

QUANTIFICATION OF CORTICAL FOLDING USING MR IMAGE
DATA

ROBERT WRIGHT

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

Doctor of Philosophy

Imperial College London
Department of Computing

20th May 2016

ABSTRACT

The cerebral cortex is a thin layer of tissue lining the brain where neural circuits perform important high level functions including sensory perception, motor control and language processing. In the third trimester the fetal cortex folds rapidly from a smooth sheet into a highly convoluted arrangement of gyri and sulci. Premature birth is a high risk factor for poor neurodevelopmental outcome and has been associated with abnormal cortical development, however the nature of the disruption to developmental processes is not fully understood. Recent developments in magnetic resonance imaging have allowed the acquisition of high quality brain images of preterms and also fetuses *in-utero*. The aim of this thesis is to develop techniques which quantify folding from these images in order to better understand cortical development in these two populations.

A framework is presented that quantifies global and regional folding using curvature-based measures. This methodology was applied to fetuses over a wide gestational age range (21.7 to 38.9 weeks) for a large number of subjects ($N = 80$) extending our understanding of how the cortex folds through this critical developmental period. The changing relationship between the folding measures and gestational age was modelled with a Gompertz function which allowed an accurate prediction of physiological age.

A spectral-based method is outlined for constructing a spatio-temporal surface atlas (a sequence of mean cortical surface meshes for weekly intervals). A key advantage of this method is the ability to do group-wise atlasing without bias to the

anatomy of an initial reference subject. Mean surface templates were constructed for both fetuses and preterms allowing a preliminary comparison of mean cortical shape over the postmenstrual age range 28-36 weeks. Displacement patterns were revealed which intensified with increasing prematurity, however more work is needed to evaluate the reliability of these findings.

DECLARATION OF ORIGINALITY

I declare that the work in this thesis is my own except where explicitly acknowledged in the text.

May 20, 2016

Robert Wright

COPYRIGHT

The copyright of this thesis rests with the author and is made available under a Creative Commons Attribution Non-Commercial No Derivatives licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

PUBLICATIONS

Some ideas and figures in this thesis have appeared previously in the following publications:

R. Wright, V. Kyriakopoulou, C. Ledig, M. Rutherford, J. Hajnal, D. Rueckert, and P. Aljabar. Automatic quantification of normal cortical folding patterns from fetal brain MRI. *NeuroImage*, 91:21 – 32, 2014. ISSN 1053-8119. <http://dx.doi.org/10.1016/j.neuroimage.2014.01.034>.

R. Wright, A. Makropoulos, V. Kyriakopoulou, P. Patkee, L. Koch, M. Rutherford, J. Hajnal, D. Rueckert, and P. Aljabar. Construction of a fetal spatio-temporal cortical surface atlas from in utero MRI: Application of spectral surface matching. *NeuroImage*, 120:467 – 480, 2015. ISSN 1053-8119. <http://dx.doi.org/10.1016/j.neuroimage.2015.05.087>.

ACKNOWLEDGEMENTS

I am eternally grateful to my supervisors Daniel Rueckert and Paul Aljabar for their support, guidance and wisdom. I could not have wished for better supervisors and this thesis would not exist without them. I would also like to thank Jo Hajnal and Mary Rutherford for their enthusiasm and thoughtful feedback.

CONTENTS

| | | |
|-------|---|----|
| 1 | INTRODUCTION | 39 |
| 1.1 | The cerebral cortex and cortical folding | 39 |
| 1.2 | Abnormal folding and neuropsychiatric disease | 41 |
| 1.3 | Magnetic resonance imaging | 43 |
| 1.4 | Fetal imaging | 45 |
| 1.5 | Preterm imaging | 49 |
| 1.6 | Datasets | 50 |
| 1.7 | Objective and Contributions | 55 |
| 1.8 | Outline of thesis | 57 |
| 2 | BACKGROUND | 59 |
| 2.1 | Image Registration | 60 |
| 2.1.1 | Background | 60 |
| 2.1.2 | Transformation models | 63 |
| 2.1.3 | Image similarity | 66 |
| 2.1.4 | Optimization | 70 |
| 2.2 | Image Segmentation | 72 |
| 2.2.1 | Background | 72 |
| 2.2.2 | Atlas-based segmentation | 73 |
| 2.2.3 | Probabilistic atlas | 75 |
| 2.2.4 | Expectation-Maximisation Segmentation | 79 |
| 2.3 | Surface Modelling | 88 |
| 2.4 | Conclusion | 92 |
| 3 | AUTOMATIC QUANTIFICATION OF CORTICAL FOLDING | 93 |

| | | |
|-------|---|-----|
| 3.1 | Introduction | 93 |
| 3.2 | Material and Methods | 95 |
| 3.2.1 | Overview | 95 |
| 3.2.2 | Automatic Brain Extraction | 96 |
| 3.2.3 | Automatic Segmentation | 100 |
| 3.2.4 | Voxel-based surface representation | 102 |
| 3.2.5 | Principle curvature estimation | 103 |
| 3.2.6 | Folding measures | 106 |
| 3.3 | Results & Discussion | 108 |
| 3.3.1 | Segmentation | 108 |
| 3.3.2 | Folding Measures | 111 |
| 3.3.3 | Gompertz model | 112 |
| 3.3.4 | Regional Folding | 116 |
| 3.3.5 | Physiological Age | 119 |
| 3.4 | Conclusions | 120 |
| 4 | CONSTRUCTION OF A SPATIO-TEMPORAL SURFACE ATLAS | 123 |
| 4.1 | Introduction | 123 |
| 4.1.1 | Chapter overview | 127 |
| 4.2 | Material and Methods | 128 |
| 4.2.1 | Surface Modelling | 128 |
| 4.2.2 | Atlas construction overview | 131 |
| 4.2.3 | Spectral Matching | 132 |
| 4.2.4 | Joint spectral analysis | 136 |
| 4.2.5 | Kernel Regression | 137 |
| 4.2.6 | Surface Reconstruction | 139 |
| 4.2.7 | Temporal Correspondence | 141 |
| 4.3 | Results & Discussion | 144 |

| | | |
|-------|---|-----|
| 4.3.1 | Model Parameters | 144 |
| 4.3.2 | Sulcal Alignment Accuracy | 145 |
| 4.3.3 | Regional Overlaps | 149 |
| 4.3.4 | Embedding Regularity | 151 |
| 4.3.5 | Average Cortical Surfaces | 153 |
| 4.3.6 | Variability of average cortical surfaces | 157 |
| 4.4 | Conclusion | 158 |
| 5 | A COMPARISON OF CORTICAL SHAPE FOR FETUSES AND PRETERMS | 161 |
| 5.1 | Introduction | 161 |
| 5.2 | Atlas Construction | 163 |
| 5.2.1 | Accurate mean surface estimation from irregularly distributed samples | 165 |
| 5.3 | Results | 171 |
| 5.3.1 | Surface displacement | 172 |
| 5.4 | Discussion | 173 |
| 5.5 | Conclusion | 176 |
| 6 | CONCLUSION | 185 |
| 6.1 | Contributions | 185 |
| 6.2 | Limitations and Future work | 187 |
| | BIBLIOGRAPHY | 191 |
| A | APPENDIX | 215 |
| A.1 | Cortical surface templates | 215 |
| A.2 | Cortical surface displacement maps | 222 |

LIST OF FIGURES

- Figure 1 **Coronal slice of a primate brain:** Cells have been dyed using the Nissl staining technique revealing their organisation within the brain. Image modified from [Brain-Maps.org](#). 40
- Figure 2 **Brodmann regions.** This diagram shows various cytoarchitectural regions of the cerebral cortex, discovered through histological methods. Each region has a different cellular organisation and relates to a high level function of the brain. Image from [OpenStax College](#). 41
- Figure 3 **Cortical folding visualized through magnetic resonance imaging.** These *in utero* magnetic resonance images show the rapid convolution of the the fetal brain in the third trimester. The cortical grey matter is seen as a dark ribbon on the surface of the brain. 42
- Figure 4 **2D ultrasound.** Image showing measurement of head circumference (HC) and biparietal diameter (BPD) for a fetus which are useful biometric markers for brain development. Image modified from [Dilmen \(2001\)](#). 46

- Figure 5 **Reconstruction of 3D volume from 2D slices.** 2D slices are acquired in the transverse plane using a fast scanning sequence to freeze motion (a). The coronal (b) and sagittal (c) views are corrupted by motion between slice acquisitions. The motion between slices is estimated and a 3D volume is reconstructed from the original slice data providing a coherent 3D image (d,e,f). 48
- Figure 6 **Comparison of reconstruction techniques.** Two reconstructions of the same MR acquisition are shown using the non-super-resolution method of Jiang et al. (top row) and the super-resolution method of Kuklisova-Murgasova et al. (bottom row). 49
- Figure 7 **Fetal dataset.** Bar chart with accompanying density trace, estimated using a Gaussian kernel ($\sigma = 1$). The mean age of the cohort is shown in red. 51
- Figure 8 **Neonatal dataset.** Preterms may be characterised by their postmenstrual age at scan and their gestational age at birth. Postmenstrual age is the time passed since the first day of the last menstrual period, which is equal to chronological age (time outside the womb) plus gestational age (time inside the womb). Note babies with a lower gestational age at birth are more preterm. 196 subjects were selected (filled black circles) from the dataset, whose ages overlapped with that of the fetal dataset, for a comparison of folding between the two populations. 52

- Figure 9 **Fetal MRI dataset examples.** Axial slices of the fetal brain are shown for random subjects from the fetal dataset used for this thesis (gestational age range: 21.7 – 38.9 weeks). 53
- Figure 10 **Preterm MRI dataset examples.** Axial slices of the preterm brain are shown for a random subset of the 196 subjects selected for comparison with the fetal dataset (postmenstrual age range: 26.57 – 37.14). 54
- Figure 11 **MR and PET combined visualization.** Images are registered and aligned allowing joint visualisation [Wikimedia Commons](#). 60
- Figure 12 **Registration of two brain images.** Four registrations of a source and target image are shown from coarse to fine, top to bottom. The top row shows an affine alignment which transforms points globally and preserves collinearity, i.e. all points that lie on a line before a transformation still lie on a line afterwards. The subsequent rows show a non-rigid alignment where the source image is deformed locally using the free-form deformation (FFD) model of [Rueckert et al. \(1999\)](#). Registration models are discussed in detail in Section 2.1.2. The difference of the two images is shown in the second column, which shows the similarity of the target and the transformed source image. The third column shows the deformation fields generated by the free-form deformation method. Note a finer control point mesh allows a finer alignment of the anatomies. 62

- Figure 13 **Neonatal brain atlas.** The atlas shown is one of twenty from the dataset created by [Gousias et al. \(2012\)](#). The brain image was manually segmented into 50 regions: 32 cortical regions and 18 sub-cortical regions. 74
- Figure 14 **Target-free atlas template construction.** For N images, each image k is registered to every other image j . The average of all transformations $T_{k,j}$ gives a transformation into an “average space” for image k . In this space, the intensities of the images can be averaged to create an unbiased atlas template. 77
- Figure 15 **Fetal spatio-temporal atlas.** This atlas was developed by [Serag et al. \(2012b\)](#) and is available at [brain-development.org](#). The first row show image templates for T2 weighted images for several gestational ages. The remaining rows show greyscale probability maps for the cerebrum, GM, CSF and the lateral ventricles (top to bottom). 78
- Figure 16 **Neonatal spatio-temporal atlas.** This atlas was developed by [Serag et al. \(2012a\)](#) and is available at [brain-development.org](#). The first two rows show image templates for T1 and T2 weighted images for several gestational ages. The remaining rows show greyscale probability maps for WM, GM, CSF and subcortical GM (top to bottom). 79

- Figure 17 **Partial volume effect.** When a tissue boundary intersects a voxel, the signal intensities of the adjacent tissues are combined and an intermediate intensity is observed. This is a major confounding factor for segmentation of fetal MR images, such as the one shown. For voxels at the cerebral boundary, the signal from GM and CSF combines, resulting in a moderate intensity similar to that of WM. Accurate segmentation is particularly challenging deep within sulci, where small amounts of CSF are present (blue arrows). 84
- Figure 18 **Improved cortical segmentation using a second order MRF.** Two subjects are shown, each segmented using a standard EM-MRF algorithm (a) and an EM algorithm incorporating a second order MRF (b). Voxels correctly labelled as WM are shown in green while PV voxels mislabelled as WM are shown in red. Note the reduction in mislabelled voxels in (b). 87
- Figure 19 **Topological errors.** This figure shows the inner cortical surface of a neonatal brain, extracted from a WM segmentation using the Marching Cubes algorithm. Two errors are shown which alter the topology of the mesh from that of a sphere. The green circle shows an implausible bridge across the sylvian fissure, caused by a segmentation error, whereas the green circle shows a hole within the central sulcus, where a relatively thin layer of WM was present. Mean curvature is colour mapped onto the surface to aid visual perception of the cortical geometry (red: convex, blue: concave). 89

- Figure 20 **Inner and outer cortical surfaces extracted using the Marching Cubes algorithm.** Note the merging of the pre- and post-central gyri and also the occipital and frontal lobes. 90
- Figure 21 **Processing pipeline overview.** Multiple acquisition loops consisting of 2D snapshots are reconstructed into a complete 3D volume using a slice-to-volume reconstruction algorithm. Then a number of preprocessing steps are carried out including bias correction and brain extraction. Automatic segmentation is performed using an EM-based algorithm, delineating seven anatomical regions. A probability map defining matter inside the inner cortical boundary is then constructed by summing the posterior probability maps of subcortical structures. Principal curvatures are evaluated at each boundary voxel location and folding measures are computed to quantify folding. A mesh representation of the surface is also extracted using the marching cubes algorithm in order to visualize the computed curvature-based descriptors. 96
- Figure 22 **Patch search and weighting.** The input image and atlas images are co-aligned so that similar patches (yellow) can be found within a local neighbourhood (green) around an input voxel. The patch similarity can then give a weighting when fusing the atlas labels. 98

- Figure 23 **Calculating the principal curvatures.** The surface voxels, where the principal curvatures are evaluated, are defined as any foreground voxel that shares a face with a background voxel within the binary WM volume. The binary volume was smoothed to produce a scalar field on which to estimate the image derivatives. The calculation of the principal curvature magnitudes and signs is then split into two processes. The structure tensor is computed from the image derivatives and allows the principal curvature directions and magnitudes to be computed. The signs of the principal curvatures are then recovered using the Hessian matrix. 104
- Figure 24 **Regional parcellation.** Example of a brain parcellated into nine different regions by registering an anatomical atlas. 108
- Figure 25 **Comparison of cortical surface meshes for manual and automatic segmentations.** Mesh representations of the surfaces were constructed using the marching cubes algorithm for visualization purposes. Colour mapping depicts normalized mean curvature (H) at each point on the cortical surface. 109

Figure 26 Global folding measures with gestational age. Folding measures computed from automated segmentations are shown as either green or blue circles, whereas curvatures computed from manual segmentations are shown as red triangles. For each subject that was segmented both manually and automatically, the corresponding points are shown in red and green respectively and joined together by a line. The remaining subjects that were only segmented automatically are represented by a blue circle. A Gompertz function (black line) was fitted to each plot except global mean curvature, which did not exhibit a Gompertz-like relationship. Upper and lower confidence intervals are shown as dashed grey lines. 113

Figure 27 Folding measures computed from atlas priors. The mean curvature L_2 norm (H_N) is shown for each subject computed from automated segmentations (either green or blue circles), manual segmentations (red triangles) and atlas priors (black triangles). For subjects segmented both manually and automatically, the corresponding points are shown in red and green respectively and joined together by a line. Note the values of H_N computed from atlas priors are much lower than for subjects of a similar age. 116

- Figure 28 **Regional differences in folding measures.** (a) The change in mean curvature L_2 norm (H_N) over individual regions with gestational age is shown. All lobar regions except the insula and anterior temporal lobe exhibit a Gompertz like growth pattern. (b) The rate of growth across gestational ages was computed from the derivative of the fitted Gompertz functions for H_N in lobar regions. The rate of growth peaked around 30 weeks gestational age for all regions, however there are differences in growth rates. 118
- Figure 29 **GA prediction error.** Gestational age was predicted from the observed value of H_N by taking the inverse of the Gompertz function fitted to the data for H_N . 120
- Figure 30 **Cortical Surface Extraction.** Cortical surface models for a selection of gestational ages are overlaid on their corresponding MR image volumes. The isosurface of the sub-cortical tissue probabilities is shown as a yellow contour while the cross-section of the cortical surface model is overlaid as a red contour. 131
- Figure 31 **Framework.** An average cortical surface template is constructed for each week of gestation, with all subjects within a week of the target age contributing to the output. A spectral analysis yields spatial correspondences between a group of cortical surfaces, allowing the average surface position to be computed and a template surface constructed. Note the red/blue surface colour mapping depicts the mean curvature of the surfaces. 133

- Figure 32 **Surface Eigenmodes.** The 1st three vibration modes of two surfaces are shown. Each mode depicts a pattern of displacement for a resonant frequency of the surface between poles, colour mapped to blue and red respectively. Figure adapted from [Lombaert et al. \(2013b\)](#). 134
- Figure 33 **Spectral Matching.** Vertex-wise correspondences between two surfaces (a) are given by the shortest Euclidean distance in the spectral domain (b). Colour mapping is given by the first three spectral coordinates which are mapped to RGB channels respectively. Figure adapted from [Lombaert et al. \(2013b\)](#). 135
- Figure 34 **3D Spectral Embedding of a Group of Surfaces.** Embedded vertices of 18 cortical surfaces are shown (29 – 31 weeks GA). Colour mapping depicts the mean curvature at each point on the original surfaces (red, convex; blue, concave). 138
- Figure 35 **Average cortical surface template for a 30 week fetus.** Kernel regression in the spectral domain is used to find the average surface position. This gives a point sampling of the average cortical surface (a). Poisson surface reconstruction, is then used to extract a closed surface (b), from (a). 141

- Figure 36 **Stepwise Embedding Coordinate Propagation.** This diagram illustrates the propagation of a set of embedding coordinates ($t' = 37$) to the cortical surfaces of younger subjects. Surface colour mapping shows the first 3 embedding coordinates encoded as RGB values. These coordinates can be viewed as a surface labelling, which may be propagated to neighbouring surfaces via surfaces that are shared across neighbouring embeddings. At each iteration, labelled subjects are treated as atlases, whose labels are propagated to any unlabelled surfaces within the same embedding using kernel regression. This process can be repeated iteratively until all surfaces are labelled, yielding a shared parameterisation for all surfaces. The dashed lines enclose the surfaces used to create each of the embeddings. 143
- Figure 37 **Multi-atlas Label Propagation.** Three examples of surfaces automatically labelled using kernel regression. (a) A 19 region parcellation. (b) The first embedding coordinates (for $t' = 37$) mapped to RGB channels respectively (See Fig. 36). 144
- Figure 38 **Aligned Central Sulci.** An example of a manual delineation of the central sulcus (depicted in red), along with 38 central sulci delineations mapped from other subjects, using the established spectral correspondence (shown in white). 146

- Figure 39 **Intra-embedding sulcal mapping error.** For each embedding, the sulcal delineations were mapped pairwise for all surfaces. The average alignment error is shown for four methods, quantified by the Fréchet distance, F_d (top), and the average Fréchet distance, F_a (bottom). Error bars show the standard deviation of the alignment error. 148
- Figure 40 **Inter-embedding sulcal mapping error.** Central sulcus delineations were mapped between surfaces that contributed to different embeddings by first establishing temporal correspondences (Section 4.2.7). The average alignment error, quantified by the average Fréchet distance, F_a , is shown for four methods (a) when mapping delineations for younger source subjects (≈ 27 weeks) to progressively older target subjects and (b) when mapping delineations for older source subjects (≈ 37 weeks) to progressively younger target subjects. Error bars show the standard deviation of the alignment error. 149
- Figure 41 **Surface Parcellation.** An example of a cortical surface model parcellated into 6 regions. 150

- Figure 42 **Embedding regularity visualisation.** Each line signifies a vertex where the local structure of the surface was not preserved in the embedding. The length of the line is proportional to the distance that the embedded vertex is from the centroid of its neighbouring vertices, when both are mapped back to the spatial domain (the exact distance in mm is given by the colour mapping). Note the irregularities seen when the initial surface links are not regularised. 153
- Figure 43 **Average cortical surface templates.** Cortical surface templates were constructed for each week of gestation, for both spectral embeddings and spherical demons. Both methods produced visually similar templates, capturing an estimate of the average growth for the cohort. 155
- Figure 44 **Effect of sample size on the variability of generated atlas templates.** The average distance between distinct templates generated from disjoint subsets of the cortices was computed to give an estimate of the variability of generated surface templates. The points plotted show an average of 10 iterations for a particular sample size and age, with the error bars showing the standard deviation of the 10 iterations. The lines show a power law fit of the data for each target age. 157

- Figure 45 **Preterm atlas construction using temporal windows.** Each circle depicts a window for atlas construction, where the surfaces for subjects contained within are averaged to generate a mean template. Colours correspond to developmental trajectories for babies born at different gestational ages. For example, purple correspondences to babies born at 26 weeks GA ± 1.5 . 164
- Figure 46 **Influence of imaging quality on extracted cortical surface models.** Typical cortical surface models automatically extracted for two subjects aged 32 weeks GA, from a 1.5T non-super resolution image (a) and a 3T super-resolution image (b). Note the reduced sulcal depth in (a) and also the implausible geometry around the calcarine fissure due to segmentation error. 175
- Figure 47 **Atlas templates (GA at birth: 25-29 weeks).** For the preterm templates, each row shows a developmental trajectory for subjects born around the same age. Note the background colour for each birth age corresponds with Fig. 45 & 49. The colour mapping shows mean curvature (red, convex; blue, concave), which enhances visual perception of the template geometries. 177

- Figure 48 **Atlas templates (GA at birth: 30-36 weeks)** For the preterm templates, each row shows a developmental trajectory for subjects born around the same age. Note the background colour for each birth age corresponds with Fig. 45 & 49. The colour mapping shows mean curvature (red, convex; blue, concave), which enhances visual perception of the template geometries. 178
- Figure 49 **Template volumes.** Each colour depicts a developmental trajectories for subjects born at a particular age. These colours correspond to those in Fig. 45, 47 & 48. 179
- Figure 50 **Typical template displacement for extremely preterm babies (lateral view).** The template shown was constructed for babies born at 27 ± 1.5 weeks GA and scanned at 29 ± 1.5 weeks PMA. Arrows depict displacements from matched locations on the corresponding fetal template constructed for subjects with a GA of 29 ± 1.5 weeks. The length of the arrows show the magnitude of the displacement with a one-to-one scaling. Colour mapping depicts the scalar projection of the displacement along the surface normal (red: +3 mm, blue: -3 mm). Warmer colours show positive (outward) displacements and cooler colours show negative (inward) displacements. 180

Figure 51 **Typical template displacement for extremely preterm babies (medial view).** The template shown was constructed for babies born at 27 ± 1.5 weeks GA and scanned at 29 ± 1.5 weeks PMA. Arrows depict displacements from matched locations on the corresponding fetal template constructed for subjects with a GA of 29 ± 1.5 weeks. The length of the arrows show the magnitude of the displacement with a one-to-one scaling. Colour mapping depicts the scalar projection of the displacement along the surface normal (red: +3 mm, blue: -3 mm). Warmer colours show positive (outward) displacements and cooler colours show negative (inward) displacements. [181](#)

Figure 52 **Typical template displacement for moderately preterm babies (lateral view).** The template shown was constructed for babies born at 33 ± 1.5 weeks GA and scanned at 35 ± 1.5 weeks PMA. Arrows depict displacements from matched locations on the corresponding fetal template constructed for subjects with a GA of 35 ± 1.5 weeks. The length of the arrows show the magnitude of the displacement with a one-to-one scaling. Colour mapping depicts the scalar projection of the displacement along the surface normal (red: +3 mm, blue: -3 mm). Warmer colours show positive (outward) displacements and cooler colours show negative (inward) displacements. [182](#)

- Figure 53 **Typical template displacement for moderately preterm babies (medial view).** The template shown was constructed for babies born at 33 ± 1.5 weeks GA and scanned at 35 ± 1.5 weeks PMA. Arrows depict displacements from matched locations on the corresponding fetal template constructed for subjects with a GA of 35 ± 1.5 weeks. The length of the arrows show the magnitude of the displacement with a one-to-one scaling. Colour mapping depicts the scalar projection of the displacement along the surface normal (red: +3 mm, blue: -3 mm). Warmer colours show positive (outward) displacements and cooler colours show negative (inward) displacements. [183](#)
- Figure 54 **Lateral view, templates for extremely preterm babies (25-29 weeks GA at birth).** [216](#)
- Figure 55 **Lateral view, templates for very and moderately preterm babies (30-35 weeks GA at birth).** [217](#)
- Figure 56 **Medial view, templates for extremely preterm babies (25-29 weeks GA at birth).** [218](#)
- Figure 57 **Medial view, templates for extremely preterm babies (25-29 weeks GA at birth).** [219](#)
- Figure 58 **Superior/lateral view, templates for all preterm babies.** [220](#)
- Figure 59 **Inferior/medial view, templates for all preterm babies.** [221](#)
- Figure 60 **Lateral view, templates for extremely preterm babies (25-29 weeks GA at birth).** [223](#)

| | |
|-----------|---|
| Figure 61 | Lateral view, templates for very and moderately preterm babies (30-35 weeks GA at birth). 224 |
| Figure 62 | Medial view, templates for extremely preterm babies (25-29 weeks GA at birth). 225 |
| Figure 63 | Medial view, templates for extremely preterm babies (25-29 weeks GA at birth). 226 |
| Figure 64 | Superior/lateral view, templates for all preterm babies. 227 |
| Figure 65 | Inferior/medial view, templates for all preterm babies. 228 |

