-49, 1988.
- [296] L.R. Østergaard, O.V. Larsen, J. Haase, F. van Meer, A.C. Evans, and D.L. Collins, Knowledge-Based Extraction of Cerebral Vasculature from Anatomical MRI, *Proc. SPIE Medical Imaging*, 2001.

- [297] C.B. Ostertag, Interstitial Irradiation Treatment of Low Grade Gliomas, *Stereotactic and Directed Surgery of Brain Tumours*, ed. D.G.T. Thomas, Churchill Livingstone, pp. 125-134, 1993.
- [298] S. Owen, A Survey of Unstructured Mesh Generation Technology, available online at www.andrew.cmu.edu/user/sowen/survey/index.html, 1999.
- [299] D.W. Paglieroni, Distance Transforms: Properties and Machine Vision Applications, *CVGIP: Graph. Models and Image Proc.* 54 (1), pp. 56-74, 1992.
- [300] M.R. Pamidi and S.H. Advani, Nonlinear Constitutive Relations for Human Brain Tissue, *Trans. ASME*, Vol. 100, pp. 44-48, 1978.
- [301] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, 3rd ed., 1991.
- [302] D.L. Parker and G.T. Gullberg, Signal-to-Noise Efficiency in Magnetic Resonance Imaging, *Med. Phys.*, Vol. 17, No. 2, pp. 250-257, Mar./Apr. 1990.
- [303] C.A. Pelizzari, G.T.Y. Chen, D.R. Spelbring, R.R. Weichselbaum, C.T. Chen, Accurate Three-Dimensional Registration of CT, PET, and/or MR Images of the Brain, *J. Comp. Ass. Tomog.* 13 (1), pp. 20-26, 1989.
- [304] A. Peña, M.D. Bolton, H. Whitehouse and J.D. Pickard, Effects of Brain Ventricular Shape on Periventricular Biomechanics: A Finite-element Analysis, *Neurosurg.*, Vol. 45, No. 1, July 1999.
- [305] A. Pentland and B. Horowitz, Recovery of Nonrigid Motion and Structure, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 13, No. 7, pp. 730-742, July 1991.
- [306] A. Pentland and S. Sclaroff, Closed-Form Solutions for Physically Based Shape Modelling and Recognition, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 13, No. 7, pp. 715-729, July 1991.

- [307] D.M. Perese, Superficial Veins of the Brain from Surgical Point of View, *J. Neurosurg.*, Vol. 17, pp. 402-412, 1960.
- [308] P. Perona and J. Malik, Scale-space and Edge Detection Using Anisotropic Diffusion, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 12, No. 7, pp. 629-639, July 1990.
- [309] T.M. Peters, C.J. Henri, P. Munger, A.M. Takahashi, A.C. Evans, B. Davey, and A. Olivier, Integration of Stereoscopic DSA and 3D MRI for Image-Guided Neurosurgery, *Comput. Med. Imag. and Graph.*, Vol. 18, No. 4, 1994.
- [310] T. Peters, B. Davey, P. Munger, R. Comeau, A. Evans, and A. Olivier, Three-Dimensional Multimodal Image-Guidance for Neurosurgery, *IEEE Trans. Med. Imag.*, Vol. 15, No. 2, pp. 121-128, April 1996.
- [311] Philips Gyroscan Intera I/T System Interventional and Therapeutic MR, <http://www.medical.philips.com/PageDisplay.asp?pid=4160>.
- [312] C. Picard, A. Olivier and G. Bertrand, The First Human Stereotaxic Apparatus. The Contribution of Aubrey Mussen to the Field of Stereotaxis, *J. Neurosurg.*, No. 59, pp. 673-676.
- [313] T. Poggio, H. Voorhees and A. Yuille, A Regularized Solution to Edge Detection, *J. Complexity*, Vol. 4, pp. 106-123, 1988.
- [314] W.H. Press, B. P. Flannery, S. A. Teukolsky and W. T. Vetterling, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed., Cambridge University Press, Cambridge, 1992.
- [315] M.C. Preul, Z. Caramanos, R. Kasrai, R. Leblanc, D.L. Arnold, Proton Magnetic Resonance Spectroscopic Imaging in Neurosurgery: Applications for Brain Tumors, *Advanced Neurosurgical Navigation*, ed. E. Alexander III and R.J. Maciunas, Thieme Medical Publishers, 1999.

- [316] J.G. Proakis, *Digital Signal Processing: Principles, Algorithms, and Applications*, Prentice Hall, Upper Saddle River, N.J., 1996.
- [317] M.H. Protter and H.F. Weinberger, *Maximum Principles in Differential Equations*, Prentice-Hall, 1967.
- [318] G.M. Radack and N.I. Badler, Local Matching of Surfaces Using a Boundary-Centered Radial Decomposition, *CVGIP* 45, pp. 380-396, 1989.
- [319] B. Radig, Image Sequence Analysis Using Relational Structures, *Patt. Rec.* 17 (1), pp. 161-167, 1984.
- [320] I. Ragnemalm, Neighborhoods for Distance Transformations Using Ordered Propagation, *Comput. Vis., Graph. and Image Proc.: Image Understand.*, Vol. 56, pp. 399-409, 1992.
- [321] A. Rangarajan, H. Chui, E. Mjolsness, S. Pappu, L. Davachi P. Goldman-Rakic and J. Duncan, A Robust Point-Matching Algorithm for Autoradiograph Alignment, *Med. Image Anal.* 1 (4), pp. 379-398, 1997.
- [322] H.F. Reinhardt, Sonic Stereometry in Microsurgical Procedures for Deep-Seated Brain Tumors and Vascular Malformations, *Neurosurg.*, Vol. 32, pp. 51-57, 1993.
- [323] T. Riechert, Long Term Follow-up of Results of Stereotaxic Treatment of Extrapyramidal Disorders, *Conf. Neurol.*, Vol. 22, pp. 356-363, 1962.
- [324] D.W. Roberts, A. Hartov, F.E. Kennedy, M.I. Miga, K.D. Paulsen, Intraoperative Brain Shift and Deformation: A Quantitative Analysis of Cortical Displacement in 28 Cases, *Neurosurg.*, Vol. 43, No. 4, pp. 749-760, 1998.
- [325] D.W. Roberts, M.I. Miga, A. Hartov, S. Eisner, J.M. Lemery, F.E. Kennedy, K.D. Paulsen, Intraoperatively Updated Neuroimaging Using Brain Modeling and Sparse Data, *Neurosurg.*, Vol. 45, No. 5, pp. 1199-1207, Nov. 1999, Vol. 45, No. 5.

- [326] R. Rohling, P. Munger, J.M. Hollerbach and T. Peters, Comparison of Relative Accuracy Between a Mechanical and an Optical Position Tracker for Image-Guided Neurosurgery, *J. Image Guided Surgery*, Vol. 1, pp. 30-34, 1995.
- [327] A. Rosenfeld, The Max Roberts Operator is a Hueckel-type Edge Detector, *IEEE Trans. Pattern Analysis and Machine Intelligence*, PAMI-3, pp. 101-103, 1981.
- [328] A. Rosenfeld and A.C. Kak, *Digital Picture Processing*, 2nd ed., Academic Press, 1982.
- [329] G. Roth and M.D. Levine, Extracting Geometric Primitives, *CVGIP: Image Understanding* 48 (1), pp. 1-22, 1993.
- [330] P.J. Rousseeuw and A.M. Leroy, *Robust Regression and Outlier Detection*, John Wiley and Sons, 1987.
- [331] E. Rouy and A. Tourin, A Viscosity Solutions Approach to Shape-From-Shading, *SIAM J. Numerical Analysis*, Vol. 29, No. 3, pp. 867-884, June 1992.
- [332] J.S. Ruan, T.B. Khalil, and A.I. King, Dynamic Response of the Human Head to Impact by Three-Dimensional Finite Element Analysis, *ASME J. Biomechanical Engrg.*, Vol. 116, pp. 44-50, 1994.
- [333] E. Samset and H. Hirschberg, Neuronavigation in Intraoperative MRI, *Comput. Aided Surg.*, Vol. 4, pp. 200-207, 1999.
- [334] P.T. Sander and S.W. Zucker, Inferring Surface Trace and Differential Structure from 3-D Images, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 12, No. 9, pp. 833-854, Sept. 1990.
- [335] A.A.H.J. Sauren and M.H.A. Claessens, Finite Element Modeling of Head Impact: The Second Decade, *Proc. 1993 Int. IRCOBI Conf. Biomechanics of Impacts*, pp. 241-254, 1993.