LIST OF TABLES

| | |
|---------|--|
| Table 1 | Summary of folding measures. 95 |
| Table 2 | Folding measures. The degree of gyrification was quantified using eight folding measures, each of which was formulated by summarising a shape descriptor over all surface voxels, x_i . A weighting, w , was applied to each voxel which depended on its surface voxel configuration. Each of these measures was formulated to ensure its independence of the scale and surface area. Note, K^+ , denotes the set of voxels with positive Gaussian curvature. 108 |

| | |
|---------|---|
| Table 3 | Segmentation accuracy. Each row summarises the local displacements for an automated segmentation of an individual subject from its manual “ground truth”. The median displacement and the median absolute deviation (MAD) from the median are shown, as well as a kernel estimate of the distribution of displacements ($\sigma = 0.1$). Note, a negative displacement here means that the automated surface is “inside” the manual surface. 110 |
| Table 4 | Modelling the relationship of folding measures and GA. Three models were fitted to the data: a Gompertz model, a quadratic model and a linear model. The predicted sum of square residuals $\hat{\epsilon}$ and coefficient of determination R^2 were estimated to evaluate the fit of each model. 115 |
| Table 5 | Relative peak growth rate in lobar regions. The peak increase of H_N is shown for all regions relative to the posterior temporal lobe. Peak growth occurred around 30 weeks gestational age for all lobar regions and regional differences in growth rates are apparent. 118 |
| Table 6 | Cortical region overlap. Region labels were propagated between pairs of cortical surfaces using the established spectral or spherical correspondences, before measuring their overlap using the Dice coefficient. Average Dice scores are summarised. 150 |
| Table 7 | Embedding regularity. Percentage of mesh vertices where structure is preserved (see text). 152 |

| | |
|---------|--|
| Table 8 | Atlas template variability prediction. By extrapolating the power law fit of the variability measured to sample sizes greater than 11 (see Fig. 44), it is possible to estimate the variability of surface templates generated using the full dataset size. 158 |
| Table 9 | Sample sizes for atlas construction. The number of subjects that contributed to the construction of each template is shown. Note that symmetric templates were constructed from both the left and right hemisphere of each brain, thus the number of surfaces averaged to form a template was double the sample size above. 165 |

ACRONYMS

| | |
|-----|--------------------------|
| BET | Brain Extraction Tool |
| CC | Cross Correlation |
| CHD | Congenital Heart Disease |
| CR | Correlation Ratio |
| CSF | Cerebrospinal Fluid |
| CT | Computed Tomography |

| | |
|------|-------------------------------|
| EM | Expectation-Maximisation |
| FFD | Free-Form Deformation |
| GA | Gestational Age |
| GM | Grey Matter |
| ICP | Iterative Closest Point |
| LCC | Largest Connected Component |
| MI | Mutual Information |
| MR | Magnetic Resonance |
| MRF | Markov Random Field |
| MRI | Magnetic Resonance Imaging |
| NMI | Normalised Mutual Information |
| NMR | Nuclear Magnetic Resonance |
| NICU | Neonatal Intensive Care Unit |
| PET | Positron Emission Tomography |
| PMA | Postmenstrual Age |
| PSF | Point Spread Function |
| PV | partial volume |
| RF | Radio Frequency |

| | |
|-------|--|
| ROI | Region of Interest |
| SAD | Sum of Absolute Differences |
| SBM | Surface-Based Morphometry |
| SPECT | Single-Photon Emission Computed Tomography |
| SS | Structural Similarity |
| SSD | Sum of Square Differences |
| SSFSE | Single Shot Fast Spin Echo |
| SVR | Slice to Volume |
| US | Ultrasound |
| VBM | Voxel-Based Morphometry |
| WM | White Matter |

INTRODUCTION

1.1 THE CEREBRAL CORTEX AND CORTICAL FOLDING

The cerebral cortex is a thin layer of tissue typically 2-3 mm thick that lines the cerebrum, the largest division of the vertebrate brain. It largely consists of a dense entanglement of neuron cell bodies, unmyelinated axons and glia, and is often referred to as Grey Matter (GM) due to its grey appearance. In contrast, the White Matter (WM) situated below, has relatively few neuron cell bodies and long myelinated axons. Myelination is a process where Myelin, a fatty white substance, forms an electrically insulating sheath around an axon, increasing the speed of electrical transmissions. This is what gives WM a lighter appearance compared to GM.

The cortex is the largest processing centre of the brain where neural circuits perform many high level functions including sensory perception, motor control, memory retention and language processing (Kandel, 2013). Histological techniques such as Nissl staining (Fig. 1) have revealed the cellular organisation of the brain thus allowing researchers to map out its architecture. Brodmann (1909) produced one of the first mappings of cortical cytoarchitecture, dividing the cortex into 52 distinct regions, which were later found to perform distinct functions (Fig. 2).

Cortical folding is an important neurodevelopmental process that occurs largely in the third trimester of pregnancy (Garel et al., 2001), where the cerebral cortex

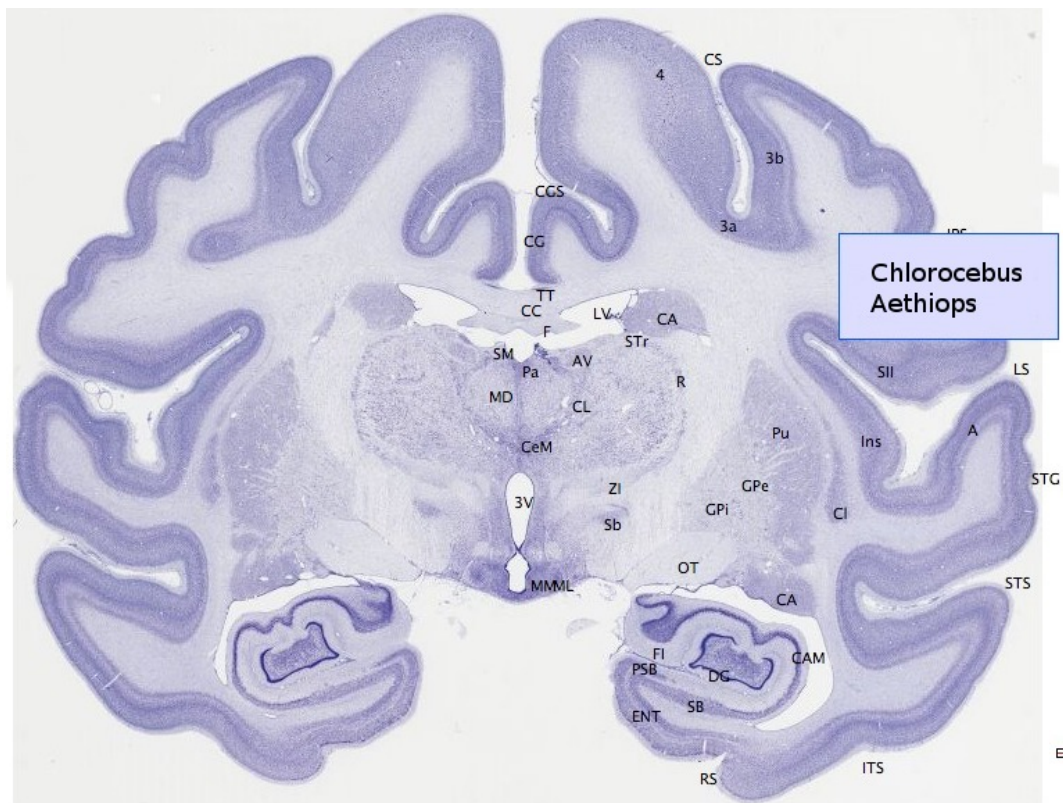


Figure 1: **Coronal slice of a primate brain:** Cells have been dyed using the Nissl staining technique revealing their organisation within the brain. Image modified from [BrainMaps.org](https://brainmaps.org).

transforms rapidly from a smooth sheet into a highly convoluted arrangement of gyri and sulci (Fig. 3). The increase in surface area allows a greater number of neuronal circuits to be packed into the skull's limited volume and is fundamental to increases in cognitive ability seen in humans and other intelligent species (Lui et al., 2011; Zilles et al., 2013). The level of gyrification is controlled by genes which determine neuronal cell birth and division in the subventricular zone, and their migration along radial glial scaffolding out towards to the cortex (Rakic, 2009). The exact mechanism that causes the cortex to fold is still unknown (Xu et al., 2010), however, several plausible theories have been put forward, including differential growth of the cortex (Richman et al., 1975) and WM axonal tension (Van Essen and Drury, 1997).

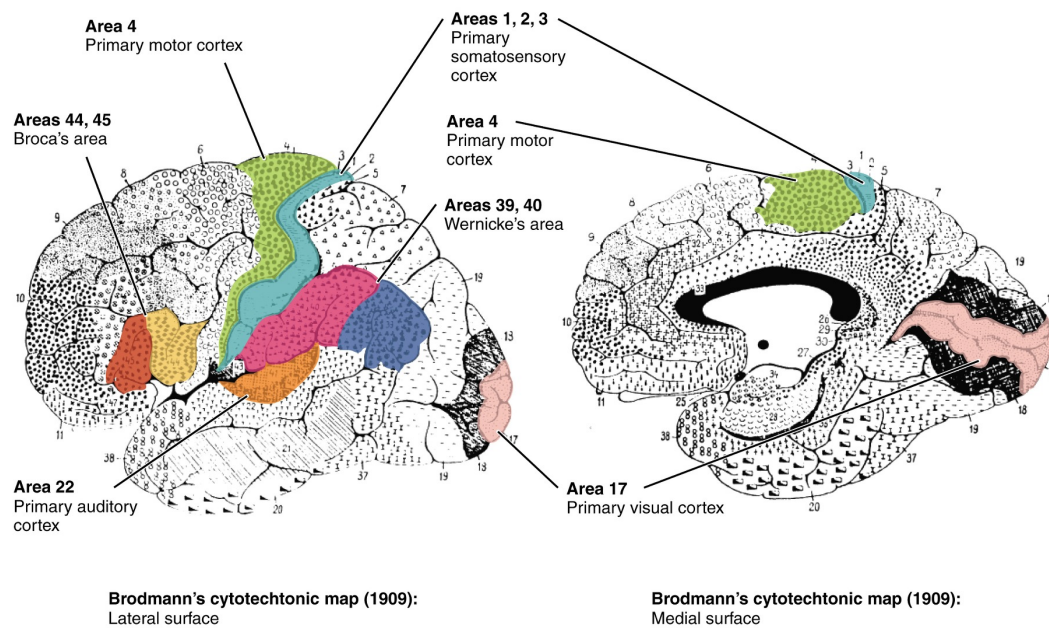


Figure 2: **Brodman regions.** This diagram shows various cytoarchitectural regions of the cerebral cortex, discovered through histological methods. Each region has a different cellular organisation and relates to a high level function of the brain. Image from [OpenStax College](#).

1.2 ABNORMAL FOLDING AND NEUROPSYCHIATRIC DISEASE

Abnormal folding of the cortex has been associated with neuropsychiatric disorders. For example, [Dubois et al. \(2008\)](#) found that fetuses with intra-uterine growth restriction, which is a risk factor for developing attention deficit hyperactivity disorder and schizophrenia ([Geva et al., 2006](#)), had measurable structural abnormalities at birth which were associated with a lower neurobehavioural development score at term equivalent age. There a number of risk factors for abnormal folding including inherited genes, premature birth and disturbances in the uterine environment. How these factors influence cortical development and neurobehavioural traits is poorly understood. By characterising cortical folding patterns in normal fetuses and neonates, and also in other groups of subjects that have specific neu-

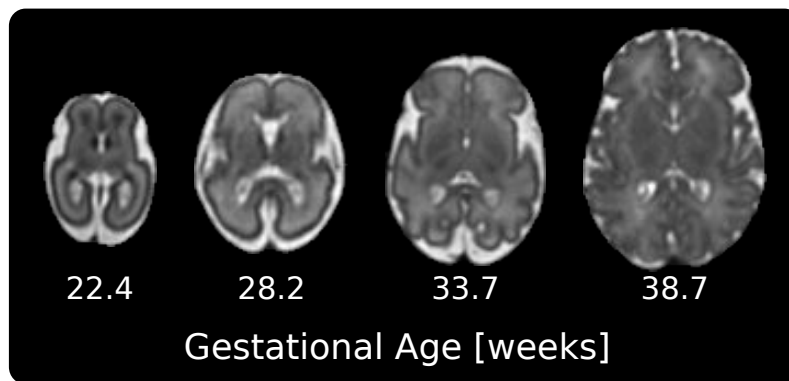


Figure 3: **Cortical folding visualized through magnetic resonance imaging.** These *in utero* magnetic resonance images show the rapid convolution of the fetal brain in the third trimester. The cortical grey matter is seen as a dark ribbon on the surface of the brain.

rocognitive deficiencies, we can begin to understand the effects of abnormal development.

One cohort of particular interest is preterm infants. In the United States, a significant proportion of the new-born population is preterm, at over 11% (Joyce A. Martin et al., 2015), although this figure has been decreasing slightly in recent years. Preterm birth is strongly associated with poor neurodevelopmental outcome (Johnson et al., 2009), however the nature of the disruption to developmental processes is not fully understood.

Medical imaging technologies such as Magnetic Resonance Imaging (MRI) have given researchers the opportunity to investigate the effect of premature birth and studies have shown that diffuse WM injury is common and predictive of poor outcome (Boardman et al., 2010). Researchers have also reported GM abnormalities such as reduced cortical tissue volume (Kapellou et al., 2006; Ball et al., 2012), reduced cortical gyrification (Ajayi-Obe et al., 2000) and also reduced thalamic volume (Ball et al., 2012). GM and WM abnormalities do not occur in isolation (Inder et al., 1999; Ball et al., 2012; Melbourne et al., 2012; 2014), suggesting that preterm birth disrupts brain development as a whole. Due to the prevalence of neurological

abnormalities in preterm babies, fetuses may ultimately provide the best control group for studying the effects of premature birth on brain development. However, acquiring high quality brain images of the fetus is a challenging task, which will be discussed in the Section 1.4.

1.3 MAGNETIC RESONANCE IMAGING

A number of different imaging techniques have been developed that exploit physical phenomena in order to visualise structures inside of the human body, including Ultrasound (US), Computed Tomography (CT), Single-Photon Emission Computed Tomography (SPECT) and Positron Emission Tomography (PET). This thesis utilises MRI for acquiring fetal and neonatal brain images, which relies on the physical phenomena known as Nuclear Magnetic Resonance (NMR), to manipulate the spin of hydrogen protons inside the body, using magnets and Radio Frequency (RF) pulses, inducing a current in a receiver coil that may be interpreted as an image.

This technique is well suited to imaging of the human body due to the abundance of water in the human body (which contains hydrogen). A major advantage is that it does not involve high frequency ionising radiation unlike other modalities such as CT (x-rays) and PET/SPECT (gamma rays), and is therefore relatively safe. This makes it an excellent choice for imaging fetuses and neonates, who are more sensitive to radiation than adults.

An MRI machine consists of a superconducting solenoid, cooled close to absolute zero using liquid helium, allowing electrons to flow without resistance. This generates a strong magnetic field (typically 1.5T or 3T), known as B_0 . Due to their intrinsic spin, protons within the body act like tiny magnets and become aligned with or against the external B_0 magnetic field. At any instant, a tiny excess of

protons are aligned with the field (a lower energy state) resulting in a net magnetisation of the body.

In order to image a subject, an RF pulse is emitted at a specific resonant frequency (the Larmor precession frequency), which is absorbed by the protons causing them to precess. As the spins return to equilibrium (spin-lattice relaxation) and become aligned with the external magnetic field (B_0) they emit radio-waves which induce a signal in the receiver coil.

Critically, the frequency of the emission is proportional to the strength of the external magnetic field (B_0). This allows spatial encoding using three sets of electromagnetic gradient coils which, combined, modify the magnetic field strength linearly along an arbitrary axis. A Fourier analysis of the signal yields the intensities for different frequencies / spatial locations.

In practice, a slice may be selected by turning on the gradient coils and emitting an RF pulse at a frequency which corresponds to the desired position along the slice selection vector ("z-axis"). This restricts the protons that are excited to the chosen slice. By turning on the gradient coils again during the readout, the precession frequency of the protons may be modified linearly along another axis within the chosen slice ("x-axis"). This is known as frequency encoding and gives spatial localisation along one axis within a slice.

Spatial localisation along the second slice axis ("y-axis"), is achieved by manipulating the phase of the precessing protons, by switching on the gradient coils for only a short duration, speeding up precession temporarily. When the gradient coil is switched off the protons precess at the same frequency again (determined by B_0) but they are now out of phase along the y-axis. Unfortunately, the Fourier transform cannot measure more than one phase at a particular frequency, however, multiple different phase encoding steps and readouts may be used to determine

the phases. Each readout becomes a row in the frequency domain or “k-space”, which yields an image when a 2D Fourier transform is applied.

The rate of signal decay (spin-lattice relaxation) after excitation is dependent on tissue type and is characterised by its T_1 value. The contrast of an image may be optimised by considering the T_1 curves for tissues of interest for a particular scan and choosing an appropriate delay after excitation before sampling the receiver coil.

Deterioration of an NMR signal is not limited to spin-lattice relaxation. After an initial excitation, local inhomogeneities in magnetic field strength cause some protons to precess faster and others more slowly, resulting in a loss of phase over time (spin-spin relaxation). This results in a loss of intensity at a particular sampling frequency as the signal broadens. Dephasing occurs at different rates for different tissues, which is characterised by its T_2 value. This gives another mechanism for obtaining contrast in images.

Typically the loss of signal due to dephasing (T_2) proceeds at a quicker rate than spin-lattice relaxation (T_1). In order to refocus the signal, an “inversion” RF pulse may be applied at time t , which flips the precessing protons causing them rephase at time $2t$, which is known as an echo. This is the most commonly used pulse sequence for acquiring T_1 images.

1.4 FETAL IMAGING

2D US is the most widely used imaging modality for assessing fetal health as it is relatively safe, inexpensive and portable. Its prevalence has allowed researchers to establish normal ranges for basic brain parameters, such as biparietal diameter,



Figure 4: **2D ultrasound.** Image showing measurement of head circumference (HC) and biparietal diameter (BPD) for a fetus which are useful biometric markers for brain development. Image modified from [Dilmen \(2001\)](#).

occipitofrontal diameter and ventricle diameter, with a high confidence from large datasets ([Snijders and Nicolaides, 1994](#)) ($N = 1040$), thereby facilitating biometric screening for abnormalities (Fig. 4). The introduction of 3D [US](#) imaging has enabled volumetry of vital organs, such as the heart ([Chang et al., 1997](#); [Peralta et al., 2006](#)), lungs ([Pöhls and Rempen, 1998](#); [Peralta et al., 2006](#)) and brain ([Chang et al., 2003](#)), thus allowing additional biometric assessment of fetal health. [Chang et al. \(2003\)](#) found that brain volume was highly correlated with Gestational Age ([GA](#)) and that measurements were reproducible with high accuracy, suggesting that this is an excellent marker for neural development. Unfortunately, measuring the volume of a brain is relatively time consuming (≈ 15 minutes) compared to alternative but simpler 2D measurements such as biparietal diameter and occipitofrontal diameter, which may be used to approximate volume. Automatic methods would be useful for this purpose, however precise delineation of anatomical structures from [US](#) data is a challenging task given the low signal-to-noise ratio and the difficulty in obtaining ‘canonical’ views of structures.

The development of ultra-fast scanning techniques such as Single Shot Fast Spin Echo (SSFSE) (Glastonbury and Kennedy, 2002) has allowed 2D MRI of the fetal brain. Unlike adults, who are requested to remain still during scanning, fetuses move freely, therefore a fast acquisition sequence is necessary to avoid motion blurring artifacts in images. MRI has since become an important tool for clinical investigation of abnormalities (Hosny and Elghawabi, 2010). With significantly greater tissue contrast and spatial resolution, MRI affords an opportunity for more detailed investigation, thus enabling more accurate diagnoses for subjects referred after an initial US screening (Rutherford et al., 2008). Additionally, MRI may be preferable when US image acquisition is affected by external factors such as maternal obesity or awkward fetal positioning.

A stack of 2D slices may be acquired in a sequence to create a 3D Magnetic Resonance (MR) image volume, however due to fetal motion between the acquisition of slices, the spatial correspondence is corrupted (Fig. 5). Slice-to-volume registration techniques (Rousseau et al., 2006; Jiang et al., 2007) have allowed coherent 3D images to be reconstructed by estimating the optimum slice alignment. Super-resolution techniques (Rousseau et al., 2010; Kim et al., 2010; Gholipour et al., 2010; Kuklisova-Murgasova et al., 2012) have subsequently boosted spatial resolution by exploiting repeated overlapping slice samples from multiple acquisition loops (see Fig. 6). These reconstruction techniques are relatively computationally intensive. However, GPU acceleration provides a significant and scalable reduction in computation time (Kainz et al., 2015), allowing the use of 3D MR images for clinical purposes. The increased resolution and signal-to-noise ratio of MR images allows excellent visualisation of complex brain structures such as the cortex and this has stimulated new research into the structural development of the fetal brain in the neuroimage analysis community (Gholipour et al., 2012; Habas et al., 2012;

Jacob et al., 2011; Rajagopalan et al., 2011; Scott et al., 2011; Serag et al., 2012b; Caldairou et al., 2011; Habas et al., 2010; Scott et al., 2013).

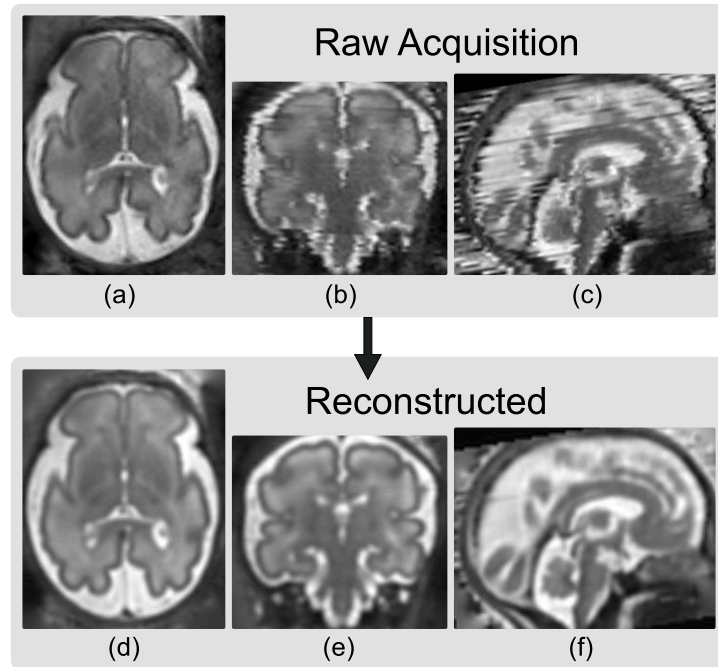
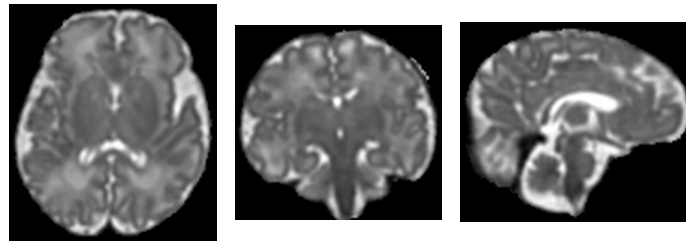


Figure 5: **Reconstruction of 3D volume from 2D slices.** 2D slices are acquired in the transverse plane using a fast scanning sequence to freeze motion (a). The coronal (b) and sagittal (c) views are corrupted by motion between slice acquisitions. The motion between slices is estimated and a 3D volume is reconstructed from the original slice data providing a coherent 3D image (d,e,f).

Non-super-resolution reconstruction



Super-resolution reconstruction

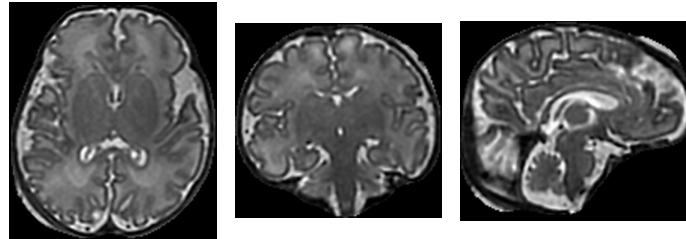


Figure 6: **Comparison of reconstruction techniques.** Two reconstructions of the same MR acquisition are shown using the non-super-resolution method of [Jiang et al.](#) (top row) and the super-resolution method of [Kuklisova-Murgasova et al.](#) (bottom row).

1.5 PRETERM IMAGING

MR imaging of preterms presents its own unique set of practical challenges ([Rutherford, 2002](#), ch. 1 & 2). Inside the womb, the mother provides nutrition for the fetus through the placenta and amniotic fluid, its temperature is well-regulated and it is also protected against infection. Outside this protective environment, preterm babies are more fragile and need to be cared for inside a Neonatal Intensive Care Unit (NICU), where vital functions such as temperature, blood oxygenation, and cardiac function may require monitoring. Babies are often placed in an incubator to help regulate their temperature and some babies may need a ventilator to help them breath. Typically, a dedicated MR scanner is used to scan preterms, which is placed as close as possible to the NICU to avoid disrupting care. Additionally,

transportation procedures are optimised to minimise handling of the baby. Inside the scanning room, the temperature must be regulated to keep the babies warm, and for extremely preterm babies extra blankets may be required. Additionally, any monitoring equipment or ventilators used must be compatible with scanner technology, in particular, the strong magnetic fields.

As with fetal imaging, motion presents an obstacle for preterm image acquisition. High quality images can be acquired, however, with immobilisation of the baby and also by using a smaller radio frequency coil, such as an adult knee coil, to get as close to the head as possible, thus increasing the signal-to-noise ratio. For immobilisation, sedatives are not necessary if babies are imaged after a feed as this helps to induce sleep. Before scanning, a vacuum-pack bag of polystyrene balls is placed around the baby's head and the air evacuated, providing snug but comfortable immobilisation while reducing the exposure to noise generated from the scanner. Noise is a primary concern when performing imaging as it can agitate the baby. This not only impedes acquisition of images free of motion artefacts but has physiological effects on the baby, such as increased heart rate, respiratory rate and intracranial pressure (Long et al., 1980). Scanners are designed, therefore, to reduce acoustic noise as much as possible. Additionally, mini ear muffs are used to further reduce noise exposure.