- [336] S. Sclaroff and A. Pentland, Modal Matching for Correspondence and Recognition, *M.I.T. Media Lab. Perceptual Computing Section Tech. Rep. No. 201*, 1993.
- [337] H. Schiffbauer, *Neuronavigation in Brain Tumor Surgery- Clinical beta-phase of the Oulu Neuronavigator System*, Academic Dissertation, Faculty of Medicine, Univ. Münster, available online at <http://herkules.oulu.fi/isbn9514252543>, 1999.
- [338] W. Schroeder, K. Martin and B. Lorensen, *The Visualization Toolkit*, 2nd ed., Prentice Hall PTR, 1998.
- [339] G.L. Scott and H.C. Longuet-Higgins, An Algorithm for Associating the Features of Two Images, *Proc. Royal Soc. London B* 244, pp. 21-26, 1991.
- [340] R. Sedgewick, *Algorithms in C*, 3rd ed., Addison-Wesley, 1998.
- [341] J.A. Sethian, Fast Marching Level Set Method for Monotonically Advancing Fronts, *Proc. Nat. Acad. Sci. USA*, Vol. 93, pp. 1591-1595, Feb. 1996.
- [342] J.A. Sethian, A Review of the Theory, Algorithms, and Applications of Level Set Methods for Propagating Interfaces, *Acta Numerica* 5, pp. 309-395, 1996.
- [343] J.A. Sethian, *Level Set Methods and Fast Marching Methods: Evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science*, 2nd ed., Cambridge University Press, 1999.
- [344] L.S. Shapiro, and J.M. Brady, Feature-Based Correspondence: an Eigenvector Approach, *Image and Vision Comput.* 10 (5), pp. 283-288, 1992.
- [345] D.W. Shattuck, S.R. Sandor-Leahy, K.A. Schaper, D.A. Rottenberg, and R.M. Leahy, Magnetic Resonance Image Tissue Classification Using a Partial Volume Model, *NeuroImage*, Vol. 13, pp. 856-876, 2001.
- [346] J. Shen and S. Castan, An Optimal Linear Operator for Edge Detection, *IEEE Proc. Conf. Comp. Vision and Patt. Rec.*, pp. 109-114, 1986.

- [347] T.A. Shugar and M.G. Katona, Development of Finite Element Head Injury Model, *J. Amer. Soc. Civil Engineers*, Vol. 101, No. E173, pp. 223-239, 1975.
- [348] T.A. Shugar, A Finite Element Head Injury Model, *Report No. DOT HS 289-3-550-TA*, Vol. 1, 1977.
- [349] R. Sibson, Studies in the Robustness of Multidimensional Scaling: Perturbational Analysis of Classical Scaling, *J. Royal Stat. Soc. B*, Vol. 41, pp. 217-229, 1979.
- [350] K. Siddiqi, Y. Berube, A. Tannenbaum, and S.W. Zucker, Area and Length Minimizing Flows for Shape Segmentation, *IEEE Trans. Image Proc.*, Vol. 7, No. 3, pp. 433-443, March 1998.
- [351] K. Siddiqi, S. Bouix, A. Tannenbaum and S.W. Zucker, The Hamilton-Jacobi Skeleton, *Proc. IEEE Int. Conf. Comp. Vision*, pp. 828-834, 1999.
- [352] J. Sijbers, P. Scheunders, N. Bonnet, D. Van Dyck, and E. Raman, Quantification and Improvement of the Signal-to-Noise Ratio in a Magnetic Resonance Image Acquisition Procedure, *Magn. Res. Imag.*, Vol. 14, No. 10, pp. 1157-1163, 1996.
- [353] J. Sijbers, *Signal and Noise Estimation from Magnetic Resonance Images*, Ph.D. thesis, Univ. Antwerp, 1998.
- [354] D.A. Simon, M. Hebert, T. Kanade, Real-time 3-D Pose Estimation Using a High-Speed Range Sensor, *Proc. IEEE Int. Conf. Rob. and Auto.*, pp. 2235-2241.
- [355] D.A. Simon, M. Hebert, T. Kanade, Techniques for Fast and Accurate Intra-Surgical Registration, *Proc. Int. Symp. Med. Rob. and Comp. Assis. Surg.*, pp. 90-97.
- [356] M. Sinasac and T.M. Peters, Use of 3-D Deformable Surface Models for Intra-operative Visualization and Quantification of Cerebral Tissue Resection, *Medical Physics*, Vol. 24, No. 7, p. 1211, July 1997.