1.6 DATASETS

Two image datasets were used for this thesis: a fetal dataset consisting of 80 normal fetal subjects (GA range: 21.7 – 38.9 weeks) (Fig. 7) and a neonatal dataset consisting of 468 subjects of which 196 were selected (Postmenstrual Age (PMA) range: 26.57 – 37.14), whose ages overlapped with that of the fetal dataset, for a compari-

son of cortical shape between the two populations (Fig. 8). For the fetal dataset, T2-weighted MRI was performed on a 1.5T Philips Achieva system with a 32 channel coil. A SSFSE sequence was used to acquire 2D slices ($TR = 15000$ ms, $TE = 160$ ms, flip angle = 90° and voxel resolution = $1.25 \times 1.25 \times 2.5$ mm). An additional Slice to Volume (SVR) reconstruction step was then applied (Jiang et al., 2007) to produce coherent 3D volumes from these 2D slices (See Section 1.4 and Fig. 5) because the slices are disoriented due to fetal motion during acquisition. Correcting for this requires registration of the slices to find the best alignment. After reconstruction, the final 3D volumes had an isotropic resolution of $1.18 \times 1.18 \times 1.18$ mm. (Jiang et al., 2007) For the neonatal dataset, T2-weighted MRI was performed on 3T Philips Intera system ($TR = 8670$ ms, $TE = 160$ ms, flip angle = 90° , field of view = 220 mm and voxel resolution = $0.86 \times 0.86 \times 1$ mm). Both datasets were processed using N4 (Tustison et al., 2010) to correct for intensity non-uniformity.

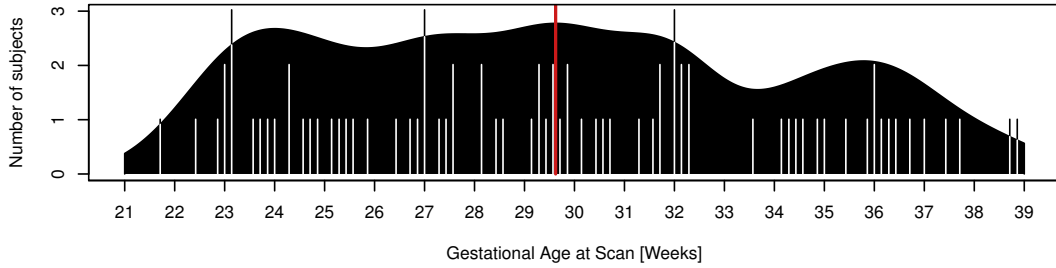


Figure 7: **Fetal dataset.** Bar chart with accompanying density trace, estimated using a Gaussian kernel ($\sigma = 1$). The mean age of the cohort is shown in red.

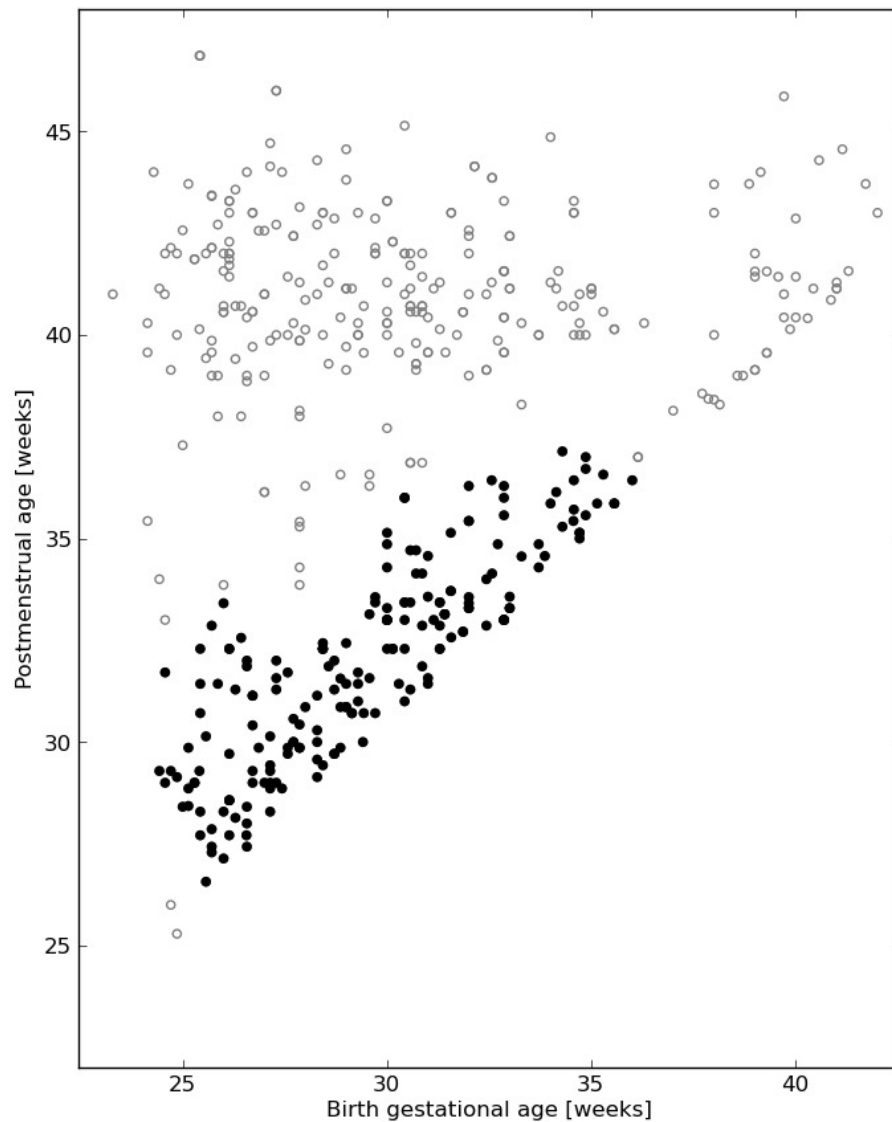


Figure 8: **Neonatal dataset.** Preterms may be characterised by their postmenstrual age at scan and their gestational age at birth. Postmenstrual age is the time passed since the first day of the last menstrual period, which is equal to chronological age (time outside the womb) plus gestational age (time inside the womb). Note babies with a lower gestational age at birth are more preterm. 196 subjects were selected (filled black circles) from the dataset, whose ages overlapped with that of the fetal dataset, for a comparison of folding between the two populations.

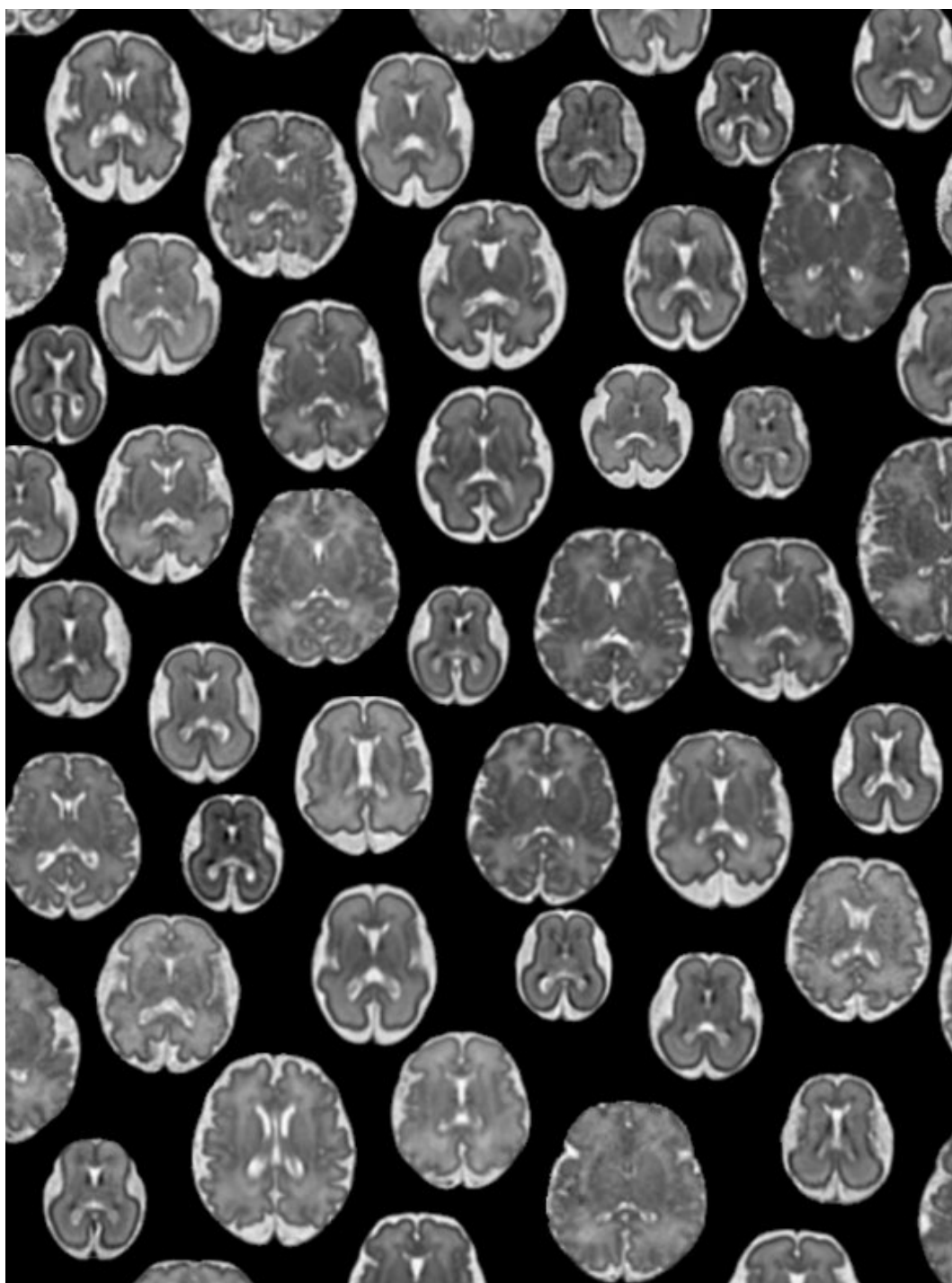


Figure 9: **Fetal MRI dataset examples.** Axial slices of the fetal brain are shown for random subjects from the fetal dataset used for this thesis (gestational age range: 21.7 – 38.9 weeks).

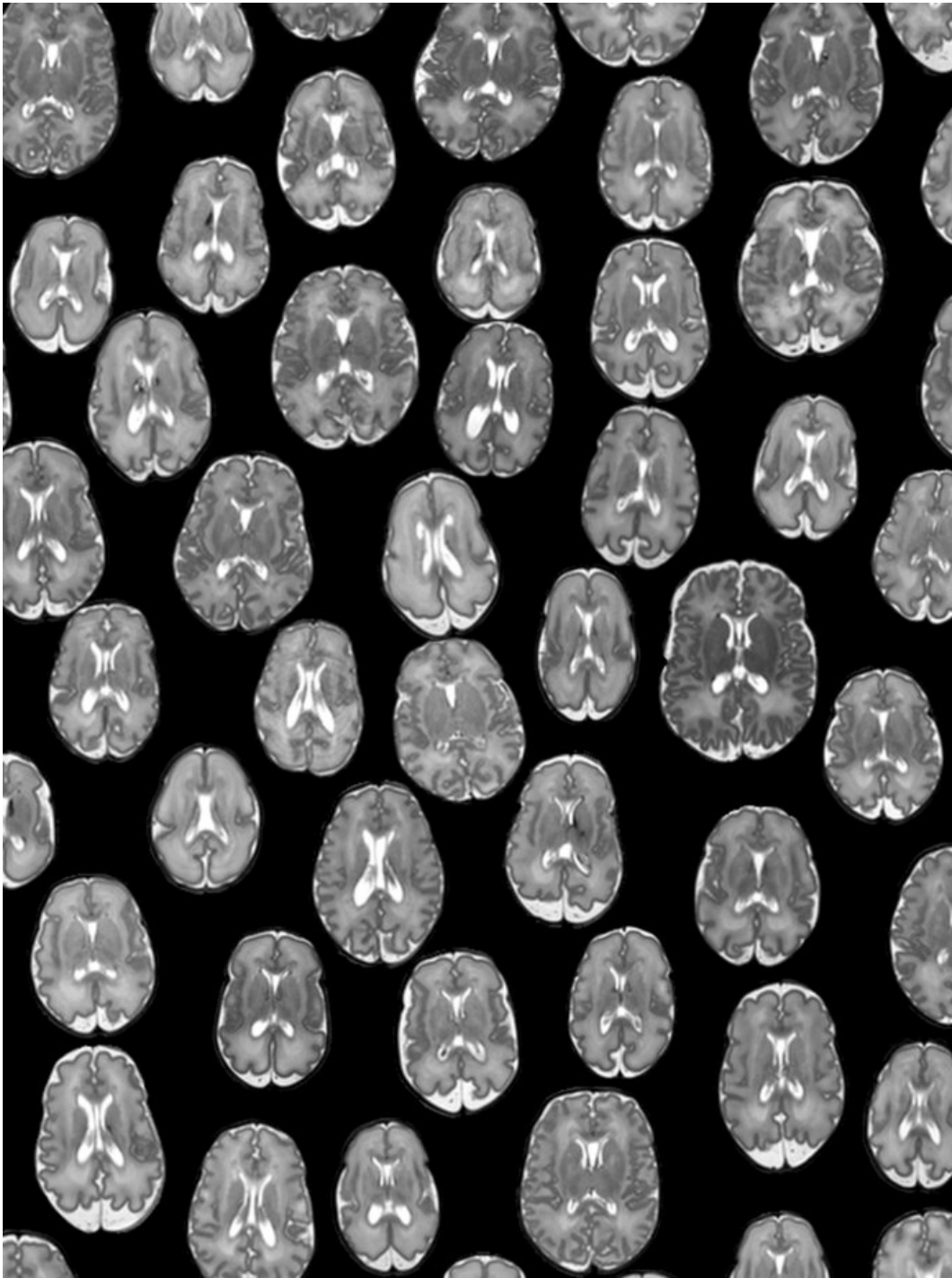


Figure 10: **Preterm MRI dataset examples.** Axial slices of the preterm brain are shown for a random subset of the 196 subjects selected for comparison with the fetal dataset (postmenstrual age range: 26.57 – 37.14).

1.7 OBJECTIVE AND CONTRIBUTIONS

The objective of this thesis is to develop techniques that will allow cortical folding to be quantified automatically for both fetuses and neonates from MR images in order to gain a greater understanding of the folding process *in utero* and how this deviates in preterm babies.

The contributions in this thesis comprise of two methods and one application as follows:

- A framework for quantifying global and regional folding using curvature-based folding measures (Chapter 3), which was applied to fetal data. Brain structures, including the cortex, were delineated automatically using an Expectation-Maximisation (EM)-based algorithm (Ledig et al., 2012) and a spatio-temporal atlas of brain structure location (Serag et al., 2012b). The accuracy of this procedure was evaluated in sixteen of the subjects for which structures were also delineated manually. For each subject, surface curvatures were evaluated in the image volume domain over the delineated cortical boundary using techniques proposed in (Rieger et al., 2004; Rodriguez-Carranza et al., 2008). Eight curvature-based folding measures from the literature (Van Essen and Drury, 1997; Magnotta et al., 1999; Batchelor et al., 2002; Awate et al., 2008; Rodriguez-Carranza et al., 2008) were adapted to quantify the degree of gyrification of the subjects within the cohort. Local gyrification was also quantified in nine different regions by parcellating each of the image volumes using an anatomical atlas (Gousias et al., 2012). Folding was quantified over a wide gestational age range (21.7 to 38.9 weeks) for a large number of subjects ($N = 80$) extending our understanding of how the cortex folds through this critical developmental period. The changing relationship

between the folding measures and gestational age was modelled with a Gompertz function which allowed an accurate prediction of physiological age.

- A framework for constructing a spatio-temporal surface atlas of cortical folding, using spectral methods to find accurate correspondences between surfaces and kernels to generate temporally and spatially weighted average surface templates (Chapter 4), which was applied to fetal data. In this work, brain structures were again delineated automatically using an EM-based algorithm (Makropoulos et al., 2014) and a spatio-temporal atlas of brain structure location (Serag et al., 2012b). Topologically corrected surface models were extracted by deforming a mesh with the correct topology towards the delineated cortical boundary using a technique similar to previous authors (Dale and Sereno, 1993; Davatzikos and Bryan, 1996; MacDonald et al., 2000; Kim et al., 2005). A group-wise extension to the spectral-based method of Lombaert et al. (2013b) was used to determine correspondences between meshes and embed them in the same domain. Kernels (Nadaraya, 1964), were then used to compute a spatially and temporally weighted estimate of the average surface boundary in this domain, which was readily converted to a mesh representation (Kazhdan and Hoppe, 2013). Sulcal alignment, regional overlaps, and embedding regularity were found to be comparable with that of state of the art methods. Templates generated from disjoint subsets of the data were also found to be highly consistent.
- An application of surface atlasing to compare mean cortical shape for fetuses and neonates (Chapter 5). In this work, surface atlas templates were constructed for groups of preterms with both a similar GA birth and PMA at scan, allowing the effects of premature birth to be investigated along different developmental trajectories. The mean cortical shape was compared between each preterm template and its corresponding fetal template (con-

structured for the same PMA). For this, correspondences were established via a joint spectral analysis and displacements were computed for matched surface locations. Patterns of displacement were then analysed for extremely preterm babies and also very and moderately preterm babies.

1.8 OUTLINE OF THESIS

Analysis of cortical folding is a challenging task that requires a combination of image analysis techniques, including image registration (aligning anatomies), image segmentation (delineating brain structures), surface modelling (extracting an explicit mesh representation from an image), surface registration (finding correspondences between surface geometries), and surface-based morphometry (detecting localised statistical differences in anatomy between cohorts).

In the next chapter an overview of neuroimage analysis tasks and concepts that this thesis builds upon will be given, including image registration, segmentation and surface modelling. Specific challenges that relate to fetal and neonatal datasets, and methods from the literature that tackle these, will also be discussed. The two subsequent chapters (3 & 4) will detail methods developed for this thesis that help to characterise cortical folding in fetuses and neonates. In Chapter 3 a framework for quantifying gyrification using curvature-based shape descriptors will be described while Chapter 4 will outline a framework for constructing a spatio-temporal cortical surface atlas. In Chapter 5, cortical shape is compared for fetuses and preterms by application of surface atlasing. The conclusions of this thesis will then be presented in the final chapter (Chapter 6).

BACKGROUND

This chapter gives an overview of neuroimage analysis concepts and tasks that this thesis builds on, along with a description of important methods in these areas. Difficulties relating to fetal and neonatal [MR](#) data will also be highlighted along with methodology developed specifically to tackle these challenging data. The content is split into three sections: image registration, image segmentation and cortical surface modelling. Image registration is concerned with aligning images, whereas image segmentation is the process of labelling anatomical structures on an image voxel grid. Registration facilitates numerous medical image analysis tasks, which will be discussed in the next section. Most importantly, in the context of this thesis, it gives a mechanism for propagating a prior estimate of tissue location from a digital brain atlas, facilitating automatic tissue segmentation. A mesh model of a tissue boundary may then be readily extracted from a tissue segmentation. Several techniques have been developed to extract cortical surface meshes which, use prior knowledge of cortical topology to improve accuracy. These will be discussed in the last section of this chapter.

2.1 IMAGE REGISTRATION

2.1.1 Background

Image registration is the process of automatically aligning two images (a “source” and a “target”) and involves estimating the optimal transformation between them (Fig. 12). Aligning images is extremely useful as it allows inspection and analysis in a common coordinate system. For example, in a clinical setting, a radiologist may wish to track disease progression in a single patient from a recent scan and several older scans. When these images are well aligned, changes may be tracked much more easily. Additionally, images acquired from different imaging modalities that provide complementary information may be registered (multi-modal registration), thus allowing simultaneous visualization (Fig. 11). For example, a neurosurgeon may co-align a PET image, which allows visualization of metabolic processes, with a structural MR image, improving tissue localisation before neurosurgery (Maciunas et al., 1992).

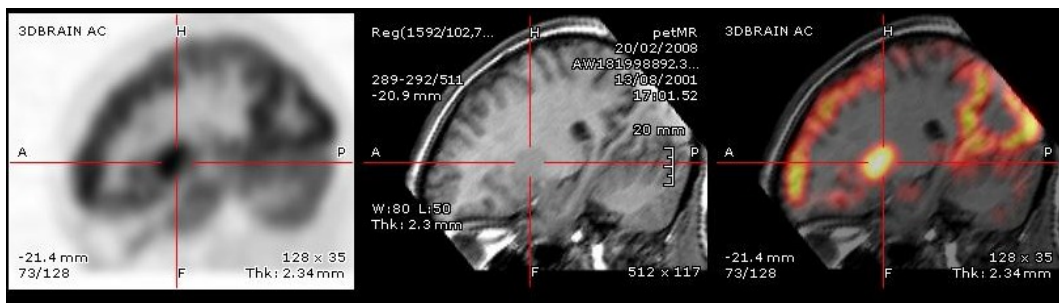


Figure 11: **MR and PET combined visualization.** Images are registered and aligned allowing joint visualisation [Wikimedia Commons](#).

The registration applications discussed so far involve registering the imaged anatomy of a single subject (intra-subject registration), however, registration may

also be used to register the anatomies of different subjects (inter-subject registration). Inter-subject registration is frequently used in cross-sectional image analysis studies, allowing morphometric analysis of the brain between different populations. For example, aligning the imaged anatomy of two populations in a common coordinate system allows a statistical analysis of their differences on a voxel-wise basis. This experimental design is known as Voxel-Based Morphometry (VBM) (Ashburner and Friston, 2000). Often a population with a specific disease is compared with a normal control population to discover abnormalities. This methodology has been successfully applied to a number of diseases including Alzheimer’s disease (Karas et al., 2004), Parkinson’s Disease (Price et al., 2004) and epilepsy (Keller et al., 2004). The deformations that map subjects into a reference space may also be used for morphometric analysis. For example, the determinant of the Jacobian operator of a deformation field can be used to quantify voxel-wise volume change for all images with respect to a reference image. Population differences in localised tissue volume may then be inferred (Boardman et al., 2006). This type of analysis is referred to as deformation-based morphometry, but the term tensor-based morphometry may also be used when other features of the Jacobian tensor are employed.

A typical image registration algorithm consists of three components: a transformation model, an image similarity metric and an optimization procedure. A transformation is applied to one of the images, which is usually referred to as the “source”, bringing it into alignment with the other image, referred to as the “target”. The goal of a registration algorithm is to estimate the best transformation for aligning a source image with a target image, by optimizing the parameters of a transformation model in order to maximise an image similarity metric. In the remainder of this section these three components will be discussed in further detail.

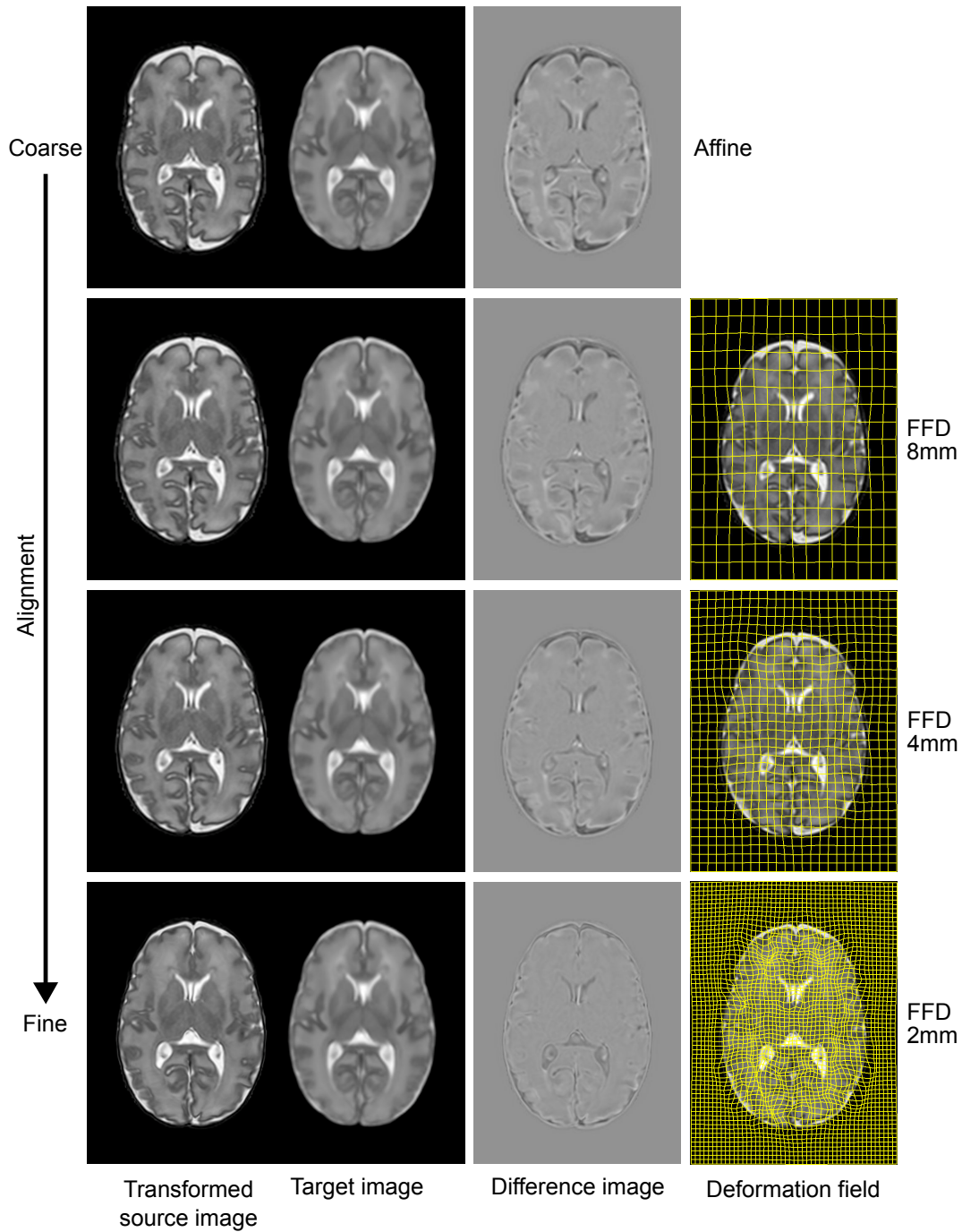


Figure 12: **Registration of two brain images.** Four registrations of a source and target image are shown from coarse to fine, top to bottom. The top row shows an affine alignment which transforms points globally and preserves collinearity, i.e. all points that lie on a line before a transformation still lie on a line afterwards. The subsequent rows show a non-rigid alignment where the source image is deformed locally using the free-form deformation (FFD) model of [Rueckert et al. \(1999\)](#). Registration models are discussed in detail in Section 2.1.2. The difference of the two images is shown in the second column, which shows the similarity of the target and the transformed source image. The third column shows the deformation fields generated by the free-form deformation method. Note a finer control point mesh allows a finer alignment of the anatomies.