- [357] E.P. Sipos, S.A. Tebo, S.J. Zinreich, D.M. Long, H. Brem, In Vivo Accuracy Testing and Clinical Experience with the ISG Viewing Wand, *Neurosurg.*, Vol. 39, pp. 194-204, 1996.
- [358] O. Škrinjar, D. Spencer and J. Duncan, Brain Shift Modeling for Use in Neurosurgery, *Medical Image Computing and Computer-Assisted Intervention- MIC-CAI'98*, pp. 641-649, 1998.
- [359] O. Škrinjar and J. Duncan, Real Time 3D Brain Shift Compensation, *Proc. Inform. Proc. Med. Imag.- IPMI*, pp. 42-55, 1999.
- [360] J.G. Sled, A.P. Zijdenbos and A.C. Evans, A Nonparametric Method for Automatic Correction of Intensity Nonuniformity in MRI Data, *IEEE Trans. Medical Imaging*, Vol. 17, No. 1, pp. 87-97, Feb. 1998.
- [361] C.P. Slichter, *Principles of Magnetic Resonance*, Springer-Verlag, 3rd ed., 1990.
- [362] P.K. Sneed and P.H. Gutin, Interstitial Radiotherapy for High Grade Gliomas, *Stereotactic and Directed Surgery of Brain Tumours*, ed. D.G.T. Thomas, Churchill Livingstone, pp. 135-147, 1993.
- [363] SNN, The Network for Image-guided Surgery, www.surgnav.com.
- [364] H.W. Sorenson, *Parameter Estimation: Principles and Problems*, Marcel Dekker, New York, 1980.
- [365] Medtronic Sofamor-Danek, StealthStation Mach 3.0 Cranial Navigation Software, www.stealthnavigator.com/frames.htm.
- [366] G. Soucy and F.P. Ferrie, Surface Recovery from Range Images Using Curvature and Motion Consistency, *Comp. Vision and Image Under.* Vol. 65, No. 1, pp. 1-18, 1997.
- [367] E.A. Spiegel, H.T. Wycis and M. Marks, Stereotaxic Apparatus for Operations on the Human Brain, *Science*, Vol. 106, pp. 349-350, 1947.

- [368] E.A. Spiegel and H.T. Wycis, Pallido-thalamotomy in Chorea, *Arch. Neurol.*, Vol. 64, pp. 495-496, 1950.
- [369] E.A. Spiegel and H.T. Wycis, *Stereoencephalotomy Part I*, Grune and Stratton, New York, 1952.
- [370] E.A. Spiegel and H.T. Wycis, Pallidotomy and Pallidoamygdalotomy in Certain Types of Convulsive Disorders, *Arch. Neurol.*, Vol. 80, pp. 714-728, 1956.
- [371] M.R. Spiegel, *Mathematical Handbook of Formulas and Tables*, Schaum's Outline Series in Mathematics, McGraw-Hill, 1968.
- [372] R.L. Spilker and J.-K. Suh, Formulation and Evaluation of a Finite Element Model for the Biphasic Model of Hydrated Soft Tissues, *Computers and Structures*, Vol. 35, No. 4, pp. 425-439, 1990.
- [373] L.H. Staib and J.S. Duncan, Model-based Deformable Surface Finding for Medical Images, *IEEE Trans. Med. Imag.*, Vol. 15, No. 5, pp. 720-731, Oct. 1996.
- [374] Y.P. Starreveld, D.G. Gobbi, and T.M. Peters, Reusable Software Components for Rapid Development of Medical Image Visualization and Surgical Planning Applications, *SPIE Medical Imaging*, San Diego, 2001.
- [375] Ph. St-Jean, A.F. Sadikot, D.L. Collins, D. Clonda, R. Kasrai, A. Evans, T.M. Peters, Automated Atlas Integration and Interactive Three-Dimensional Visualization Tools for Planning and Guidance of Functional Neurosurgery, *IEEE Trans. Medical Imaging*, Vol. 17, No. 5, pp. 672-680, Oct. 1998.
- [376] G. Stockman, Object Recognition and Localization via Pose Clustering, *Comp. Vis. Graph. Image Proc.*, Vol. 40, pp. 361-387, 1987.
- [377] G. Subsol, J.Ph. Thirion, N. Ayache, Non Rigid Registration for Building 3D Anatomical Atlases, *IEEE Proc. IEEE Int. Conf. Patt. Rec.*, pp. 576-578, 1994.

- [378] T. Sumanaweera, G. Glover, S. Song, J. Adler, and S. Napel, Quantifying MRI Geometric Distortion in Tissue, *Magnetic Resonance in Medicine*, vol. 31, No. 1, pp. 40-47, 1994.
- [379] K. Surry, www.irus.rrri.on.ca/~kath, 1999.
- [380] D.L. Swets and J. Weng, Using Discriminant Eigenfeatures for Image Retrieval, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 18, No. 8, pp. 831-836, Aug. 1996.
- [381] R. Szeliski and D. Tonnesen, Surface Modeling with Oriented Particle Systems, *Comput. Graph.*, Vol. 26, No. 2, pp. 185-194, July 1992.
- [382] R. Szeliski and S. Lavallee, Matching 3-D Anatomical Surfaces with Non-Rigid Deformations using Octree-Splines, *Int. J. Computer Vision*, Vol. 18, No. 2, pp. 171-186, 1996.
- [383] Y. Tada and T. Nagashima, Modeling and Simulation of Brain Lesions by the Finite-Element Method, *IEEE Engng. in Medicine and Biology*, pp. 497-503, Aug./Sept 1994.
- [384] J. Talairach, H. Hecasn, M. David, M. Monnier, et J. De Ajuriaguerra, Recherches sur la coagulation therapeutique des structures sous-corticales chez l'homme, *Rev. Neurol.*, Vol. 81, pp. 4-24, 1949.
- [385] J. Talairach, M. David, P. Tournoux, H. Corredor, and T. Kvasina, *Atlas d'anatomie stéréotaxique*, Masson, Paris, 1957.
- [386] J. Talairach, G. Szikla and P. Tournoux, *Atlas d'anatomie stéréotaxique du télencéphale*, Masson, Paris, 1967.
- [387] K.K. Tan, R. Grzeszczuk, D.N. Levin, C.A. Pelizzari, G.T.Y. Chen, R.K. Erickson, D. Johnson, and G.J. Dohrmann, A Frameless Stereotactic Approach to Neurosurgical Planning Based on Retrospective Patient-Image Registration, *J. Neurosurg.*, Vol. 79, pp. 296-303.

- [388] D. Terzopoulos, Regularization of Inverse Visual Problems Involving Discontinuities, *IEEE Trans. Patt. Anal. and Mach. Intell.*, Vol. PAMI-8, No. 4, pp. 413-424, 1986.
- [389] D. Terzopoulos, and D. Metaxas, Dynamic 3D Models with Local and Global Deformations: Deformable Superquadrics, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 13, No. 7, 703-714, 1991.
- [390] D. Terzopoulos, R. Szeliski, Tracking with Kalman Snakes, *Active Vision*, ed. by Blake, A., Yuille, A., MIT Press, Cambridge, Mass., pp. 3-20, 1992.
- [391] D. Terzopoulos, A. Witkin and M. Kass, Constraints on Deformable Models: Recovering 3D Shape and Nonrigid Motion, *Artificial Intelligence*, Vol. 36, pp. 91-123, 1988.
- [392] B. ter Haar Romeny, ed., *Geometry Driven Diffusion in Computer Vision*, Kluwer, 1994.
- [393] J.-Ph. Thirion, Extremal Points: Definition and Application for 3D Image Registration, *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 587-592, 1994.
- [394] A.N. Tikhonov and V.Y. Arsenin, *Solution of Ill-Posed Problems*, John Wiley and Sons, 1977.
- [395] A.W. Toga and P.M. Thompson, The Role of Image Registration in Brain Mapping, *Image and Vision Computing*, Vol. 19, No. 1-2, pp. 3-24, Jan. 2001.
- [396] F.I. Toriwaki, S. Yokoi, Voronoi and Related Neighbors on Digitized Two-Dimensional Space with Applications to Texture Analysis, in *Computational Morphology*, ed. Toussaint, G.T., North-Holland, Amsterdam, 1988.
- [397] Traxtal Technologies, www.traxtal.com.
- [398] C. Truesdell and W. Noll, *The Non-Linear Field Theories of Mechanics*, 2nd. ed., Springer-Verlag, 1992.