2.1.2 Transformation models

Transformation models can be grouped in two major categories: linear and non-linear. A linear or affine transformation preserves collinearity, i.e. all points that lie on a line before a transformation still lie on a line afterwards. Consequently, a linear transformation operates on points globally, i.e. it does not allow localised deformation. An affine transformation can be any combination of translation, rotation, reflection, scaling and shearing. This type of transformation can be used to align brain images of different subjects with different sizes.

A rigid transformation is also a linear transformation but it only permits any combination of rotation and translation. This preserves distances and angles between points (isometry). This is particularly useful when registering two images of the same brain that have been captured using different modalities. As there is little change in brain shape over short periods of time, a global transformation is sufficient to align the imaged anatomy.

An affine transformation can be represented using homogeneous coordinates. In this coordinate system, a 3D position vector $[x \ y \ z]^t$ becomes a 4-vector $[x \ y \ z \ 1]^t$. Any 3D affine transformation may then be encoded by a 4×4 transformation matrix with 12 parameters (\mathbf{T}_{global}).

$$\mathbf{T}_{global} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

The location of a transformed point in homogeneous coordinates $[x' \ y' \ z' \ 1]^t$ is computed by multiplying a transformation matrix \mathbf{T}_{global} with an initial position vector $[x \ y \ z \ 1]^t$:

$$\begin{bmatrix} x' \\ y' \\ z' \\ 1 \end{bmatrix} = \mathbf{T}_{global} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}.$$

Five examples of affine transformation matrices are shown below for different transformation operations, including scaling (**s**), translation (**t**) and rotation around the z -axis (**r_z**), y -axis (**r_y**) and x -axis (**r_x**):

$$\mathbf{s}(s_x, s_y, s_z) = \begin{bmatrix} s_x & 0 & 0 & 0 \\ 0 & s_y & 0 & 0 \\ 0 & 0 & s_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \mathbf{t}(t_x, t_y, t_z) = \begin{bmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{r}_z(\gamma) = \begin{bmatrix} \cos \gamma & -\sin \gamma & 0 & 0 \\ \sin \gamma & \cos \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{r}_y(\beta) = \begin{bmatrix} \cos \beta & 0 & \sin \beta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin \beta & 0 & \cos \beta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \mathbf{r}_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \alpha & \sin \alpha & 0 \\ 0 & -\sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Any number of such transformation matrices may be composed, resulting in another 4×4 transformation matrix encoding the whole sequence of transformation operations.

The shape of the human brain and its constituent substructures are highly variable. A global transformation model is sufficient to bring two brains into coarse alignment, however, more complex transformation models that allow localised deformation are required to accurately align brain regions. Models that allow location

deformation are referred to as non-linear or non-rigid transformation. Image volumes have been modelled as physical entities, such as an elastic body (Yanovsky et al., 2008) which resists deformation under the influence of an external force that attempts to match the images. Viscous fluid models (Christensen et al., 1996) have also been used, which can flow allowing large-scale deformations. Other prevalent methods are derived from interpolation theory (Bookstein, 1989; Rueckert et al., 1999; Hellier et al., 2001), where displacements at discrete locations are interpolated over an entire image volume, constructing a smoothly varying displacement field. These models have the advantage of being parametrised by relatively few parameters. For a thorough and recent review of non-rigid registration models the interested reader is referred to (Sotiras et al., 2013).

In this thesis, the Free-Form Deformation (FFD) model of (Rueckert et al., 1999) is employed for non-rigid registration, which is an interpolation-based method. This model comprises of a regular grid of control points each with an associated displacement vector. A smoothly varying displacement field can be constructed by blending these vectors using cubic B-spline basis functions.

Let ϕ denote a $n_x \times n_y \times n_z$ mesh of control points with uniform spacing, then the displacement for the image coordinate (x, y, z) is calculated below:

$$T_{local}(x, y, z) = \sum_{l=0}^3 \sum_{m=0}^3 \sum_{n=0}^3 B_l(u) B_m(v) B_n(w) \phi_{i+l, j+m, k+n}$$

where $i = \lfloor x/n_x \rfloor - 1$, $j = \lfloor y/n_y \rfloor - 1$, $k = \lfloor z/n_z \rfloor - 1$,

$$u = x/n_x - \lfloor x/n_x \rfloor, \quad v = y/n_y - \lfloor y/n_y \rfloor, \quad w = z/n_z - \lfloor z/n_z \rfloor$$

and B_l is the l -th B-spline basis function of which are given below:

$$B_0(u) = \frac{(1-u)^3}{6}, \quad B_1(u) = \frac{3u^3 - 6u^2 + 4}{6},$$

$$B_2(u) = \frac{-3u^3 + 3u^2 + 3u + 1}{6}, \quad B_3(u) = \frac{u^3}{6}.$$

2.1.3 Image similarity

The alignment of a target image and a transformed source image may be estimated by an image similarity metric. A source image must first be “resliced” onto the target image’s voxel grid using an interpolation scheme. Common interpolation schemes for this purpose include nearest neighbour, linear and sinc. Intensities can then be compared at matched locations across the target’s voxel grid. The choice of metric is dependent on the imaging data and some experimentation is usually necessary to determine the most suitable one. Many factors may be relevant when evaluating the performance of an image similarity metric for a particular application, including robustness in the presence of noise, precision of alignment and ease of optimisation, especially if images are initially poorly aligned. When registering images of the same modality, a metric may compare voxel intensities directly. This approach is unsuitable for multi-modal registration, however, and information theoretic techniques are commonly used in this instance.

2.1.3.1 Mono-modal metrics

One of the simplest and most commonly used measures for mono-modal registration is the Sum of Square Differences (SSD) of the image intensities. Note that SSD is strictly a dissimilarity metric but may be converted into a similarity metric by negating its value.

Let the intensities of a target image with n voxels be denoted as a 1D vector $[t_1, t_2, \dots, t_n]$ and the intensities of a resliced source image be denoted

$[s_1, s_2, \dots, s_n]$, where t_k and s_k are the intensities of corresponding voxels. Then the SSD is given:

$$\text{SSD} = \frac{1}{n} \sum_{i=1}^n (t_i - s_i)^2.$$

The use of SSD as an image similarity metric is coupled with the assumption that equivalent tissues between images have the same intensity. Unfortunately MR image intensities do not correspond directly to tissue composition, even for a standardized protocol. Therefore, images must first be normalized, which can be accomplished by transforming an image histogram to match a standardized model (Nyúl and Udupa, 1999).

The accuracy of alignment using SSD can be effected by intensity outliers, for example, from image noise or reconstruction artefacts. Sum of Absolute Differences (SAD) provides an alternative to SSD that provides a more robust estimate of the optimum transformation in the presence of outliers:

$$\text{SAD} = \frac{1}{n} \sum_{i=1}^n |t_i - s_i|.$$

Another popular similarity metric for mono-modal registration is Cross Correlation (CC). This metric does not require normalization of images as long as their intensities have a linear relationship:

$$\text{CC} = \frac{\sum_{i=1}^n (t_i - \bar{t})(s_i - \bar{s})}{\sqrt{\sum_{i=1}^n (t_i - \bar{t})^2} \sqrt{\sum_{i=1}^n (s_i - \bar{s})^2}}.$$

2.1.3.2 Multi-modal metrics

The mono-modal metrics discussed in the previous section assume a linear relationship between image intensities. Therefore, they perform poorly when registering images acquired using different modalities, where no such relationship exists. Mutual Information (MI), which was independently proposed by [Collignon et al. \(1995\)](#) and [Viola and Wells \(1995\)](#), looks for a predictable, non-linear, relationship between intensities by constructing the joint image histogram.

Constructing a normalised histogram for an image, by first binning its intensities, allows the probability of a particular image intensity to be estimated. Let the probability for target and source image intensities t and s be denoted $p(t)$ and $p(s)$ respectively, then their Shannon entropies are given:

$$H_T = - \sum_t p(t) \log(p(t)), \quad H_S = - \sum_s p(s) \log(p(s)).$$

The Shannon entropy quantifies the amount of information in an image. For example, an image with mainly one value has a low amount of information whereas an image with intensities that are more evenly distributed has more information. A joint histogram may also be constructed allowing the probability of a target intensity t and a source intensity s at the same location to be estimated, $p(t, s)$. Thus, the joint entropy is given:

$$H_{T,S} = - \sum_t \sum_s p(t, s) \log(p(t, s))$$

and MI is defined as the sum of the independent image entropies minus the joint entropy of the images:

$$\text{MI} = H_T + H_S - H_{T,S}.$$

Consider the example of registering two structural images captured using different modalities, where a target image has mean tissue intensities t_{gm} , t_{wm} and t_{csf} and a source image has mean tissue intensities s_{gm} , s_{wm} and s_{csf} . If the images are well aligned then the joint probabilities $p(t_{gm}, s_{gm})$, $p(t_{wm}, s_{wm})$ and $p(t_{csf}, s_{csf})$ will be high and the probability of other intensity pairs will be low, leading to a low joint entropy and a high [MI](#).

[MI](#) is not, however, independent of the overlap between two images. This is problematic as the image background may contribute disproportionately to the mutual information for a given image alignment. For example, consider rigidly aligning two images of a spherical phantom whose volume is significantly smaller than the image volume. The best rotational alignment of the images is arbitrary, however, the lowest joint entropy and highest [MI](#) is given when the background overlaps the most. [Studholme et al. \(1999\)](#) later introduced Normalised Mutual Information ([NMI](#)), which is independent of image overlap and is defined as the ratio of the sum of marginal entropies, H_T and H_S , and the joint entropy, $H_{T,S}$:

$$\text{NMI} = \frac{H_T + H_S}{H_{T,S}}.$$

[Roche et al. \(1998\)](#) proposed the Correlation Ratio ([CR](#)) as an image similarity which offers a robust alternative to [MI](#). This measures the functional dependence of two images, i.e. how well the intensities in one image predicts the intensities of another. This is achieved by binning the image intensities for a target image and looking at the variance of the image intensities for the corresponding voxels in a source image.

Let the image data for a target image T be binned by intensity such that Ω_k is the set of voxels for bin k . Then the mean and variance may be computed for a source image S within each bin and also globally:

$$\sigma^2 = \frac{1}{n} \sum_{i=0}^n s_i^2 - \bar{s}^2, \quad \bar{s} = \frac{1}{n} \sum_{i=0}^n s_i.$$

$$\sigma_k^2 = \frac{1}{|\Omega_k|} \sum_{i \in \Omega_k} s_i^2 - \bar{s}_k^2, \quad \bar{s}_k = \frac{1}{|\Omega_k|} \sum_{i \in \Omega_k} s_i.$$

The correlation ratio is then defined:

$$\text{CR} = \frac{1}{n\sigma^2} \sum_k |\Omega_k| \sigma_k^2$$

where n is the number of voxels in an image.

2.1.4 Optimization

The parameters of a transformation model (e.g. the matrix elements of an affine transformation matrix or a set of [FFD](#) deformation vectors) can be optimised iteratively to maximise a similarity metric and bring a pair of images into alignment. A parameter may be perturbed by a small amount, and if this results in an increase in the similarity metric, then it may be modified permanently. This process may be iterated until convergence, where the image similarity is no longer increasing and a locally optimal solution is found. This style of optimisation is referred to as “hill climbing”, as the parametric space is traversed in order to find a peak image similarity value. A number of variants of this procedure exist, such as “steepest ascent hill climbing” where all parameters are perturbed in turn, and the parameter change that results in the greatest increase in similarity is modified permanently. Hill climbing algorithms employ a local search strategy and may therefore converge to a local maximum which may not be the global maximum. A good

initialisation of the parameters, i.e. close to the optimal solution, can mitigate this problem.

Hill climbing methods can be slow to converge to a solution, particularly where ridges exist in the parameter space, where a zigzagging path is taken. For this reason, gradient descent methods, which compute the direction of maximum increase for an image similarity metric (or maximum decrease for an image dissimilarity metric), are more commonly used as they converge faster. Note the term gradient ascent is most appropriate when maximising an image similarity metric. However, the term gradient descent is more commonly used in the literature, where the problem is often formulated as a minimisation, so this convention is followed here. Let the values for k parameters of a transformation model at the n -th iteration be denoted as a vector $\Theta^n = [\theta_1, \theta_2, \dots, \theta_k]$ and $f(\Theta)$ be a function on the parameter space that returns the image dissimilarity. $f(\Theta)$ is minimised by iteratively stepping along the inverse of the gradient ∇f . Thus, the parameter update step is given:

$$\Theta^{n+1} = \Theta^n - \delta \nabla f,$$

$$\nabla f = \left(\frac{\partial f(\Theta)}{\partial \theta_1}, \frac{\partial f(\Theta)}{\partial \theta_2}, \dots, \frac{\partial f(\Theta)}{\partial \theta_k} \right).$$

Here δ controls the step size through the parametric space. If this value is set too high then we may oscillate, skipping past the minimum, however lower values result in slower convergence. The gradient may be computed analytically for some metrics such as [SSD](#), however where an analytical solution does not exist or is costly to compute it may be estimated using a finite difference method. A number of other gradient based methods exist that aim to improve the speed

of convergence over gradient descent including conjugate gradient descent, the Levenberg–Marquardt algorithm and the Broyden–Fletcher–Goldfarb–Shanno algorithm.

2.2 IMAGE SEGMENTATION

2.2.1 *Background*

Image segmentation involves delineating the boundary of an anatomical structure by assigning an appropriate label to each image voxel. Structures of interest may be major subdivisions of the brain such as the cerebrum, cerebellum or brainstem, or tissues with different cellular compositions such as [GM](#), and [WM](#). Once a structure has been segmented, a morphological analysis may be undertaken in the region, for example estimating its volume or shape. This thesis is mainly focused on obtaining accurate segmentations of the cerebral cortex in order to extract its surface boundary and examine its shape.

Accurate segmentation may be performed manually by a trained expert, which can be both laborious and time-consuming. This is especially true for term-born infants and for fetuses at later gestational ages, where a single segmentation of the cortex may take a whole day, due to its complexity. Therefore, a robust automatic segmentation algorithm is highly desirable for large population based studies.

An [MR](#) imaging sequence may be optimized to maximise the contrast in signal intensities between different tissues such as cortical [GM](#) and [WM](#), facilitating automatic segmentation using intensity-driven algorithms. However automatic segmentation is not a straightforward task with fetal and neonatal data as the inten-

sity distributions for developing tissues change as a function of age and are also more heterogeneous. Other confounding factors with these data include image reconstruction artefacts and the partial volume (PV) effect (where multiple tissues occupy a single voxel mimicking the intensity of another tissue).

2.2.2 *Atlas-based segmentation*

In the context of image segmentation, the term “brain atlas” refers to two associated images: a template, which may be an MR brain image of an individual subject or a digitally generated image of average anatomy, and a corresponding label image defining the spatial location of anatomical regions within the template volume (Fig. 13). By registering an atlas template with the imaged anatomy of a new subject, we can automatically segment the defined regions in the new anatomy (for example by assigning the label of the nearest propagated atlas voxel). This approach is particularly useful when anatomical structures cannot be distinguished by image intensities alone. For example, while cortical GM is easily distinguished from neighbouring WM due to their inherently different cellular compositions, subdividing the cortex into smaller functional units, such as those defined by Brodmann (1909), requires prior anatomical knowledge, which a brain atlas provides. Confusingly, in the context of cross-sectional imaging studies, the terms atlas and template are used interchangeably in the literature to describe a reference anatomy with which a population of subjects may be aligned and image analysis performed (e.g. VBM). For this purpose, a digitally created average anatomical image is commonly used as it provides an unbiased reference space for analysis.

Atlas-based segmentation relies on accurate registration of an atlas to an unlabelled target subject. Unfortunately, image registration is an ill-posed problem in

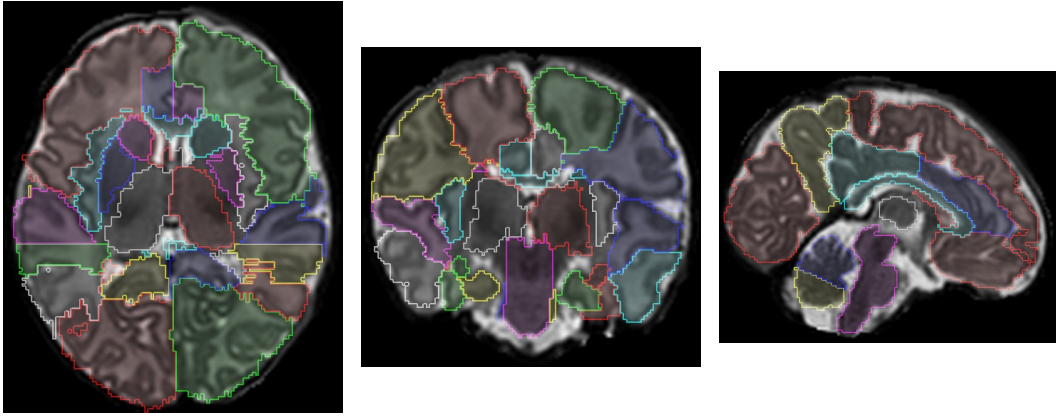


Figure 13: **Neonatal brain atlas.** The atlas shown is one of twenty from the dataset created by Gousias et al. (2012). The brain image was manually segmented into 50 regions: 32 cortical regions and 18 sub-cortical regions.

the sense that many local maxima may exist for an image similarity metric in the parametric search space. Errors in registration are, therefore, common and to overcome this, multiple atlases may be registered to an unlabelled target image and their segmentations transformed and fused to gain a consensus classification (Rohlfing et al., 2003; Heckemann et al., 2006; Aljabar et al., 2009).

Once aligned with a target image, atlas labellings may be fused using a number of voting schemes, which have proven to be effective. A majority vote, where the most frequent label at each voxel location is chosen, produces a “hard” segmentation where each voxel belongs to a single region only and boundaries are explicit between voxels. A probabilistic, or “soft”, segmentation may also be obtained where a probability map is constructed for each tissue and boundaries are not explicitly defined. The probability of a tissue may be inferred from the frequency its label occurs among the atlases at each target voxel location.

Classification may be improved by selecting the most anatomically similar atlases from an atlas set, estimated by an image similarity metric (Aljabar et al., 2009). Image similarity may also give a weighting for label fusion on a global, local (Artaechevarria et al., 2009) and also non-local basis (Coupé et al., 2011). A local

weighting can improve classification accuracy by down-weighting an atlas vote where a localised registration error has occurred or if local anatomy is highly dissimilar. Local weightings may be derived from the image similarity between the atlas templates and a target image, in a 3D patch around a voxel which is being classified. [Coupé et al. \(2011\)](#) proposed a non-local, many-to-one, patch-based method for atlas fusion. In this work, numerous voxel labels from each atlas are taken from a local neighbourhood around a target voxel and fused to give a classification. Chosen atlas labels within the neighbourhood are weighted according to their patch similarity with the target patch around the voxel being classified.

Another notable method for fusing atlas labels is STAPLE ([Warfield et al., 2004](#)). This method iterates between two steps: generating a consensus segmentation and estimating the reliability of each classifier (i.e. atlas). The reliability of each classifier is determined by its sensitivity and specificity and is used to weight its contribution when constructing the consensus segmentation. The algorithm iterates between these two steps until convergence.

2.2.3 Probabilistic atlas

A probabilistic atlas is an average model constructed from a population of subjects and consists of an average image template and a set of tissue probability maps. An average template ideally provides an unbiased target for aligning a group of subjects and a coordinate system for conducting image analysis (e.g. for a [VBM](#)-based study). Additionally, a probabilistic atlas is readily combined into a Bayesian framework for automatic segmentation (Section [2.2.4](#)). Aligning the tissue probability maps with an unseen subject gives a prior estimate of tissue

location that can be refined using the intensities of the new image or additional knowledge of tissue morphology (e.g. shape, topology and connectedness).

An image template is constructed by averaging the image intensities for a group of subjects. Images may be co-aligned using a target-free approach (Seghers et al., 2004; Kuklisova-Murgasova et al., 2011; Serag et al., 2012a), where no single image is used as a target for alignment (Fig. 14). In this approach, all images are registered pairwise and a transformation into an “average space” is computed for each image, by averaging its transformations to every other image. An image template is then constructed by averaging the intensities of all the images once transformation into the “average space”. Brain structures in the resulting template are spatially normalized, i.e. average in both shape and size for the group, as they are not biased towards any individual subject.

During gestation and early childhood, the brain undergoes rapid morphological changes. The size and shape of structures can change significantly in a matter of weeks and a single image template cannot represent the diverse anatomies examined in this thesis. Previous authors (Habas et al., 2010; Kuklisova-Murgasova et al., 2011; Serag et al., 2012b;a) have addressed this issue by constructing spatio-temporal atlases (Fig. 15,16) that consist of multiple age-specific image templates, each with a set of associated tissue probability maps. Each image template is constructed from subjects within a time-window where anatomies are similar. Modelling the rapid change in fetal and neonatal cohorts requires separate templates for each week of gestation. Additionally, specific tissue probability maps may be necessary for different periods of gestation to map out transient structures and also newly formed structures. For example, between between 8 and 28 weeks, the enormous rate of proliferation of neurons and glial cells in the subventricular zone results in a hypointense signal in T2 *in utero* MR images (Habas et al., 2010). These cells later migrate through the intermediate zone to the cortex and contribute to

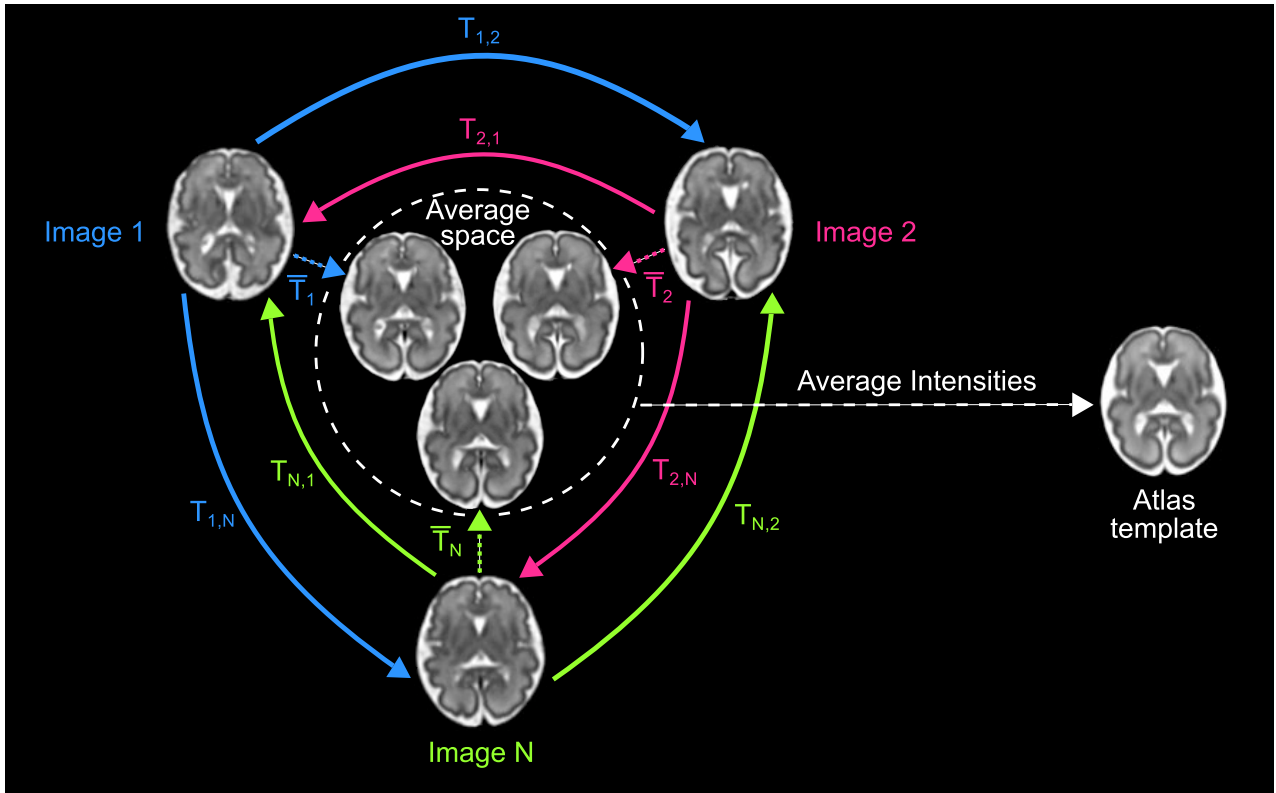


Figure 14: **Target-free atlas template construction.** For N images, each image k is registered to every other image j . The average of all transformations $T_{k,j}$ gives a transformation into an “average space” for image k . In this space, the intensities of the images can be averaged to create an unbiased atlas template.

the massive increase in cortical surface area. Another example includes the deep grey nuclei, which are barely visible at 25 weeks but become much more distinct at around 30 weeks. In this thesis we make use of two spatio-temporal atlases for image segmentation and analysis, one constructed from a population of 80 fetuses (Fig. 15) and another constructed from a population of 204 preterm neonates (Fig. 16)

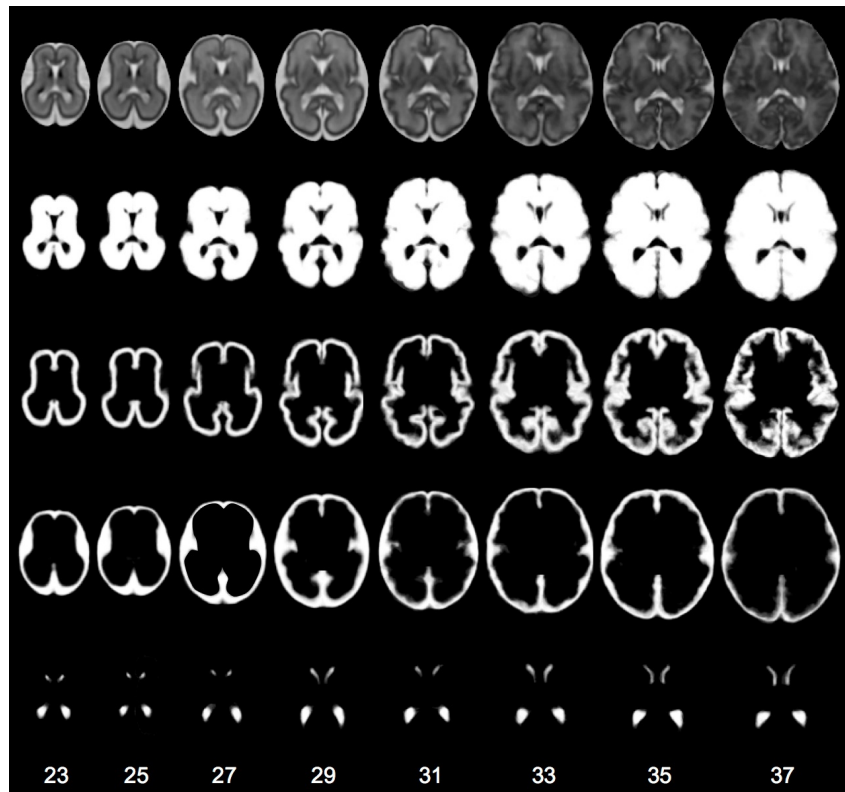


Figure 15: **Fetal spatio-temporal atlas.** This atlas was developed by [Serag et al. \(2012b\)](#) and is available at brain-development.org. The first row show image templates for T2 weighted images for several gestational ages. The remaining rows show greyscale probability maps for the cerebrum, GM, CSF and the lateral ventricles (top to bottom).

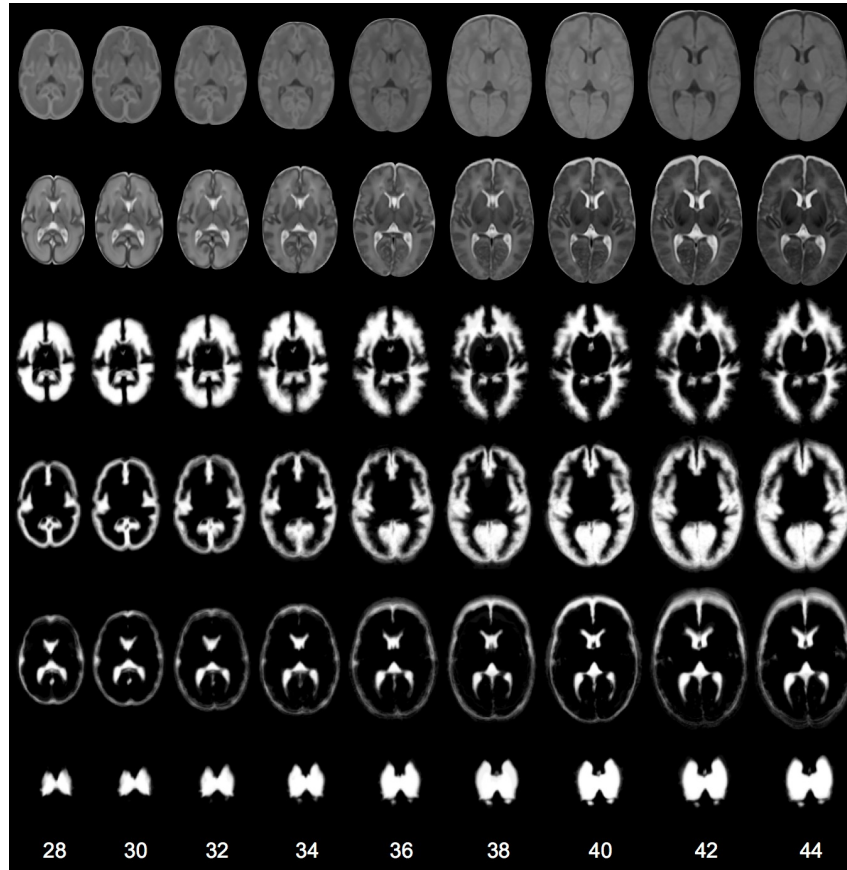


Figure 16: **Neonatal spatio-temporal atlas.** This atlas was developed by [Serag et al. \(2012a\)](#) and is available at brain-development.org. The first two rows show image templates for T1 and T2 weighted images for several gestational ages. The remaining rows show greyscale probability maps for WM, GM, CSF and subcortical GM (top to bottom).

2.2.4 Expectation-Maximisation Segmentation

2.2.4.1 Background

In this thesis we employ variants of an expectation-maximisation algorithm ([Ledig et al., 2012](#); [Makropoulos et al., 2014](#)) for accurate tissue segmentation that was first proposed by ([Wells et al., 1996](#)). This algorithm automatically segments a

brain image into several tissue classes, using prior knowledge of tissue location, by building specific intensity models for each tissue class. This technique was originally developed for adult imaging data, however, a number of authors have applied this approach to fetal and neonatal imaging data (Prastawa et al., 2005; Weisenfeld et al., 2006; Bach Cuadra et al., 2009; Habas et al., 2010; Cardoso et al., 2011; Ledig et al., 2012; Makropoulos et al., 2014), which will be discussed later on in this section.

Each iteration of the algorithm is composed of two steps: classifying image voxels and estimating the intensity distribution of each tissue class. To begin the process, an initial classification is obtained from a set of tissue probability maps associated with a probabilistic atlas.

2.2.4.2 Estimating tissue intensity models

Let the intensities of an MR image with N voxels be denoted as a 1D vector $y = [y_1, y_2, \dots, y_n]$ and its segmentation be denoted $z = [z_1, z_2, \dots, z_n]$. For K tissue labels enumerated $1 \dots K$, let $z_i = e_k$ indicate that label k has been assigned to voxel i , where e_k is a $K \times 1$ unit vector with the k th element equal to one. Let the prior tissue probabilities for the i -th voxel be denoted π_i , which is also a $K \times 1$ vector, where the k th element gives the prior probability for tissue class k , i.e. $P(z_i = e_k) = \pi_{ik}$. The intensity distribution of each tissue class is modelled by a Gaussian distribution parametrised by a mean μ_k and a standard deviation σ_k . Let the complete parameter set be denoted $\Phi_y = [\mu_1, \mu_2, \dots, \mu_K, \sigma_1, \sigma_2, \dots, \sigma_K]$.

At each iteration m , a classification or expectation step is performed (Eq. 1). The probability p_{ik}^m of tissue k being present at the i -th voxel, is computed given its intensity y_i and the current model parameters $\Phi_y^{[m-1]}$. An application of Bayes' rule allows this computation to be reformulated in terms of the prior probability of

tissue k being present at the i -th voxel, $P(z_i = e_k)$, and the conditional probability of the observed voxel intensity, given a tissue classification k and the current model parameters, $P(y_i|z_i = e_k, \Phi_y^{[m-1]})$. Note that the prior probabilities may be taken as an initial classification, i.e. $p_{ik}^0 = P(z_i = e_k) = \pi_{ik}$. Thus, the exception step is given:

$$p_{ik}^m = P(z_i = e_k|y_i, \Phi_y) = \frac{P(y_i|z_i = e_k, \Phi_y^{[m-1]})P(z_i = e_k)}{\sum_{j=1}^K P(y_i|z_i = e_j, \Phi_y^{[m-1]})P(z_i = e_j)} \quad (1)$$

Here the probability of a voxel intensity y_i , given a tissue classification k , is given by a Gaussian probability density function G :

$$P(y_i|z_i = e_k, \Phi_y^{[m-1]}) = G(y_i|\mu_k, \sigma_k) = \frac{1}{\sigma_k \sqrt{2\pi}} \exp\left(-\frac{(y_i - \mu_k)^2}{2\sigma_k^2}\right).$$

A maximisation step follows that updates the tissue model parameters to give $\Phi_y^{[m]}$. Tissue means and variances are calculated by weighting the influence of each voxel by their classification probabilities:

$$\mu_k^{[m]} = \frac{\sum_{i=1}^N p_{ik}^{[m]} y_i}{\sum_{i=1}^N p_{ik}^{[m]}}. \quad (2)$$

$$(\sigma_k^{[m]})^2 = \frac{\sum_{i=1}^N p_{ik}^{[m]} (y_i - \mu_k^{[m]})^2}{\sum_{i=1}^N p_{ik}^{[m]}}. \quad (3)$$

2.2.4.3 Markov random field regularisation

Although MRI acquisition is typically optimized so that tissue classes such as GM, WM and Cerebrospinal Fluid (CSF) have as much contrast between them as possible, voxels may be mislabelled for a number of reasons including image noise,

reconstruction artefacts and the [PV](#) effect. Additionally, the intensities of other tissues close to the brain, such as the skull and meninges (and also maternal tissue for the fetus), may have a similar intensity to brain tissue and may be erroneously labelled as such.

Each voxel is classified independently in the core [EM](#) algorithm outlined above, therefore tissue topology, connectedness and adjacency are ignored. By incorporating spatial dependence into the prior term of the classification model, a regularisation effect can be achieved, thus improving the accuracy and plausibility of tissue classifications. An Markov Random Field ([MRF](#)) can be used to achieve this goal ([Van Leemput et al., 1999](#)). The [MRF](#) formulation presented here (Eq.4) is a multi-class extension of the Potts model from ([Cardoso et al., 2013](#); [Makropoulos et al., 2014](#)). The prior probability of voxel i having tissue class k is dependent on the tissue classes of its 6 first-order neighbours N_i (i.e. face neighbours).

$$P(z_i = e_k) = \frac{\pi_{ik} \exp(-U_{MRF}(e_k|N_i))}{\sum_{j=1}^K \pi_{ij} \exp(-U_{MRF}(e_j|N_i))}. \quad (4)$$

The [MRF](#) energy function U_{MRF} is defined:

$$U_{MRF}(e_k|N_i) = \sum_{j=1}^K A_{kj} v_j \quad (5)$$

where

$$v_j = \left(\sum_{l \in N_i^x} s_x p_{lj} + \sum_{l \in N_i^y} s_y p_{lj} + \sum_{l \in N_i^z} s_z p_{lj} \right). \quad (6)$$

Here, $s = \{s_x, s_y, s_z\}$ is the anisotropic voxel resolution; \mathbf{A} is a $K \times K$ matrix, where A_{kj} defines the prior connectivity strength of tissues k and j ; and \mathbf{v} is a vector of size K , where v_j is the summed probability of class j within the neighbourhood.

This energy function can be used to independently penalise both small disconnected regions and implausible tissue adjacencies in the classification. A_{kj} may be set to a relatively high value for non-adjacent tissues, for example **WM** and extra-cerebral **CSF**, decreasing the prior probability of neighbourhood configuration involving these tissues. For adjacent structures, e.g. **GM** and **WM**, A_{kj} may be set to a relatively low value. Where k and j are the same tissue, A_{kj} is usually set to zero. In this instance, if the neighbours of a voxel i all have the same tissue k , then the prior probability of voxel i also having the same label is given directly by the atlas, i.e. π_{ik} .

2.2.4.4 *Partial volume effect*

The **PV** effect occurs at voxels that sample the boundary between two or more tissues, where the signal intensity of the tissues is combined, resulting in an intermediate signal intensity. The early **EM** approaches of (Wells et al., 1996; Van Leemput et al., 1999) did not address this problem and assumed that a single tissue is present at each voxel. This is a major confounding factor for the **EM** segmentation algorithm and other methods that use intensity models for tissue classification, when applied to both fetal and neonatal data (Fig. 17). For these data, T2-weighted images are primarily used for image analysis as they have greater tissue contrast compared with T1-weighted images. For T2-weighted images, **GM** has the lowest intensity, **WM** has an intermediate intensity and **CSF** has the highest intensity. At the cerebral boundary the intensities of **GM** and **CSF** are combined, resulting in a moderate intensity similar to **WM**. This effect is particularly apparent in the deep

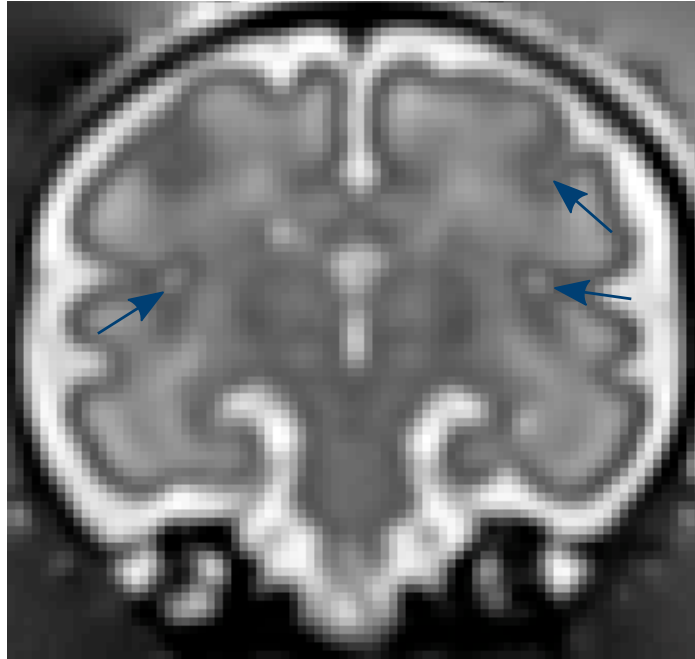


Figure 17: **Partial volume effect.** When a tissue boundary intersects a voxel, the signal intensities of the adjacent tissues are combined and an intermediate intensity is observed. This is a major confounding factor for segmentation of fetal MR images, such as the one shown. For voxels at the cerebral boundary, the signal from GM and CSF combines, resulting in a moderate intensity similar to that of WM. Accurate segmentation is particularly challenging deep within sulci, where small amounts of CSF are present (blue arrows).

narrow sulci of older fetuses and neonates. This problem is exacerbated by the lower image resolution of neonatal and fetal datasets.

After birth, the signal intensities of WM and GM change progressively until the age of two years and appear to reverse in intensity (Hillenbrand and Huisman, 2012). This is mainly due to progressive myelination of the WM, a process where a fatty white substance forms a sheath around neuronal axons, providing electrical insulation, which increases the speed of signal transmission. After two years the GM/WM intensity distributions remain similar into adulthood. In contrast to fetal and neonatal imaging, T1-weighted images are typically acquired for adult brain image analysis as they have better GM/WM tissue contrast. In these images, WM has the largest intensity followed by GM and then CSF. In this case, a PV voxel at the

cortical boundary will not replicate the signal of another tissue, thus the PV effect is a lesser confound. When segmenting this data, GM/WM PV voxels are usually be labelled as GM, in the absence of image noise due to the greater variance in its intensity distribution (Van Leemput et al., 1999), resulting in an overly thick cortical segmentation.

A number of authors have had some success tackling this problem by explicitly modelling GM/WM PV voxels as a separate class in the EM framework (Ruan et al., 2000; Van Leemput et al., 2003; Bach Cuadra et al., 2009; Cardoso et al., 2013). First an initial segmentation is performed, then an additional PV class is introduced that models the mixture of the two tissues before continuing. Its mean may be estimated as the average of the individual class means weighted by their variance, $\mu_{GM/CSF} = (\sigma_{GM}\mu_{GM} + \sigma_{CSF}\mu_{CSF})/(\sigma_{GM} + \sigma_{CSF})$, and its prior probability may be given by the geometric mean of the individual posterior tissue distributions, $\omega_{i(GM/CSF)} = \sqrt{p_{i(GM)}p_{i(CSF)}}$ (Cardoso et al., 2013). The WM prior probability may also be reduced in regions where GM and CSF are both likely *a priori*, $\omega_{i(WM)} = \omega_{i(WM)}(1 - \sqrt{\omega_{i(GM)}\omega_{i(CSF)}})$. Note the priors must be normalised after these steps so that they to sum one.

Xue et al. (2007) took a different approach to this problem using knowledge based rules to adjust the prior tissue probabilities from Eq.4. If a voxel has been classified as WM and it has neighbours that have been classified as CSF and also GM, this voxel is likely to be a partial volume voxel. In this instance, the prior probability of WM was penalised and redistributed to the GM and CSF priors. Additionally, small clusters of voxels that have been classified as WM and are mainly surrounded by CSF are likely to be partial volume deep within a sulci. In this instance, the prior probability of WM was penalized and redistributed to CSF only. This approach was found to be effective at reducing the number of misclassifications in a dataset of 25 neonatal images.

2.2.4.5 *Adopted methods*

In this thesis, two EM segmentation algorithms are employed for brain tissue segmentation (Ledig et al., 2012; Makropoulos et al., 2014). Both of these methods explicitly tackle the problem of mislabelled PV voxels in neonatal and fetal datasets. Ledig et al. (2012) employed a similar strategy to Xue et al. (2007) of penalising WM which is adjacent to both CSF and GM, but using a second order MRF field instead of knowledge-based rules.

$$P(z_i = e_k | N_i) = \frac{\pi_{ik} \exp(-U_{MRF}(e_k | N_i)) \exp(-U_{2nd}(e_k | N_i))}{\sum_{j=1}^K \pi_{ij} \exp(-U_{MRF}(e_j | N_i)) \exp(-U_{2nd}(e_j | N_i))}, \quad (7)$$

where U_{2nd} is a new energy function:

$$U_{2nd}(e_j | N_i) = \mathbf{v} \mathbf{T}_k \mathbf{v}.$$

Here \mathbf{v} is a vector of size K of neighbourhood probabilities (Eq. 6) and \mathbf{T}_k is a $K \times K$ penalty matrix. If $\mathbf{T}_{kjl} > 0$, a penalty is applied to class k when there is a high probability of class j and l within the neighbourhood of i .

By penalising WM when it is implausibly adjacent to both cortex and CSF, segmentation accuracy is improved at the cortical boundary (Fig. 18).

Makropoulos et al. (2014) proposed a novel hierarchical modelling of tissue classes and adopted the method of Xue et al. (2007) to correct for mislabelled PV voxels. An extension to this algorithm described in (Makropoulos, 2014, ch. 6) is also employed, which makes use of Laplace's equation to estimate cortical thickness (Jones et al., 2000) and correct segmentation anomalies. For example, in deep sulci, the

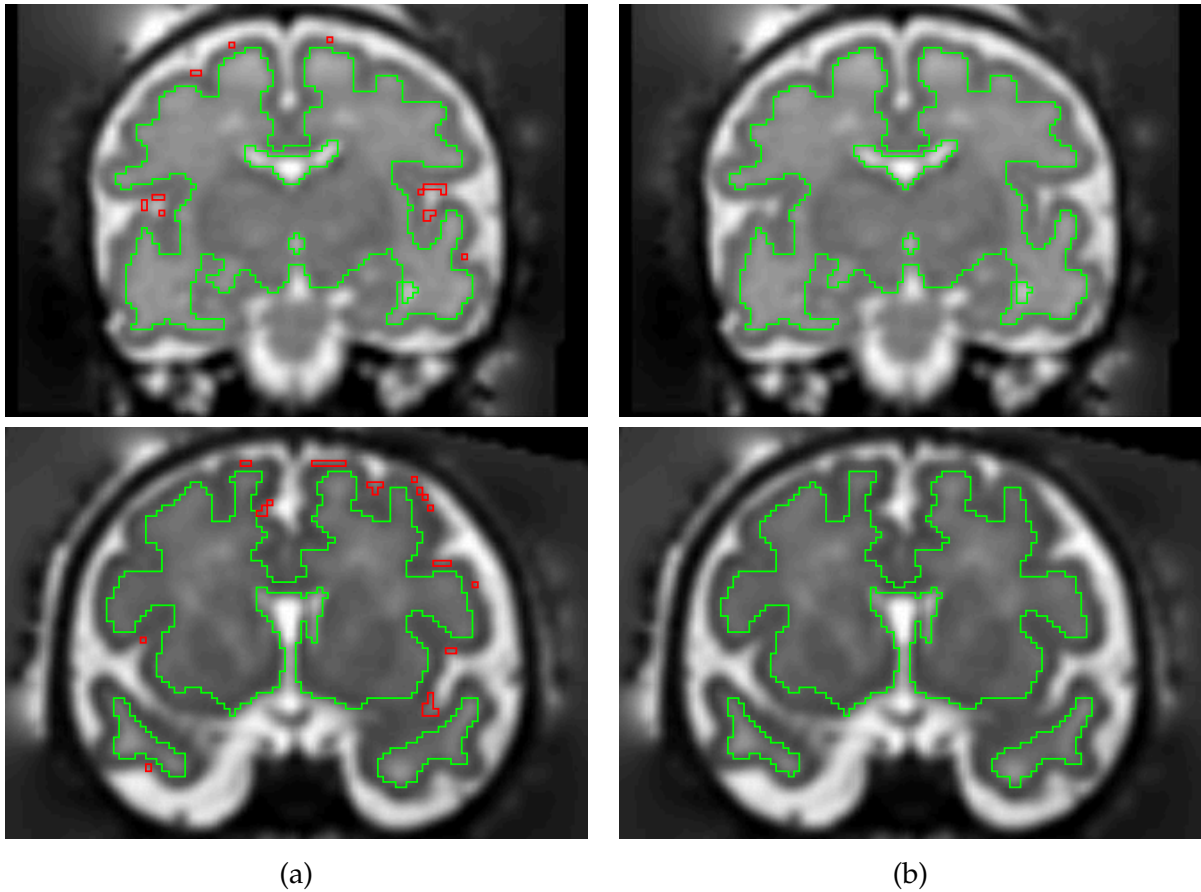


Figure 18: **Improved cortical segmentation using a second order MRF.** Two subjects are shown, each segmented using a standard EM-MRF algorithm (a) and an EM algorithm incorporating a second order MRF (b). Voxels correctly labelled as WM are shown in green while PV voxels mislabelled as WM are shown in red. Note the reduction in mislabelled voxels in (b).

small amount of CSF present here means that PV voxels have a similar intensity to GM. In this instance PV voxels may be mislabelled as GM, resulting in an overly thick cortical segmentation.

2.3 SURFACE MODELLING

Obtaining accurate representations of the cortical geometry is highly desirable in order to investigate folding patterns. A tissue boundary may be inferred from a probabilistic segmentation as an isosurface where the probability equals 0.5. Extracting a mesh representation of an isosurface from volumetric data is a well-studied problem and a number of general purpose algorithms exist for this purpose (Lorensen and Cline, 1987; Nielson and Hamann, 1991; Gibson, 1998; Treece et al., 1999; Ju et al., 2002). However, these algorithms have no prior knowledge of cortical topology and may extract implausible surface representations that have handles or holes (Fig. 19). Correct topology is important, for example, in order to determine correspondences between surfaces. The cortical boundary is a thin continuous sheet of neurons lining the cerebrum, with no holes or handles, and is thus topologically equivalent to a sphere. This knowledge can be applied to extract topologically correct surfaces.

Previous researchers (Habas et al., 2012; Clouchoux et al., 2012) have focused on delineating the inner cortical boundary, which is extracted more easily. However, the extraction of both outer and inner boundaries is necessary to estimate cortical thickness, which has been shown to be an important parameter for the characterisation of neurodevelopmental psychiatric disorders such as attention deficit hyperactivity disorder Shaw et al. (2006). At later gestations, where the cortex is more convoluted, neighbouring gyri may squeeze against each other, occupying

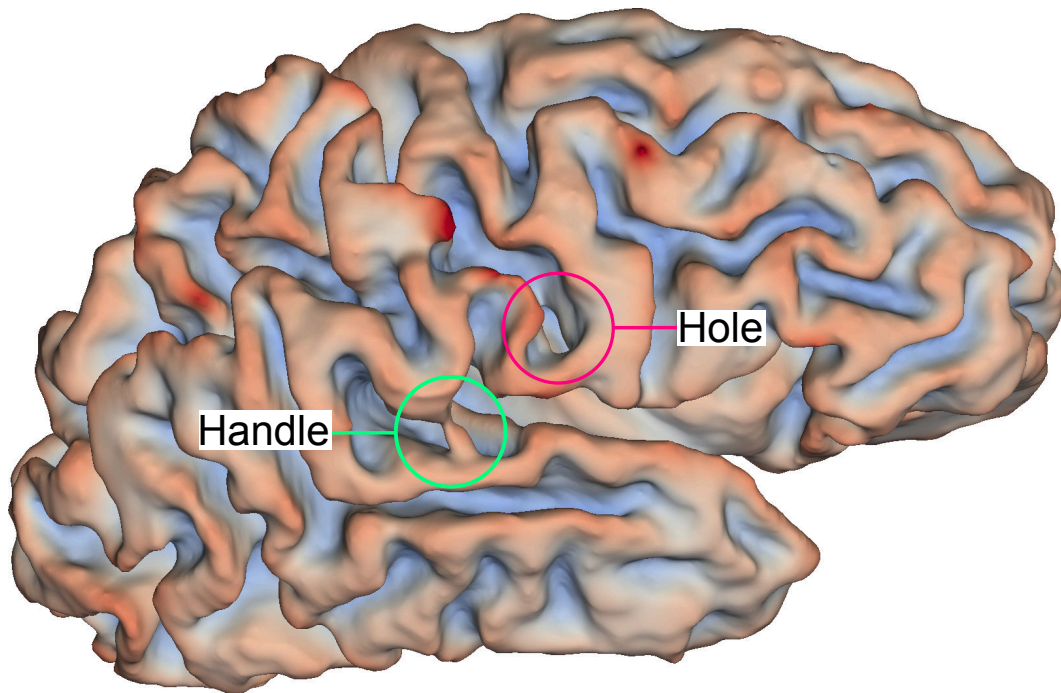


Figure 19: **Topological errors.** This figure shows the inner cortical surface of a neonatal brain, extracted from a WM segmentation using the Marching Cubes algorithm. Two errors are shown which alter the topology of the mesh from that of a sphere. The green circle shows an implausible bridge across the sylvian fissure, caused by a segmentation error, whereas the pink circle shows a hole within the central sulcus, where a relatively thin layer of WM was present. Mean curvature is colour mapped onto the surface to aid visual perception of the cortical geometry (red: convex, blue: concave).

neighbouring voxels in the image data. In these instances, segmented gyri may become connected within an isosurface boundary (Fig. 20). A similar, although less common, problem exists when extracting the GM/WM cortical boundary. When WM becomes relatively thin, for example inside a tight gyrus, the segmented cortical GM becomes connected inside a gyrus, creating a hole in the WM isosurface (Fig. 19).