- [399] M. Turk and A. Pentland, Eigenfaces for Recognition, *J. Cognitive Neuroscience*, Vol. 3 (1), pp. 71-86, 1991.
- [400] <http://nessus.loni.ucla.edu/icbm/index0.html>.
- [401] J.K. Udupa, Interactive Segmentation and Boundary Surface Formation for 3-D Digital Images, *Comp. Graph. and Image Proc.* 18 (3), pp. 213-35, March 1982.
- [402] D.R. Uecker, C. Lee, Y.F. Wang, and Y. Wang, Automated Instrument Tracking in Robotically Assisted Laparoscopic Surgery, *J. Image Guided Surg.*, Vol. 1, pp. 308-325, 1995.
- [403] M. Unser, A. Aldroubi, and M. Eden, B-Spline Signal Processing: Part I- Theory, *IEEE Trans. Signal Proc.*, Vol. 41, No. 2, pp 821-833, Feb. 1993.
- [404] M. Unser, A. Aldroubi, and M. Eden, B-Spline Signal Processing: Part II- Efficient Design and Applications, *IEEE Trans. Signal Proc.*, Vol. 41, No. 2, pp. 834-848, Feb. 1993.
- [405] Vitana Corporation, *ShapeGrabber Reference Manual*, 1997.
- [406] J.T. Vlaardingerbroek and J.A. den Boer, *Magnetic Resonance Imaging*, Springer, 1996.
- [407] L. Voo, S. Kumaresan, FA Pinter, N. Yoganandan and A. Sances, Finite-element Models of the Human Head, *Medical and Biological Engineering and Computing*, pp. 375-381, Sept. 1996.
- [408] M.W. Walker, L. Shao and R.A. Volz, Estimating 3-D Location Parameters Using Dual Number Quaternions, *CVGIP: Image Understanding*, Vol. 54, No. 3, pp. 358-367, 1991.
- [409] M.Y. Wang, C.R. Maurer Jr., J.M. Fitzpatrick, and R.J. Maciunas, An Automatic Technique for Finding and Localizing Externally Attached Markers in CT and MR Volume Images of the Head, *IEEE Trans. Biomed. Eng.*, Vol. 43, No. 6, June 1996.

- [410] C.C. Ward and R.B. Thompson, The Development of a Detailed Finite Element Brain Model, *Proc. 19th Stapp Car Crash Conf.*, pp. 641-674, 1975.
- [411] C.C. Ward, Finite Element Models of the Head and Their Use in Brain Injury Research, *Proc. 26th Stapp Car Crash Conf.*, SAE Paper 821154, pp. 71-85, 1982.
- [412] R. Wasserman, R. Acharya, C. Sibata, K. Shin, A Patient Specific In Vivo Tumor Model, *Math. Biosci.*, Vol. 136, No. 2, pp. 111-140, 1996.
- [413] E. Watanabe, T. Watanabe, S. Manaka, Y. Mayanagi, and K. Takakura, Three-Dimensional Digitizer (Neuronavigator): New Equipment for Computed Tomography-Guided Stereotaxic Surgery, *Surg. Neurol.*, Vol. 27, pp. 543-547, 1987.
- [414] E. Watanabe, Y. Mayanagi, and Y. Yosugi, Open Surgery Assisted by the Neuronavigator, a Stereotactic, Articulated, Sensitive Arm, *Neurosurg.*, Vol. 28, No. 6, pp. 792-800, 1991.
- [415] K. Waters, A Physical Model of Facial Tissue and Muscle Articulation Derived from Computer Tomography Data, *Proc. Visual. in Biomed. Comput.- SPIE Vol. 1808*, pp. 574-583, 1992.
- [416] W.M. Wells III, P. Viola, and R. Kikinis, Multi-modal Volume Registration by Maximization of Mutual Information, *Med. Image Anal.*, Vol. 1, No. 1, pp. 35-51, 1996.
- [417] A.P. Witkin, Scale-space Filtering, *Proc. 8th Int. Joint Conf. Artificial Intelligence*, pp. 1019-1022, 1983.
- [418] R.P. Woods, J.C. Maziotto and S.R. Cherry, MRI-PET Registration with Automated Algorithm, *J. Comp. Assis. Tomog.*, Vol. 17, No. 4, pp. 536-546, 1993.
- [419] A. Yezzi Jr., S. Kichenassamy, A. Kumar, P. Olver, and A. Tannenbaum, A Geometric Snake Model for Segmentation of Medical Imagery, *IEEE Trans. Medical Imaging*, Vol. 16, No. 2, pp. 199-209, Apr. 1997.

- [420] A.I. Yuille and T.A. Poggio, Scaling Theorems for Zero Crossings, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 8, No. 1, pp. 15-25, Jan. 1986.
- [421] L.J. Zamorano, L. Nolte, A.M. Kadi and Z. Jiang, Interactive Intraoperative Localization Using an Infrared-Based System, *Stereotact. Funct. Neurosurg.*, Vol. 63, pp. 84-88.
- [422] X. Zeng, L.H. Staib, R.T. Schultz and J.S. Duncan, Segmentation and Measurements of the Cortex from 3D MR Images, *Medical Image Computing and Computer-Assisted Intervention- MICCAI'98*, pp. 519-530, 1998.
- [423] J. Zhang, M.F. Levesque, C.L. Wilson, R.M. Harper, J. Engel, R. Lufkin, and E.J. Behnke, Multimodality Imaging of Brain Structures for Stereotactic Surgery, *Radiology*, Vol. 175, pp. 435-441.
- [424] Z. Zhang, Iterative Point Matching for Registration of Free-Form Curves and Surfaces, *Int. J. Comp. Vision*, Vol. 13, No. 2, pp. 119-152, 1994.
- [425] C. Zhou, T.B. Khalil, and A.I. King, A New Model Comparing Impact Responses of the Homogeneous and Inhomogeneous Human Brain, *Society of Automotive Engineers, Inc. report #952714*, 1995.
- [426] Carl Zeiss Neurosurgical Microscope Systems, www.zeiss.com/spd.
- [427] Q. Zhu and J.K.F. Suh, Dynamic Biphasic Poroviscoelastic Model Simulation of Hydrated Soft Tissues and its Potential Application for Brain Impact Study, *BED-Vol. 50, Bioeng. Conf., ASME*, pp. 835-836, 2001.
- [428] O.C. Zienkiewicz, *The Finite Element Method*, McGraw-Hill, 1977.
- [429] A.P. Zijdenbos, *MRI Segmentation and the Quantification of White Matter Lesions*, Ph.D. Thesis, Vanderbilt University, Dec. 1994.
- [430] A. Zijdenbos, R. Forghani and A. Evans, Automatic Quantification of MS Lesions in 3D MRI Brain Data Sets: Validation of INSECT, *Medical Image Computing and Computer-Assisted Intervention- MICCAI'98*, pp. 439-447, 1998.

- [431] H. Zinreich, CEO IZI Medical Systems, *personal communication*, 2001.
- [432] S.J. Zinreich, S. Tebo, D.M. Long, H. Brem, D. Mattox, M. E. Loury, C. Vander Volk, W. Kotch, D.W. Kennedy, and R.N. Bryan, Frameless Stereotaxic Integration of CT Imaging Data: Accuracy and Initial Applications, *Radiol.*, Vol. 188, No. 3, pp. 735-742, Sept. 1993.
- [433] S.W. Zucker and R.A. Hummel, A Three-Dimensional Edge Detector, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. PAMI-3, No. 3, pp. 324-331.
- [434] W. Zylka and H.A. Wischmann, On Geometric Distortions in CT Images, *Proc. IEEE Eng. in Med. and Biol. Soc.*, 1996.