A number of authors have developed techniques to correct topology errors (handle/hole pairings) (Fischl et al., 2001; Ségonne et al., 2005; Ségonne et al., 2007; Yotter et al., 2011). A hole may be detected by mapping a cortical surface mesh onto a sphere, either using spherical harmonics (Yotter et al., 2011) or through a

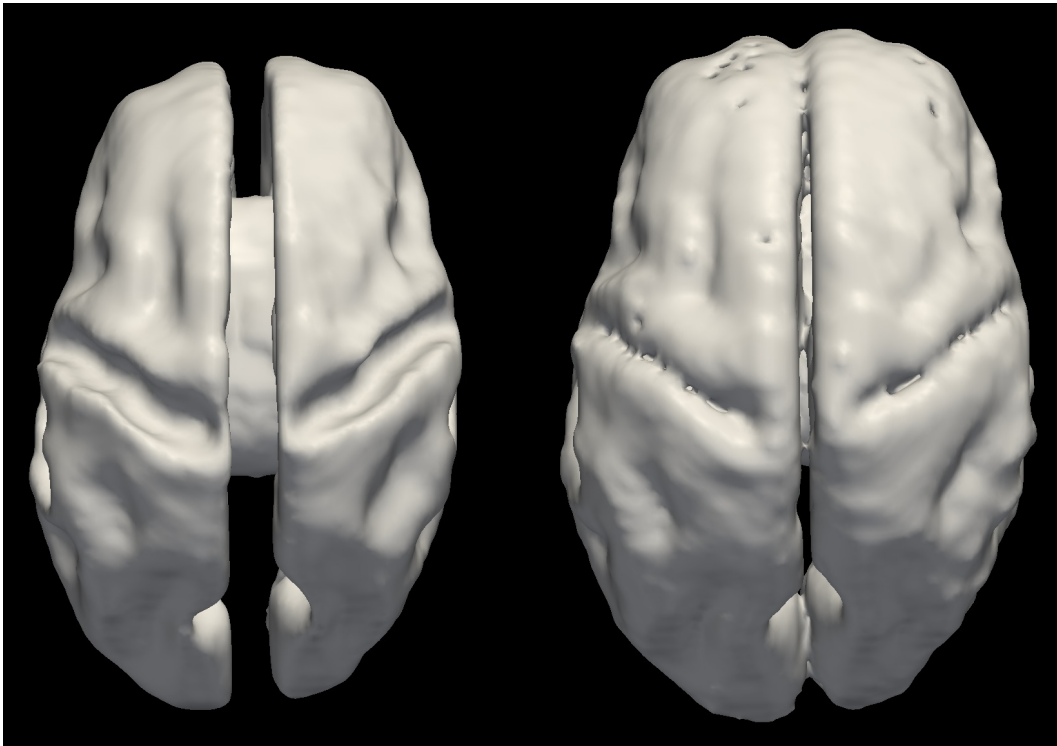


Figure 20: **Inner and outer cortical surfaces extracted using the Marching Cubes algorithm.** Note the merging of the pre- and post-central gyri and also the occipital and frontal lobes.

process of inflation (Fischl et al., 2001). Edge intersection in the spherical space indicates a topological error. Fixing an error requires either filling a hole or removing a handle. A trained operator may easily determine the correct fix manually but automating this process is more challenging.

In (Fischl et al., 2001), the decision is made implicitly using a greedy algorithm. Intersecting edges and their vertices are first removed from the mesh. Potential edges are ordered according to some metric that assess their suitability. Retessellation then involves adding the potential edges sequentially if they do not intersect the current tessellation. This method produces reasonable results, however, the plausibility of a retessellation can only be assessed in its entirety and that the “goodness of fit” of an edge does not exist in isolation (Ségonne et al., 2005).

Ségonne et al. have proposed two methods to tackle this problem, both of which select the best retessellation from a set of candidates in a Bayesian framework. In (Ségonne et al., 2005), a genetic algorithm is used to generate candidate solutions. The goodness of a solution is measured by two properties: the smoothness of the retessellation and the MRI intensity inside and outside of the surface. At each iteration the population of solutions increased, generating new solutions through “mutation” (random modification of a solution) and “crossover” (combining two solutions to create another). The best solutions are then kept for the next iteration and consequently the “fitness” of the population increases over a number of iterations until convergence. In (Ségonne et al., 2007), non-separating loops (Guskov and Wood, 2001) are used to generate candidate solutions instead. Using this methodology results in a more diversified set of candidates, thus improving tessellation.

An alternative strategy to fixing a topologically incorrect mesh, involves deforming a mesh with the correct topology (such as a sphere) towards a cortical boundary (Dale and Sereno, 1993; Davatzikos and Bryan, 1996; MacDonald et al., 2000; Kim et al., 2005). Deformation may be guided by image intensities, tissue segmentations or geometric smoothness, or a combination of these. This has the advantage of guaranteeing the correct topology, however, the global topology constraint may cause large geometric inaccuracies where segmentation or image reconstruction errors occur. Some form of volume preprocessing, such as a connected components analysis, may therefore be desirable.

In this thesis, the strategy of deforming a spherical mesh was adopted for extracting an explicit representation of the cortical surface. This was made possible by the high quality segmentations obtained using the algorithm of Makropoulos et al. (2014), which was used to drive the extraction process. A full description of this

process is given in Chapter 4 (Section 4.2.1), which makes use of surface meshes to construct a spatio-temporal cortical surface atlas.

2.4 CONCLUSION

This chapter has presented a review of some fundamental neuroimage analysis tasks and described the chosen algorithms that this thesis depends on. We have seen how image registration can be used to align imaged anatomies on a coarse and fine scale, using linear and non-rigid techniques. This can be used, for example, as a mechanism to propagate tissue labels for an individual subject to another unlabelled subject. An unbiased atlasing method was described to construct an average image template and tissue probability maps, by first aligning a group of anatomies. This type of atlas is particularly useful as it provides prior knowledge of anatomy in a Bayesian segmentation framework, such as an EM-based segmentation algorithm. Confounding factors specific to fetal and neonatal datasets were also examined, such as misclassification of PV voxels, along with methods from the literature that deal with these. Finally, algorithms that reconstruct topologically correct cortical surface models were also reviewed.

The next two chapters (3 & 4) will discuss techniques for analysing cortical folding that build upon the methods presented here. The literature discussed in these chapters will be specific to the methodological contributions within. In the next chapter, a framework developed for this thesis to quantify gyrification will be outlined and applied to a fetal dataset (see Section 1.4 for dataset details) and in Chapter 4, another framework, developed for this thesis, for constructing a spatio-temporal surface atlas will be presented and applied to the same dataset.

AUTOMATIC QUANTIFICATION OF CORTICAL FOLDING

3.1 INTRODUCTION

In this chapter, a complete image analysis framework for quantifying cortical folding is described and applied to a normal fetal dataset ($N = 80$, details of this dataset are given in Section 1.4). Cortical folding is quantified over a large gestational age range of 21.7 to 38.9 weeks, extending our understanding of how the cortex folds through this critical developmental period.

Zilles et al. (1988) proposed the Gyrification Index, one of the first popular measures for cortical folding quantification that is independent of scale. This slice-based measure is defined as the ratio of the length of two contours: the outer cortical boundary and a contour skirting the exposed cerebrum (i.e. not entering deep folds). Unfortunately, this measure is sensitive to slice orientation, and additionally it cannot be applied to different subregions of the brain.

More recent folding measures have made use of curvature-based surface descriptors. The curvature of a curve at a surface point p is defined as the reciprocal of the radius of a circle passing through p and a pair of additional points on the curve infinitesimally close to p . The curvature can be measured along any plane intersecting the surface which contains the surface normal. The principal curvatures k_1 and k_2 are the maximum and minimum curvatures for a surface point p

(i.e. $k_1 \leq k_2$). The principal curvatures give rise to two classical shape descriptors: mean curvature (H) and Gaussian Curvature (K).

$$H = (k_1 + k_2)/2 \qquad K = k_1 k_2$$

Each of these gives a description of local shape at a surface point. Convex and concave regions have positive and negative mean curvature respectively, while elliptical and hyperbolic points have positive and negative Gaussian curvature respectively. Koenderink and van Doorn proposed two additional shape descriptors, that are complimentary to each other: curvedness (C) and shape index (S).

$$C = \sqrt{(k_1^2 + k_2^2)/2} \qquad S = \frac{2}{\pi} \arctan \frac{k_2 + k_1}{k_2 - k_1}$$

Curvedness is closely related to the bending energy of a surface when deformed from a flat plate (van Vliet and Verbeek, 1993), while values of shape index can be used to distinguish 9 different surface features: spherical cup, trough, rut, saddle rut, saddle, saddle ridge, ridge, dome and spherical cap.

Integrating shape descriptors over a cortical surface provides the basis for a number of folding measures in the literature, including folding index (FI) (Van Essen and Drury, 1997), intrinsic curvature index (K_I) (Van Essen and Drury, 1997), global mean curvature (H_G) (Magnotta et al., 1999), global Gaussian curvature (K_G) (Magnotta et al., 1999) L^2 norm of mean curvature (H_N) (Batchelor et al., 2002), L^2 norm of Gaussian curvature (K_N) (Batchelor et al., 2002), global curvedness (C_G) (Awate et al., 2008; Rodriguez-Carranza et al., 2008) and global shape index (Awate et al., 2008; Rodriguez-Carranza et al., 2008). Formulations of these measures are shown in Table 1.

A useful folding measure must be invariant to position, orientation, scaling and surface area (Batchelor et al., 2002). Invariance to brain size is particularly important when measuring folding in the developing brain as brain volume increases

| | |
|----------------------------------|---|
| L^2 norm of mean curvature | $H_N = \frac{1}{4\pi} \sqrt{\int_A H^2 dA}$ |
| L^2 norm of Gaussian curvature | $K_N = \frac{1}{4\pi} \sqrt{A \int_A K^2 dA}$ |
| Intrinsic curvature index | $K_I = \frac{1}{4\pi} \int_A K dA$ where $K > 0$ |
| Folding index | $FI = \frac{1}{4\pi} \int_A k_1 (k_1 - k_2) dA$ |
| Global mean curvature | $H_G = \frac{1}{A} \int_A H dA$ |
| Global Gaussian curvature | $K_G = \frac{1}{A} \int_A K dA$ |
| Global curvedness | $C_G = \frac{1}{A} \int_A C dA$ |
| Global shape index | $S_G = \frac{1}{A} \int_A S dA$ |

Note A is the surface area of the region of interest.

Table 1: **Summary of folding measures.**

substantially throughout gestation. Unfortunately, some of the measures discussed so far are directly dependent on the surface area over which they are computed and consequently a meaningful comparison of folding measures between subregions with different sizes is prohibited. [Rodriguez-Carranza et al., 2008](#) proposed two normalization factors to address this issue, one derived from cerebral surface area and volume, $T = 3V/A$, and the other the absolute global mean curvature H_G .

3.2 MATERIAL AND METHODS

3.2.1 Overview

A spatio-temporal atlas was used to automatically segment images into a number of key tissue types using an [EM](#) based segmentation algorithm that specifically corrects mislabelled [PV](#) voxels ([Ledig et al., 2012](#)). An implicit cortical surface representation is formed on a super-sampled voxel grid and principal curvatures are evaluated over the surface voxels using the structure tensor and Hessian ma-

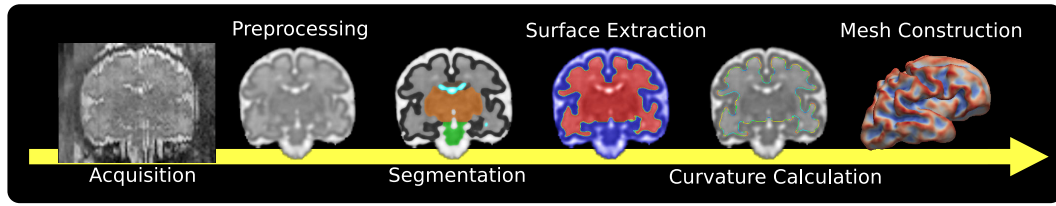


Figure 21: **Processing pipeline overview.** Multiple acquisition loops consisting of 2D snapshots are reconstructed into a complete 3D volume using a slice-to-volume reconstruction algorithm. Then a number of preprocessing steps are carried out including bias correction and brain extraction. Automatic segmentation is performed using an EM-based algorithm, delineating seven anatomical regions. A probability map defining matter inside the inner cortical boundary is then constructed by summing the posterior probability maps of subcortical structures. Principal curvatures are evaluated at each boundary voxel location and folding measures are computed to quantify folding. A mesh representation of the surface is also extracted using the marching cubes algorithm in order to visualize the computed curvature-based descriptors.

trix (Rieger et al., 2004; Rodriguez-Carranza et al., 2008). Eight curvature-based folding measures from the literature were adapted to quantify the degree of gyrification of the subjects within the cohort. Local gyrification was also quantified in nine different regions by parcellating each of the image volumes using an anatomical atlas. A Gompertz function was used to model the relationship between the folding measures and gestational age, which subsequently allowed an accurate prediction of physiological age. See Fig. 21 for an overview of the pipeline.

3.2.2 Automatic Brain Extraction

Restricting a segmentation algorithm to a region of interest bounded by the skull is useful to avoid misclassifying extra-cranial matter as brain tissue. With adult and infant subjects, the strong signal contrast between the skull and intra-cranial matter can facilitate automatic extraction of the intra-cranial region. For example, the Brain Extraction Tool (BET) (Smith, 2002) deforms a sphere, which is initialised inside the brain, one vertex at a time until the surface coincides with superficial

cerebral boundary, driven only by the image intensities. With fetal images this approach often fails, as the boundary of the skull is less distinct due to the maternal tissue immediately surrounding the head of the fetus.

The patch-based approach of (Eskildsen et al., 2012), which incorporates prior knowledge of brain location in the form of atlases, was applied and found to accurately extract fetal brains. This approach relies on the assumption that a similar patch of voxels in an input image and atlas image are likely to be centred on the same tissue. This gives a mechanism for non-local, many-to-one, label propagation. By aligning an input image with an atlas, the search for similar patches can be restricted to a local neighbourhood, which is likely to be sufficient, given the limited variability in both shape and location of the skull. Numerous similar atlas patches form a labelling consensus for each voxel in an input image, with each atlas label weighted by the similarity of its surrounding patch and input image patch, ensuring the most similar patches are the most influential (Fig. 22).

Formally, an image voxel x_i is labelled by taking a cubic patch P_i centred at x_i and computing the similarity to all patches $P_{a,j}$ of a single atlas a in a local cubic neighbourhood η_i of x_i . Firstly, patch similarity was compared using the Structural Similarity (SS) metric (Wang et al., 2004): where μ is the patch mean and σ is the patch standard deviation (Equation 8). Atlas patches with a $SS \leq 0.95$ were discarded.

$$SS = \frac{2\mu_i\mu_{a,j}}{\mu_i^2 + \mu_{a,j}^2} \times \frac{2\sigma_i\sigma_{a,j}}{\sigma_i^2 + \sigma_{a,j}^2}. \quad (8)$$

A vote was taken from the remaining atlas patch labels $l(x_{a,j})$ weighted by their similarity to the target patch, $w(P_i, P_{a,j})$, to obtain a label $L_{a,i}$ for atlas a :

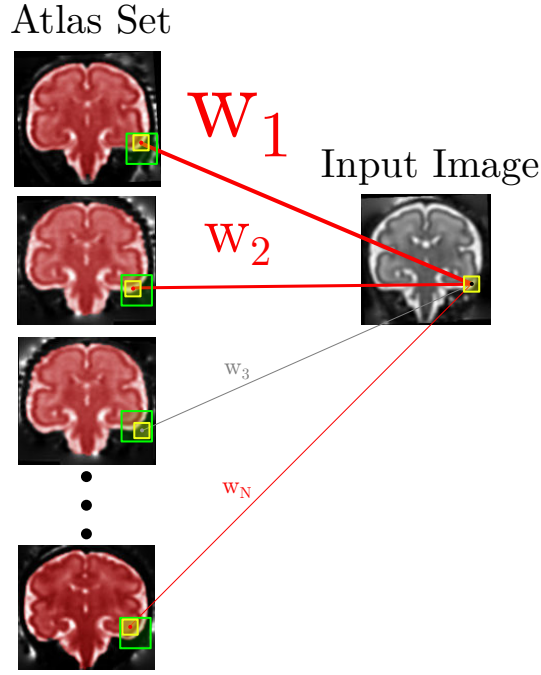


Figure 22: **Patch search and weighting.** The input image and atlas images are co-aligned so that similar patches (yellow) can be found within a local neighbourhood (green) around an input voxel. The patch similarity can then give a weighting when fusing the atlas labels.

$$L_{a,i} = \left\lfloor 0.5 + \frac{\sum_j w(P_i, P_{a,j}) l(x_{a,j})}{\sum_j w(P_i, P_{a,j})} \right\rfloor. \quad (9)$$

$$w(P_i, P_{a,j}) = \exp\left(\frac{-SSD(P_i, P_{a,j})}{h_a}\right). \quad h_a = \min(SSD(P_i, P_{a,j})). \quad (10)$$

Finally, after computing a binary labelling $L_{a,i}$ using each atlas in turn, a majority vote was taken between all atlases to reach a final binary labelling \hat{L}_i :

$$\hat{L}_i = \left\lfloor 0.5 + \frac{1}{N} \sum_a L_{a,i} \right\rfloor. \quad (11)$$

The limited variation in the shape of the cranial cavity was exploited to speed up extraction by defining a Region of Interest (ROI) in which to label voxels. The

union of all N atlas labels a_i defined a region where brain voxels were possible, $a_1 \cup a_2 \cup \dots \cup a_N$, while the intersection defined a region where brain voxels were almost certain, $a_1 \cap a_2 \cap \dots \cap a_N$. The set difference, $(a_1 \cup a_2 \cup \dots \cup a_N) \setminus (a_1 \cap a_2 \cap \dots \cap a_N)$, yielded a ROI over which to apply patch-based brain extraction, which was dilated to account for possible variation not captured by the atlas set. The most similar atlases were chosen to label a new image based on their global similarity, measured by the SSD within the ROI.

As the accuracy of the atlas set is critical for accurate extraction, fetal atlases were constructed from manual delineations of the cerebrum, brain-stem, lateral ventricles and extra-cerebral CSF, filling the small remaining voids (cavum, third ventricle) automatically. A set of 35 atlases was generated, each consisting of an MR image and a binary mask with foreground voxels defining the matter inside the skull. Each atlas was then mirrored along the centre-line of the brain, effectively doubling the size of the atlas set to 70 in total.

Before extraction, the atlas set and input image were all co-aligned using affine transformations and intensity normalised. By aligning the images, and only considering patches in a small local neighbourhood, the computational cost is reduced considerably and the influence of dissimilar distant patches is removed. For this purpose, all atlases were affinely registered to a target subject of mean gestational age within the cohort and each new input image was then registered to this target before labelling. Patch similarity was evaluated using the SSD, therefore all the images were normalized using histogram matching (Nyúl and Udupa, 1999) to standardise the intensity scale, allowing SSD to be used as a meaningful similarity metric.

To determine the best parameters for brain extraction, a leave-one-out cross-validation was performed which resulted in a high mean Dice coefficient of 0.983

using 24 atlases, a patch diameter of 7 voxels and a neighbourhood diameter of 5 voxels.

3.2.3 *Automatic Segmentation*

For brain tissue segmentation purposes, the algorithm of [Ledig et al. \(2012\)](#) was adopted, which employs a second order [MRF](#) term to reduce misclassification of [PV](#) voxels on the cortical boundary, by penalizing the prior probability of [WM](#) when there is a high probability of both [GM](#) and [CSF](#) within its vicinity (see Section 2.2.4, adopted methods subsection).

[MRI](#) volumes were automatically segmented into a background class and seven foreground classes: cortical [GM](#), deep grey nuclei, germinal matrix, [WM](#), non-cerebral tissue (brain stem and cerebellum), intra-cerebral [CSF](#), and extra-cerebral [CSF](#). For this, a previously created non-rigid spatio-temporal atlas, comprising of a sequence of [MR](#) image templates at one week intervals ([Serag et al., 2012b](#)) and associated tissue probability maps ([Wright et al., 2012](#)) was used as prior information to guide the segmentation process. Each template was constructed using the target free approach of [Seghers et al. \(2004\)](#). All contributing images were non-rigidly registered in a pairwise manner using the [FFD](#) framework ([Rueckert et al., 1999](#)), and for each image in turn, the resulting transformations were averaged, giving a mapping towards a mean [MR](#) template. Templates and associated tissue probability maps were then created by averaging the transformed image intensities and segmentations.

Subject specific tissue probability maps were constructed on a leave-one-out basis to remove bias from manual segmentations, i.e. manual segmentations for each subject did not contribute to its own prior tissue probabilities. For further details

of the atlas construction, the reader is referred to (Serag et al., 2012b). Before segmentation, the atlas template with the closest age to the subject was non-rigidly registered to the input image. The associated tissue probability maps were subsequently transformed onto the native voxel grid of the subject before segmentation.

The FFD control point spacing used for pairwise registration during atlas construction controls the flexibility of the local deformation between subjects and has a considerable impact on the “sharpness” of the average MR intensity template and its associated tissue probability maps. As the control point spacing is reduced, the number of parameters increases and the registration becomes less robust and inherently more susceptible to misalignment errors, due to convergence to local minima. Furthermore, as atlas priors become sharper, they have stronger influence over a resulting segmentation. In this case, an EM algorithm will be dominated by the prior information and fail to compensate for registration misalignments, leading to a segmentation that does not appropriately describe the subject’s anatomy. Conversely, with a coarse control point spacing, the MR template and probability maps will be considerably less sharp and less informative. In this case, when an atlas prior is aligned with an input image, there can be significant WM probability overlapping cortical sulci and vice versa. This can lead to an increase in partial volume misclassification (Serag et al., 2012b).

Therefore, a relatively fine control point spacing of 1.5 mm was used for registration during atlas construction, which produces crisp tissue priors that can be relaxed (e.g. by blurring) during segmentation to compensate for misalignment of the original tissue priors (Cardoso et al., 2013). This relaxation step occurs after the EM algorithm has first converged to a solution before continuing. A wider FFD spacing of 3 mm was chosen for registering an age-matched template to each input image, giving a more robust registration.

After segmentation, a post-processing step was used to restrict the sub-cortical matter to a single mass, removing WM probability that is topologically implausible. Posterior probability maps for sub-cortical regions (germinal matrix, WM, deep grey nuclei) and intra-cerebral CSF were combined and then binarized using a threshold of 0.95. The Largest Connected Component (LCC) was assumed to be brain tissue, while all other residual components were assumed to be sulcal PV voxels. The binarisation threshold was chosen heuristically as the highest value which did not result in WM voxels inside thin gyri becoming disconnected from the LCC. All voxels with probability below the threshold were then classified as either brain tissue or sulcal PV voxels, based on proximity to the LCC and residual components. Finally, for all voxels determined to be sulcal PV voxels, WM posterior probability was redistributed to external CSF. This procedure was repeated iteratively using a descending threshold to remove all mislabelled WM.

3.2.4 *Voxel-based surface representation*

The surface of the brain is not an ideal location for measuring curvatures. As gyri-fication proceeds, sulci narrow, adjacent gyri begin to squeeze against each other, and it becomes increasingly difficult to delineate the cerebrum boundary using a voxel-based representation. Furthermore, inexact brain extraction may remove small amounts of cortical GM where the cerebrum is very close to the skull. Therefore curvatures were estimated for the inner cortical boundary which is easier to delineate and robustly segment.

A 1 mm voxel grid is too coarse to give an accurate surface representation using a binary volume and sub-voxel accuracy is required to capture the curvature of the convoluted cortex. Therefore, the input image was upsampled by a factor of

3 and a second segmentation step was performed using the previously calculated posteriors as priors. This resolution was found to be sufficient for accurate surface representation while retaining acceptable computation time and memory requirements. Further upsampling was also tested, but there was little visible difference between surfaces when visualised as a mesh.

For each subject, the inner cortical surface was approximated directly on the up-sampled voxel grid by combining the posterior probability maps for four tissues: germinal matrix, WM, the deep grey nuclei and intra-cerebral CSF, . This map was smoothed using a Gaussian kernel with a full width at half maximum of 2 mm which was found to ameliorate segmentation irregularities arising from artefacts in the reconstruction process, while retaining the fidelity of the estimated boundary. The map was then binarised using a threshold that preserved its volume and the largest connected component was retained. A surface voxel was then defined as any foreground voxel that shared a face with a background voxel.

3.2.5 Principle curvature estimation

For each surface voxel at location x_i , the principal curvatures k_1 and k_2 were calculated using the structure tensor and Hessian matrix using the method of (Rieger et al., 2004), which have previously been applied to neonatal images (Rodriguez-Carranza et al., 2008) (see Fig. 23 for an overview). The structure tensor is a matrix derived from the partial derivatives of a scalar function and can be used to analyse the gradient orientation of a neighbourhood. Numerical estimates of the gradient were obtained from a scalar map I obtained by smoothing the binary WM volume.

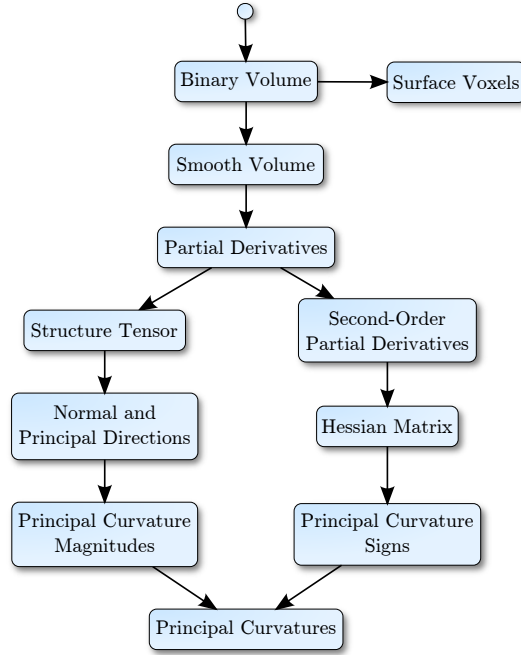


Figure 23: **Calculating the principal curvatures.** The surface voxels, where the principal curvatures are evaluated, are defined as any foreground voxel that shares a face with a background voxel within the binary WM volume. The binary volume was smoothed to produce a scalar field on which to estimate the image derivatives. The calculation of the principal curvature magnitudes and signs is then split into two processes. The structure tensor is computed from the image derivatives and allows the principal curvature directions and magnitudes to be computed. The signs of the principal curvatures are then recovered using the Hessian matrix.

The structure tensor $\mathbf{S}(x_i)$ for a surface voxel x_i is defined as the sum of outer products $\mathbf{S}^0(x_j)$ of the derivative vector $\mathbf{v}(x_j)$ with Gaussian weighting $w(x_i, x_j)$ for each voxel x_j in a local neighbourhood η_i :

$$\mathbf{S}(x_i) = \sum_{x_j \in \eta_i} \mathbf{S}^0(x_j) w(x_i, x_j).$$

$$\mathbf{S}^0(x_j) = \mathbf{v}(x_j) \mathbf{v}(x_j)^T.$$

$$\mathbf{v}(x_j) = \left[\frac{\partial I}{\partial x}(x_j) \quad \frac{\partial I}{\partial y}(x_j) \quad \frac{\partial I}{\partial z}(x_j) \right]^T.$$

Eigen-decomposition of the tensor summarises the gradient directions. Assuming an ordering of the eigenvalues: $\lambda_1 > \lambda_2 > \lambda_3$, then the gradient, the eigenvector \mathbf{e}_1 and \mathbf{n} the surface normal are all aligned. The eigenvectors \mathbf{e}_2 and \mathbf{e}_3 are aligned with the maximum and minimum curvature directions \mathbf{t}_1 and \mathbf{t}_2 respectively.

Unfortunately, the normal cannot be differentiated with respect to the principal directions to get the principal curvatures, as the vector field obtained from the first eigenvector is discontinuous i.e. $\mathbf{e}_1 = \pm \mathbf{n}$. Instead, a mapping $\mathbf{M}(\mathbf{e}_1) = \mathbf{e}_1 \mathbf{e}_1^T / \|\mathbf{e}_1\|$ is used to calculate the magnitudes and the sign of the curvature is recovered using the Hessian matrix. The gradient of $\mathbf{M}(\mathbf{e}_1)$ along the second and third eigenvectors is linearly related to the norm of the first and second principle curvatures respectively, by a constant $\sqrt{2}$:

$$|k_1| = \frac{1}{\sqrt{2}} \|\nabla_{\mathbf{e}_2} \mathbf{M}\| = \frac{1}{\sqrt{2}} \left\| \sum_{n=1}^3 \frac{\partial \mathbf{M}}{\partial \phi^n} e_2^n \right\|,$$

$$|k_2| = \frac{1}{\sqrt{2}} \|\nabla_{\mathbf{e}_3} \mathbf{M}\| = \frac{1}{\sqrt{2}} \left\| \sum_{n=1}^3 \frac{\partial \mathbf{M}}{\partial \phi^n} e_3^n \right\|,$$

$$\|\mathbf{M}\| := \text{Frobenius norm of } \mathbf{M}.$$

To recover the sign of a principal curvatures $k_{1,2}$ the Hessian matrix is employed \mathcal{H} (Equation 12) and the second derivative is examined along the corresponding principal direction $\mathbf{t}_{1,2}$, specifically $\mathbf{t}^T \mathcal{H} \mathbf{t}$ is evaluated. If the result is less than zero,

the principal curvature $k_{1,2}$ is positive (the curve is convex), otherwise the principal curvature is negative (the curve is concave).

$$\mathcal{H} = \begin{bmatrix} \frac{\partial I}{\partial x^2} & \frac{\partial I}{\partial xy} & \frac{\partial I}{\partial xz} \\ \frac{\partial I}{\partial xy} & \frac{\partial I}{\partial y^2} & \frac{\partial I}{\partial yz} \\ \frac{\partial I}{\partial xz} & \frac{\partial I}{\partial yz} & \frac{\partial I}{\partial z^2} \end{bmatrix}. \quad (12)$$

3.2.6 Folding measures

To quantify gyrification, eight curvature-based folding measures from the literature were used including global mean curvature (H_G) (Magnotta et al., 1999), global Gaussian curvature (K_G) (Magnotta et al., 1999), global curvedness (C_G) (Awate et al., 2008; Rodriguez-Carranza et al., 2008), mean curvature L_2 norm (H_N) (Batchelor et al., 2002), Gaussian curvature L_2 norm (K_N) (Batchelor et al., 2002), intrinsic Gaussian curvature (K_I) (Van Essen and Drury, 1997), mean curvature norm ratio (H_R) (Rodriguez-Carranza et al., 2008) and Gaussian curvature norm ratio (K_R) (Rodriguez-Carranza et al., 2008). Each of these measures was adapted to be independent of scale and surface area (Table 2). These measures were evaluated globally for each cortical surface and also for different regions. An atlas with nine cortical labels derived from (Gousias et al., 2012) was registered to each image volume, parcellating the cortical surface representations (Fig. 24) .

Rodriguez-Carranza et al. (2008) used a normalization factor $T = 3V/A$ to correct folding measures so that they are independent of scale and also surface area, which is required in order for values to be comparable over regions with different surface areas (Note $r = 3V/A$ for a sphere). An alternative approach is taken in

this work where the principal curvatures are first normalized for scale using a different normalization factor, so that the integration of a shape descriptor over a surface is normalized simply by dividing by its area.

The principal curvatures and the derived descriptors do not reflect the complexity of a surface as they are dependant on its scale. For example, the curvature of a sphere varies depending on its size. To remove the influence of scale on our measures, the principal curvatures were first normalized by multiplying by a correction factor r , which captures the scale of the brain. r was chosen as the radius of a sphere with equivalent volume to the WM, i.e. $r = \sqrt[3]{3V_{WM}/4\pi}$. This has the desirable property of producing a dimensionless quantity which is 1 for any point on the surface of a sphere. As a result, all of the derived folding measures in table 2 are also unitless and their values are also 1 for a sphere of any size. During experimentation, this method was also found to give better correlations between the folding measures and GA, possibly because it gives a better normalization of scale.

A voxel weighting term, w , was introduced to address the issue of unequal sampling density over the surface due to the voxel-based representation. For example, a plane perpendicular to an axis is more densely sampled than a diagonal plane. To compensate for this, the influence of each voxel was weighted depending on its connectivity configuration with background voxels and other surface voxels when computing a folding measure using the method described in (Windreich et al., 2003).

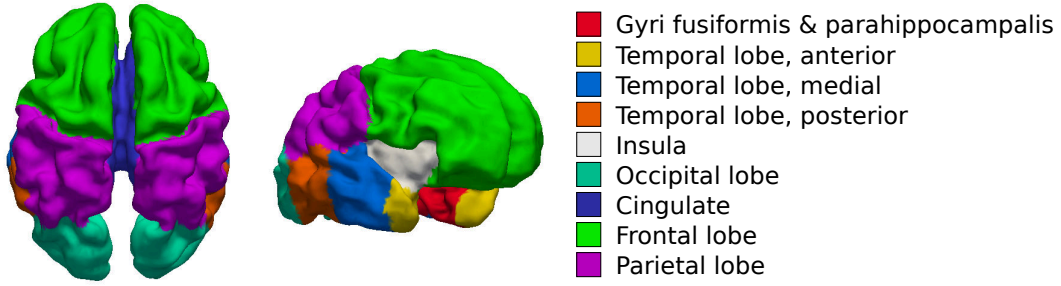


Figure 24: **Regional parcellation.** Example of a brain parcellated into nine different regions by registering an anatomical atlas.

| | |
|-------------------------------|---|
| Global mean curvature | $H_G = \sum_{x_i} Hw / \sum_{x_i} w$ |
| Global Gaussian curvature | $K_G = \sum_{x_i} Kw / \sum_{x_i} w$ |
| Global curvedness | $C_G = \sum_{x_i} Cw / \sum_{x_i} w$ |
| Mean curvature L_2 norm | $H_N = \sqrt{\sum_{x_i} H^2 w / \sum_{x_i} w}$ |
| Gaussian curvature L_2 norm | $K_N = \sqrt[4]{\sum_{x_i} K^2 w / \sum_{x_i} w}$ |
| Intrinsic Gaussian curvature | $K_I = \sqrt{\sum_{x_i \in K^+} Kw / \sum_{x_i \in K^+} w}$ |
| Mean curvature norm ratio | $H_R = \sum_{x_i} H^2 w / \sum_{x_i} H w$ |
| Gaussian curvature norm ratio | $K_R = \sqrt{\sum_{x_i} K^2 w / \sum_{x_i} K w}$ |

Table 2: **Folding measures.** The degree of gyrification was quantified using eight folding measures, each of which was formulated by summarising a shape descriptor over all surface voxels, x_i . A weighting, w , was applied to each voxel which depended on its surface voxel configuration. Each of these measures was formulated to ensure its independence of the scale and surface area. Note, K^+ , denotes the set of voxels with positive Gaussian curvature.

3.3 RESULTS & DISCUSSION

3.3.1 Segmentation

To assess the accuracy of the automated segmentation procedure, sixteen of the 80 subjects were manually segmented by a clinical expert and the discrepancy of the [GM/WM](#) boundary was measured. Details of the manual segmentation procedure can be found in ([Kyriakopoulou et al., 2014](#)). These subjects were chosen at evenly

spaced intervals throughout the cohort age range. A surface mesh was extracted from the manual segmentation using the marching cubes algorithm (Lorensen and Cline, 1987) and a boundary displacement map was extracted from the automated segmentation and interpolated to give a local displacement at each vertex on the mesh. A systematic bias was observed where the automated surface was displaced “inside” the manual surface by approximately 0.65 mm, on average, for the sixteen subjects. This can be explained by the tendency of the EM algorithm to overestimate the thickness of the cortex. GM/WM PV voxels are predominantly labelled as GM due to the greater variance in its intensity distribution (Van Leemput et al., 1999), thus leading to an overly thick cortical segmentation. The absolute deviation of the displacements around the median was fairly low at around 0.33 mm, on average. For comparison, intra and inter-rater error were also assessed and the median absolute displacement was found to be around 0.12 mm and 0.13 mm respectively. A visual comparison of manual and automated segmentations can be seen in Figure 25.

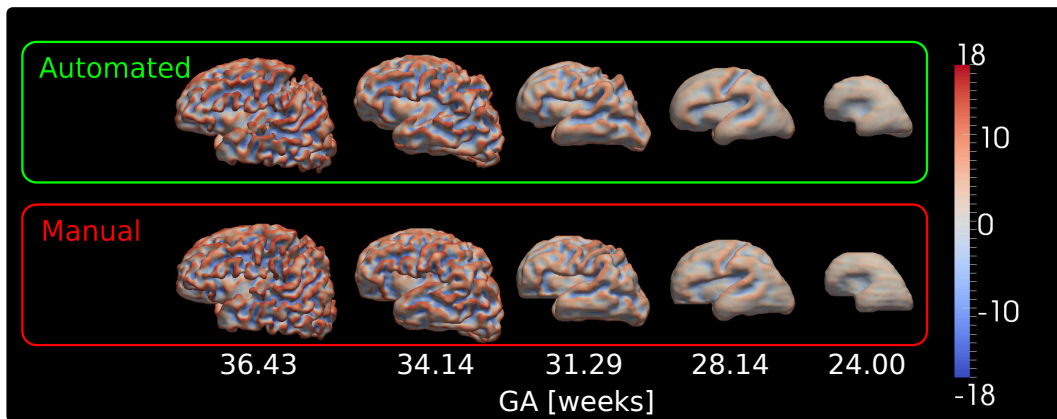


Figure 25: **Comparison of cortical surface meshes for manual and automatic segmentations.** Mesh representations of the surfaces were construction using the marching cubes algorithm for visualization purposes. Colour mapping depicts normalized mean curvature (H) at each point on the cortical surface.

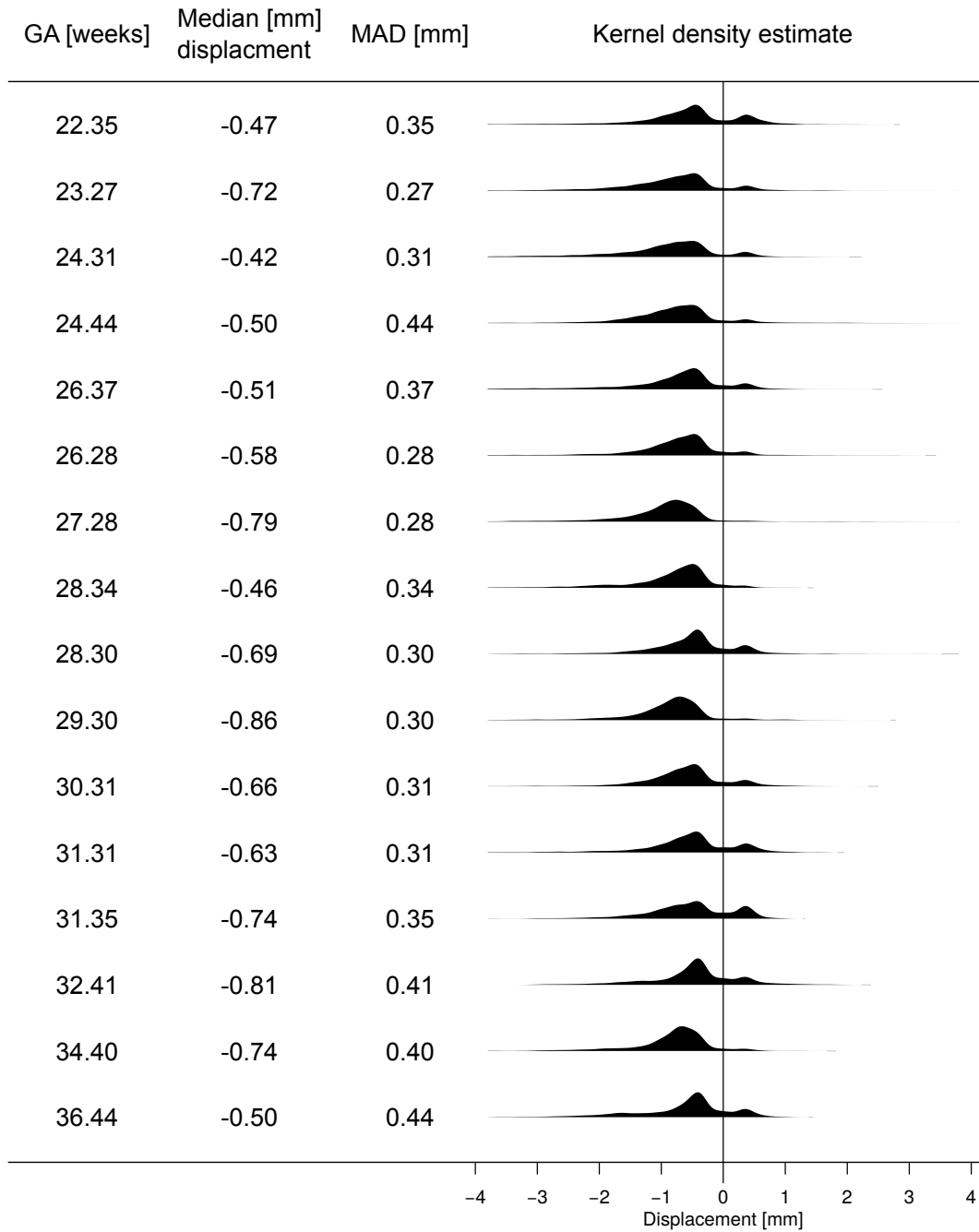


Table 3: **Segmentation accuracy.** Each row summarises the local displacements for an automated segmentation of an individual subject from its manual “ground truth”. The median displacement and the median absolute deviation (MAD) from the median are shown, as well as a kernel estimate of the distribution of displacements ($\sigma = 0.1$). Note, a negative displacement here means that the automated surface is “inside” the manual surface.

3.3.2 Folding Measures

Folding measures were computed over the automatically delineated cortical surfaces for all 80 subjects and also over the manually delineated cortical surfaces for sixteen of the subjects (Fig. 26). For six of the folding measures (C_G , H_N , K_N , K_I , H_R , K_R) a strong positive correlation with GA was observed. This suggests that these particular measures are effective for quantifying cortical folding and neurological development. Moreover, the correlation between these folding measures and GA is stronger than that of GA and volume (Table 4a), suggesting that neural development is more strongly linked to cortical complexity than brain size in this age range. These measures suggest that folding initially proceeds slowly, followed by a period of rapid change from 25 weeks onwards, which begins to slow towards birth. This reflects what has been observed visually (Garel et al., 2001), i.e. the brain is initially lissencephalic at 22 weeks with the sylvian fissure present as a shallow depression. Then primary sulci, such as the central sulcus appear first around 24-25 weeks before lobar gyrification intensifies around 30 weeks. By 34 weeks GA, all of the primary and most of the secondary sulci are present.

The computed values based on automatic segmentation were consistently lower than the computed values derived from manual segmentation and this discrepancy seemed to increase with GA. This can be explained partly by a systematic segmentation bias that led to an underestimation of the WM volume, which in turn would have decreased the magnitude of the normalized principal curvatures and the computed folding measures. The increasing discrepancy may be explained by an increase in segmentation error, for example, a missing fold may decrease the complexity of the estimated surface. The levelling off of computed values at later gestational ages is likely to be overestimated, therefore, and folding measures may in fact converge to a much higher value.

The asymptotic growth in cortical complexity suggested by these results, where the rate is slowest at the start and end of a time period, is accurately modelled by a Gompertz function, which was first proposed by Gompertz in 1825. This function has been successfully used to model various phenomena that reach a saturation point, such as mobile phone users (Tao, 2010) and tumour cells (Laird, 1964). In the instance of tumour growth, a cellular population initially increases at an exponential rate but growth slows as nutritional supplies deplete. Gyrfication is also a growth process that must eventually reach a saturation point. In this case, folding is limited by the thickness of the cortical sheet, which cannot be folded indefinitely due to physiological limitations. A Gompertz function is, therefore, a good candidate for modelling the rate of gyrfication through gestation.

3.3.3 *Gompertz model*

In this section the suitability of a Gompertz-like function for modelling the relationship between the folding measures and gestational age is assessed. The chosen function (Eqn. 13), predicts the value of a folding measure y given a gestational age t and a set of non-negative parameters β_i . In this model, parameter β_1 is the initial value of y when the brain is lissencephalic before folding commences, $\beta_1 + \beta_2$ is the upper limit of y when folding has ceased increasing, β_3 controls the growth rate that determines the increase of y during gyrfication and β_4 is a centring parameter, which determines the gestational age where y increases fastest. The fitting of the model and estimation of the confidence intervals was carried out using the Statistics Toolbox functions `nlinfit` and `nlpredci` in the MatLab Software package (MATLAB and Toolbox, 2011) .

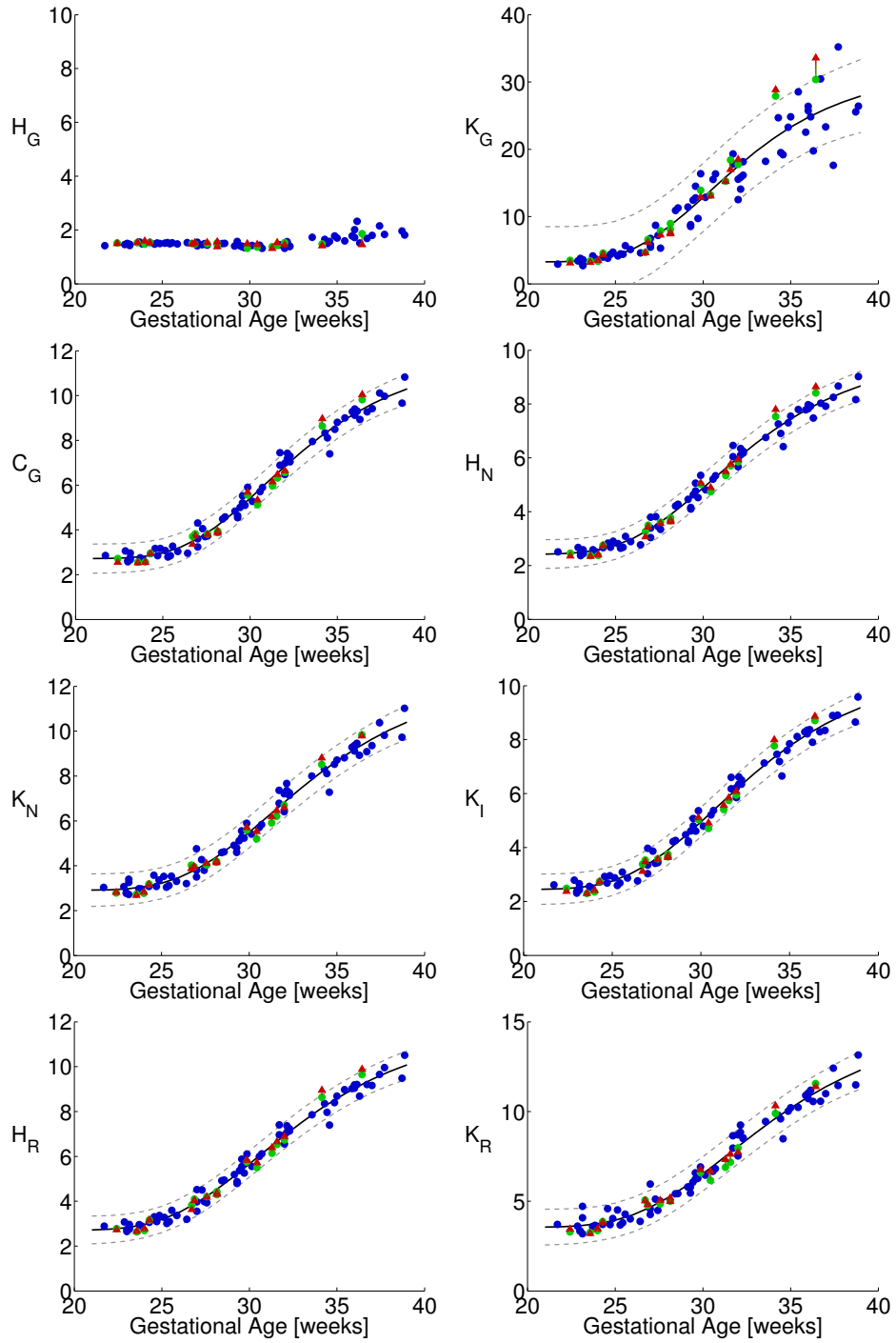


Figure 26: **Global folding measures with gestational age.** Folding measures computed from automated segmentations are shown as either green or blue circles, whereas curvatures computed from manual segmentations are shown as red triangles. For each subject that was segmented both manually and automatically, the corresponding points are shown in red and green respectively and joined together by a line. The remaining subjects that were only segmented automatically are represented by a blue circle. A Gompertz function (black line) was fitted to each plot except global mean curvature, which did not exhibit a Gompertz-like relationship. Upper and lower confidence intervals are shown as dashed grey lines.

$$\hat{t} = t - \beta_4.$$

$$f(\hat{t}) = \beta_1 + \beta_2 e^{-e^{-\beta_3 \hat{t}}}. \quad (13)$$

$$f'(\hat{t}) = \beta_2 \beta_3 e^{-\beta_3 \hat{t}} - e^{-\beta_3 \hat{t}}. \quad (14)$$

To determine whether a Gompertz function accurately models the relationship between folding measures and gestational age, the goodness of fit for unseen data was evaluated by performing 100 iterations of 10 fold cross validation. The data were partitioned into 10 equal folds and each fold was held back in turn while a model was fitted to the remaining data and a residual was predicted for each omitted observation. The sum of square prediction residuals, $\hat{\epsilon}$, gave an evaluation measure, which was averaged over 100 iterations. The Gompertz model was compared against two other models: a linear function and a quadratic function (Table 4), to determine whether a simpler model could fit the data as well. Globally, Gompertz functions were found to fit the data closely ($R^2 = 0.99$) and considerably better than a simple quadratic or linear fit (Table 4a). On average, adopting a Gompertz model reduced $\hat{\epsilon}$ by a third over a quadratic model and by almost a half over a linear model.

It must be noted that while gyrification must cease increasing at some point, and although the relationship between folding measures and age is expected to be well modelled by a Gompertz function, systematic segmentation errors towards the latter gestational ages, where segmentation is particularly challenging, may lead to a false estimate of growth. If the segmentation algorithm fails to segment an increasingly complex cortical surface, the curvature measures may be underestimated and an asymptotic value may be reached before gyrification has actually ceased, and the value of model parameter β_2 will therefore be underestimated.

| Measure | Gompertz | | Quadratic | | Linear | |
|---------|------------------|--------------|------------------|-------|------------------|-------|
| | $\hat{\epsilon}$ | R^2 | $\hat{\epsilon}$ | R^2 | $\hat{\epsilon}$ | R^2 |
| C_G | 8.4 | 0.991 | 16.3 | 0.983 | 22.7 | 0.977 |
| H_N | 5.5 | 0.992 | 10.2 | 0.984 | 11.9 | 0.982 |
| K_N | 10.3 | 0.989 | 15.4 | 0.983 | 22.6 | 0.975 |
| K_I | 6.1 | 0.992 | 11.3 | 0.985 | 15.0 | 0.980 |
| H_R | 7.2 | 0.992 | 12.3 | 0.986 | 14.3 | 0.984 |
| K_R | 19.5 | 0.984 | 24.8 | 0.980 | 35.3 | 0.971 |
| Volume | - | - | - | - | - | 0.958 |

(a) Model fit with different folding measures and volume

| Region | Gompertz | | Quadratic | | Linear | |
|-------------------------------------|------------------|--------------|------------------|-------|------------------|--------------|
| | $\hat{\epsilon}$ | R^2 | $\hat{\epsilon}$ | R^2 | $\hat{\epsilon}$ | R^2 |
| Global | 5.5 | 0.992 | 10.2 | 0.984 | 11.9 | 0.982 |
| Gyri fusiformis & parahippocampalis | 9.9 | 0.986 | 15.8 | 0.978 | 20.7 | 0.971 |
| Temporal lobe, anterior | 23.5 | 0.950 | 22.2 | 0.953 | 22.1 | 0.953 |
| Temporal lobe, medial | 6.3 | 0.989 | 9.3 | 0.984 | 13.8 | 0.976 |
| Insula | 8.8 | 0.969 | 8.8 | 0.969 | 9.3 | 0.967 |
| Occipital lobe | 6.3 | 0.990 | 11.1 | 0.983 | 16.2 | 0.975 |
| Temporal lobe, posterior | 10.5 | 0.988 | 22.8 | 0.974 | 25.4 | 0.971 |
| Cingulate | 19.0 | 0.959 | 20.6 | 0.955 | 26.8 | 0.941 |
| Frontal lobe | 5.8 | 0.991 | 10.1 | 0.984 | 11.5 | 0.982 |
| Parietal lobe | 7.3 | 0.991 | 15.6 | 0.981 | 15.3 | 0.981 |

(b) Model fit in individual regions for H_N

Table 4: **Modelling the relationship of folding measures and GA.** Three models were fitted to the data: a Gompertz model, a quadratic model and a linear model. The predicted sum of square residuals $\hat{\epsilon}$ and coefficient of determination R^2 were estimated to evaluate the fit of each model.

Given the precise timing of sulcal development (Garel et al., 2001), a high correlation between age and folding measures is expected. However, the possibility of a strong bias towards the atlas priors further reducing variability and falsely strengthening the observed correlation, is ruled out. Each atlas frame is constructed by averaging FFDs for contributing subjects towards other subjects within a time window. The anatomy and segmentations are then transformed to an av-

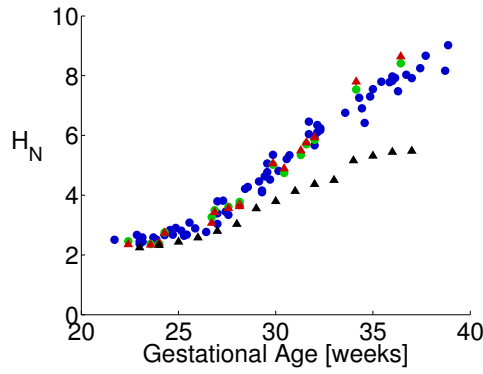


Figure 27: **Folding measures computed from atlas priors.** The mean curvature L_2 norm (H_N) is shown for each subject computed from automated segmentations (either green or blue circles), manual segmentations (red triangles) and atlas priors (black triangles). For subjects segmented both manually and automatically, the corresponding points are shown in red and green respectively and joined together by a line. Note the values of H_N computed from atlas priors are much lower than for subjects of a similar age.

erage atlas space using this deformation. The resulting atlas template does not exhibit the typical anatomy of each subject but rather an average, with reduced variability of gyral location and an overall smoother cortical surface. Therefore, curvature measures computed from atlas priors are expected to be lower than from an aged-matched subject. If the atlas priors were to dominate the segmentation and the algorithm failed to capture the subject's anatomy, then this would be reflected in the curvature measures. Fig. 27 shows that this is not the case, as the curvatures computed from atlas priors are considerably lower than those computed for individual subjects with a similar age.

3.3.4 Regional Folding

Examining folding measures locally revealed regional differences in sulcation. Lobar regions exhibited a similar relationship between GA and computed folding measures to what was observed globally, and were also well modelled by a Gom-

pertz function (Fig. 28a). The insula and the anterior temporal lobe, however, did not display the same behaviour. A steady increase in folding measures with GA was observed in the anterior temporal lobe over the time period, whereas the insula remained relatively smooth, with folding measures not increasing with GA at the same rate as other regions. Consequently, a linear model gave a better fit for these regions. (Table 4b).

The rate of growth at different gestational ages was estimated by solving the derivative of the fitted Gompertz functions (Eqn. 14) for the folding measure H_N . The models suggest that peak growth occurs around 30 weeks GA for all lobar regions, which suggests that this is a critical time period and any developmental disturbances could have a greater impact around this point (Fig. 28b). Regional differences in growth rates were also observed, with the posterior temporal lobe and parietal lobe showing the highest peak growth, with other regions such as the frontal lobe and medial temporal lobe developing more slowly, and the cingulate considerably slower (Table 5).

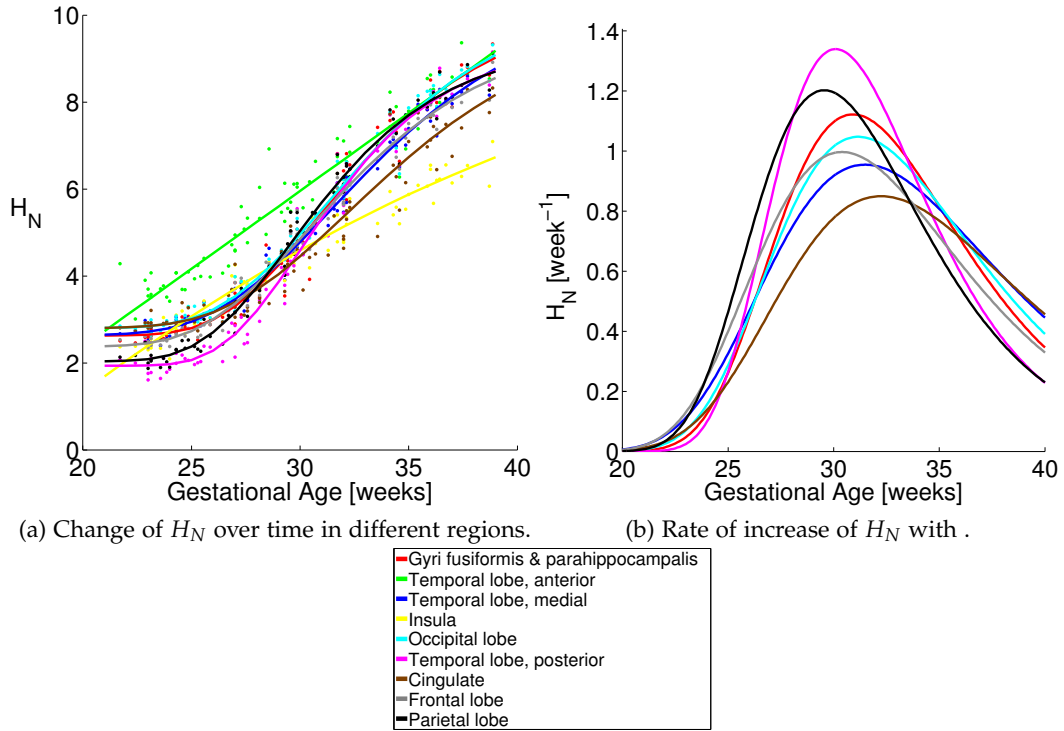


Figure 28: **Regional differences in folding measures.** (a) The change in mean curvature L_2 norm (H_N) over individual regions with gestational age is shown. All lobar regions except the insula and anterior temporal lobe exhibit a Gompertz like growth pattern. (b) The rate of growth across gestational ages was computed from the derivative of the fitted Gompertz functions for H_N in lobar regions. The rate of growth peaked around 30 weeks gestational age for all regions, however there are differences in growth rates.

| Region | Peak Growth | GA |
|-------------------------------------|-------------|------|
| Temporal lobe, posterior | 1.00 | 29.8 |
| Parietal lobe | 0.90 | 29.6 |
| Gyri fusiformis & parahippocampalis | 0.84 | 30.9 |
| Occipital lobe | 0.78 | 31.2 |
| Frontal lobe | 0.74 | 30.4 |
| Temporal lobe, medial | 0.71 | 31.5 |
| Cingulate | 0.63 | 32.2 |

Table 5: **Relative peak growth rate in lobar regions.** The peak increase of H_N is shown for all regions relative to the posterior temporal lobe. Peak growth occurred around 30 weeks gestational age for all lobar regions and regional differences in growth rates are apparent.

3.3.5 *Physiological Age*

The strong relationship between *GA* and folding measures allowed an accurate prediction of *GA* to be made. The inverse of the Gompertz function fitted for measure H_N was taken to predict *GA* from the value of H_N observed for each fetus. Prediction accuracy was inherently lower where the rate of folding was slowest. When gyrification changes very little, folding measures become a poor indicator of *GA*, as fetuses in close proximity have a similar degree of gyrification. Therefore poor age prediction is expected towards the extremities of the age range in the cohort (Fig. 29). The mean prediction error was 0.47 weeks (± 0.40) in the age range 24 – 37 weeks, and 0.88 weeks (± 0.61) outside of this range, where the rate of folding was considerably slower. Given that the real *GA* is unknown and is estimated from either an obstetric ultrasound scan or the last menstrual period, the measured prediction accuracy of the classifier is limited and may not reflect the real accuracy.

Predicting the real *GA* of a fetus from an *MR* image is not a genuinely useful application of this work, given that *GA* may be estimated by obstetric ultrasound. However, by estimating the age of a fetus based on the observed anatomy, it is actually the physiological age which is being estimated i.e. the average age of a normal fetus with the equivalent level of neurological development. A difference in real *GA* and estimated *GA* could, therefore, indicate a delay or acceleration from normal development, associated with some pathology. For example, in (Dittrich et al., 2014) age was predicted for a group of fetuses with lissencephaly, a condition which results in a smoother cortical surface from defective neuronal migration, and a mean prediction offset of -3.0 ± 3.5 days was observed.

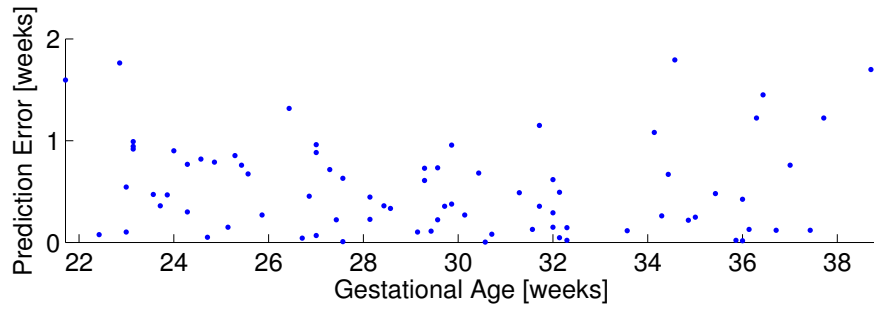


Figure 29: **GA prediction error.** Gestational age was predicted from the observed value of H_N by taking the inverse of the Gompertz function fitted to the data for H_N .

3.4 CONCLUSIONS

In this chapter, a complete framework to quantify gyrification was presented and applied to a cohort of 80 normal fetuses (age range 22 – 39 weeks). The curvature-based folding measures that were employed were strongly correlated with GA. Moreover, these measures were better correlated with GA than volume, suggesting that neurological development is more closely linked with the convolution of the cortex than the size of the cerebrum.

The increase in folding measures with GA reflected the rate of gyrification that has been observed visually (Garel et al., 2001), with gyrification accelerating rapidly between 25 and 30 weeks. A Gompertz function was found to fit the observed data well, and better than a simple linear or quadratic model. However, curvature measures were systematically underestimated at older gestational ages when using automated methods, and therefore, it is likely that curvature estimates will reach an asymptotic level while gyrification is continuing. Thus, further work is needed to assess the reliability of predictions based on this model at times when the gyrification process is nearing completion.

The Gompertz model allowed an accurate prediction of GA of a fetus with a mean error of approximately half a week between 24 – 37 weeks GA. The neurological development of a fetus can then be characterised using the notion of physiological age and a developmental delay or acceleration can be quantified in weeks. The model also allowed us to analyse growth rates over time, and peak gyrification was observed in all lobar regions around 30 weeks. Therefore, this period may be a critical time for neurological development.

This framework is well-suited for detecting gross discrepancies in gyrification associated with various neurological disorders globally, and also locally, over predefined regions. Disturbances in development may be localised and abnormal folding may remain undetected when averaging over a relatively large region. Cortical anatomy may be analysed on a fine scale when aligned to a surface template using surface-based morphometry in a procedure analogous to VBM for images. In the next chapter, we will look at the difficult task of constructing a cortical surface-atlas, which will facilitate this goal.

CONSTRUCTION OF A SPATIO-TEMPORAL SURFACE ATLAS

4.1 INTRODUCTION

The methodology presented in the previous chapter focused on quantifying cortical development at a global and lobar scale. Folding measures were computed that summarised the average extent of folding over a given region with a single scalar value. Folding anomalies that are highly localised may remain undetected using this methodology. It is desirable to be able to compare cortical anatomy on a finer scale. An average model or atlas can facilitate this by providing a densely sampled reference space where the anatomies of a population can be co-aligned and morphometric analysis performed. [MRI](#) based atlases have been used previously to align and compare neuroanatomy between cohorts on a voxel-wise basis using [VBM](#) ([Ashburner and Friston, 2000](#)). This methodology can also be applied to surfaces, undertaking morphometric analysis on a vertex-wise basis using mesh models ([Chung et al., 2008](#)).

Constructing a surface atlas requires finding smooth and accurate correspondences between cortical surfaces from different subjects at different ages, which can be a challenging task. Surfaces may be registered in 3D using transformation models commonly used for image-based registration, such as an affine transformation matrix or an [FFD](#). For each optimisation iteration, dissimilarity may be estimated as the sum of square distances, for each target mesh vertex from its nearest

source mesh vertex, once transformed. For highly convoluted surfaces, however, this procedure often converges to an incorrect solution.

Embedding cortical surfaces into some 2D domain before registration is a more robust strategy for finding correspondences. A cortical surface may be flattened by making a series of cuts, allowing it to be embedded onto a plane with minimal distortion (Van Essen and Drury, 1997). Flattened surfaces may be registered by adapting methods developed for image registration, using scalar information associated with the mesh vertices (e.g. mean curvature) as a substitute for image intensities.

Alternatively, a cortical surface may be embedded in a 2D spherical coordinate system. This avoids cutting the cortical surface, thus preserving its topology. (Fischl et al., 1999b) proposed a method for “inflating” a cortical surface by minimizing an energy functional consisting of two terms: a distance term (to reduce geometric distortions) and an oriented areal term (to unwrap folds). Note as distances are preserved, this procedure is more analogous to inflating a crumpled up paper bag than a balloon. Manifold learning techniques, such as spherical multi-dimensional scaling (Elad et al., 2005) or spherical Isomap (Tenenbaum et al., 2000), may also be used to embed a cortical surface, constraining its vertices to lie on a sphere. A number of authors have developed registration methodologies specific to a spherical coordinate systems, including Freesurfer (Fischl et al., 1999a), spherical demons (Yeo et al., 2010) and multi-modal surface matching (Robinson et al., 2014).

Spectral graph theory gives the basis for a fast alternative to the traditional registration methods discussed so far, using the eigenvalues and eigenvectors of the graph Laplacian matrix (Chung, 1997). The eigenvalues of the Laplacian may be interpreted as the natural oscillating frequencies, or resonances, of a physical shape. Each of these resonances has a distinct pattern of displacement, a so-called nor-

mal mode of the shape, which is given by the eigenvectors. At any given resonant frequency, equivalent locations on similarly shaped surfaces will have a similar relative amplitude and phase, which allows surface points to be matched in the spectral domain. Accuracy with spectral methods can be limited due to macroscopic variations in brain shape, such as expansion or compression, that can cause significant changes in a cortical surface's spectral representation. Previous attempts to address this problem include registering the derived spectral representations (Jain and Zhang, 2006; Mateus et al., 2008; Lombaert et al., 2011; 2013a).

More recently, Lombaert et al. (2013b) proposed a novel spectral-based method to acquire a mapping between surfaces that is comparable in accuracy to Freesurfer (Fischl et al., 1999a) and spherical demons (Yeo et al., 2010), while maintaining a considerable speed advantage. Moreover, atlases generated are more consistent when varying the initial reference surface for aligning a group of surfaces. The standard deviation of atlas boundaries was found to be significantly lower for the spectral method ($0.0014\text{ mm} \pm 0.0009$) compared to spherical demons ($0.28\text{ mm} \pm 0.08$). In this method, a preliminary point-wise correspondence between two surface graphs is established from their spectral representations and from additional information, such as sulcal depth. A novel dual-layered graph is then constructed by linking the corresponding vertices across the two graphs. The spectral decomposition of this dual-layered graph provides an orthonormal basis where accurate correspondences between the two cortical surfaces can be estimated. Orasanu et al. (2014) used this methodology to analyse cortical development in neonates born very preterm (around 26 weeks). Cortical surfaces for five subjects were modelled at two time-points, 30 weeks and 40 weeks equivalent gestational ages. For each individual, cortical correspondences were then estimated between the two time-points, enabling longitudinal changes in cortical geometry to be measured at matched locations.

Given the considerable diversity of cortical shape during gestation, a single anatomical atlas may not be appropriate. Indeed, a sulcus on the surface of an older, more neurologically developed brain may not yet exist on a younger brain and thus it may not be possible to establish a meaningful correspondence. Without any detectable surface sulcation, from some shape descriptor (e.g. mean curvature), the spectral properties alone can be used to determine the correspondences. Unfortunately, variations in cerebral shape can alter the spectral vibration modes, limiting matching accuracy. Furthermore, macroscopic changes in cerebral shape during gestation alter vibration modes over time, therefore accuracy will also fall as the discrepancy in age increases. A spatio-temporal approach (Serag et al., 2012a) can be used to avoid these difficulties, as an average anatomical template is constructed at several time points, each within a specific time interval where anatomy is similar or not expected to change significantly.

To date, there have only been a handful of studies which have constructed atlases of the developing cortex. Hill et al. (2010) constructed a single cortical surface template from neonates born at term to compare hemispheric asymmetries in infants and adults. Van Essen (2005) used a landmark based registration algorithm to align spherical representations of the cortex. More recently, Li et al. (2014) constructed an infant spatio-temporal surface atlas, generating atlas templates for 1, 3, 6, 9, 12, 18 and 24 months of age. This work also aligned cortices on the sphere using spherical demons (Yeo et al., 2010).

A significant portion of the cortical folding process is completed before term, including the formation of primary and secondary sulci (Garel et al., 2001), and disturbances in development during this period may have a profound effect. Preterm neonatal imaging, therefore, provides an opportunity to visualise an important part of the cortical folding process. However, the complications associated with premature birth mean that this cohort is not ideally suited for characterising nor-

mal development (Ajayi-Obe et al., 2000; Kapellou et al., 2006). For this reason, *in utero* MR imaging may ultimately provide improved visualisation and quantification of the developing cortex over this critical period.

Two studies have constructed average surface templates of the fetal cortex from *in utero* MRI. Habas et al. (2012) described an image-analysis framework to investigate early folding patterns, revealing asymmetric hemispheric development in a cohort of fetuses from 20 to 28 weeks GA. MR image volumes were co-aligned at each week of gestation using image-based registration and an average cortical surface was extracted from the combined subject tissue maps. 3D image-based registration is, however, ill-suited for cortical alignment at older gestational ages, where the geometry of the cortex is significantly more complex.

More recently, Clouchoux et al. (2012) constructed a spatio-temporal surface atlas consisting of four templates based on twelve *in utero* MR images of healthy fetuses (25.2 to 35.1 weeks gestational age), which enabled a quantitative analysis of cortical development. Surfaces were reconstructed using a deformable model, CLASP (Kim et al., 2005), and correspondences were found on the unit sphere using an iterative group registration scheme (Lyttelton et al., 2007). This methodology was also used to compare development in normal fetuses and fetuses with a severe form of congenital heart disease (Clouchoux et al., 2013), where a delay in cortical development was observed.

4.1.1 Chapter overview

Most of the atlas construction techniques discussed have utilised a spherical coordinate system to register cortical surfaces. In this chapter, an alternative spectral method is presented for constructing a spatio-temporal surface atlas from a cohort

of 80 healthy fetal subjects (Section 1.4). Average surface templates are constructed for each week of gestation from 23 to 37 weeks (15 templates in total). A natural extension to the spectral-based method of Lombaert et al. (2013b) is used to embed a group of cortical surfaces simultaneously, thus eliminating bias due to the choice of an initial reference surface for alignment. Kernels (Nadaraya, 1964) are then used to compute a spatially and temporally weighted estimate of the average surface boundary, which can be readily converted to a mesh representation (Kazhdan and Hoppe, 2013).

The methodology developed is evaluated by examining sulcal alignment, regional overlaps, and embedding regularity in comparison with spherical demons and conventional 3D surface registration algorithms. Additionally, the consistency of the average templates is demonstrated by generating multiple instances from disjoint subsamples of the dataset of varying sizes.

4.2 MATERIAL AND METHODS

4.2.1 *Surface Modelling*

For each MR image, several anatomical structures were first segmented using an EM algorithm (Makropoulos et al., 2014), including cortical GM, WM, CSF, the deep grey nuclei, lateral ventricles, brain stem and cerebellum. A probability map was then formed for matter inside the GM/WM boundary by summing the posterior probability for WM, the deep grey nuclei and the lateral ventricles. A super-sampled binary volume B was then extracted (threshold = 0.5) and from this a distance map D was computed (Maurer et al., 2003); this implicitly defines the surface boundary

in the image space. A negative distance indicated a voxel was inside the surface boundary, while a positive distance indicated a voxel was outside.

An explicit surface representation was extracted by deforming a mesh with spherical topology towards the implicit isosurface $D = 0$ using a similar approach to previous authors (Dale and Sereno, 1993; Davatzikos and Bryan, 1996; MacDonald et al., 2000; Kim et al., 2005). The deformation maintains the topology of the mesh and fits the mesh directly to the probability map rather than the image intensities.

Let surface model S be defined as pair $S = \{\mathbf{V}, \mathbf{E}\}$, where \mathbf{V} denotes a set of vertices $\mathbf{V} = (\mathbf{x}_1, \dots, \mathbf{x}_{|\mathbf{V}|})$ with position vectors \mathbf{x}_i , and \mathbf{E} denotes an adjacency matrix, where $e_{ij} = 1$ if the vertices i and j form an edge, or 0 otherwise. The cortex consists of a thin continuous sheet of GM, enclosing the WM and sub-cortical GM structures beneath. Its surface is appropriately modelled as a closed 2D manifold that is homeomorphic to a sphere, therefore S can be initialised by any convex, uniformly sampled, closed triangulated mesh with zero genus, which encloses a binary volume B .

The cortical surface is then estimated by iteratively deforming S towards the boundary defined implicitly by D (Algorithm 1). Each vertex \mathbf{x}_i is updated in turn and its new position is given by stepping along the inward surface normal direction \mathbf{n}_i , with step size γ :

$$\gamma = \begin{cases} \gamma_{max} & D(\mathbf{x}_i) > \gamma_{max} \\ -\gamma_{max} & D(\mathbf{x}_i) < -\gamma_{max} \\ D(\mathbf{x}_i) & \text{else} \end{cases}.$$

\mathbf{x}_i is only updated if a ray cast along $\gamma\mathbf{n}_i$ intersects with B , thus ensuring that vertices only move towards the boundary. Additionally, if a ray cast along $\gamma\mathbf{n}_i$ first intersects with S at \mathbf{y}_i and $\gamma + \|\mathbf{x}_i - \mathbf{y}_i\| > d$, where d is the closest permissible

distance between S and itself, then \mathbf{x}_i is not updated, to ensure the mesh does not self intersect.

After all vertices have been updated, their positions are smoothed, taking an average of the neighbouring vertex positions, ensuring each vertex moves consistently with its neighbours (Algorithm 2). To allow the surface to dynamically expand and compress without affecting the vertex density or topology of the mesh, S is remeshed after each iteration, splitting long edges and collapsing short edges. Both the update and smoothing steps are then repeated until S has converged to the implicit surface $D = 0$.

Algorithm 1: Deform Mesh.

Data: $S = \{V, E\}$, D , B

for 1 **to** n **do**

for $\mathbf{x}_i \in V$ **do**

$$\gamma = \begin{cases} \gamma_{max} & D(\mathbf{x}_i) > \gamma_{max} \\ -\gamma_{max} & D(\mathbf{x}_i) < -\gamma_{max} \\ D(\mathbf{x}_i) & \text{else} \end{cases}$$

if $\text{Intersect}(B, \mathbf{x}_i, \gamma \mathbf{n}_i)$ **and not** $\text{SelfIntersect}(S, \mathbf{x}_i, \gamma \mathbf{n}_i)$ **then**

$\mathbf{x}_i = \mathbf{x}_i + \gamma \mathbf{n}_i$

$S = \text{Smooth}(S)$;

$S = \text{Remesh}(S)$;

Algorithm 2: Smooth.

Data: $S = \{V, E\}$

for $i = 1$ **to** $|V|$ **do**

$N = \{j \mid e_{ij} = 1\}$

$$\bar{\mathbf{x}}_i = \sum_{j \in N} \mathbf{x}_j w_j \quad (w_j \propto 1 / \|\mathbf{x}_i - \mathbf{x}_j\|^2 \quad \text{and} \quad \sum_{j \in N} w_j = 1)$$

for $i = 1$ **to** $|V|$ **do**

$\mathbf{x}_i = \bar{\mathbf{x}}_i$

All meshes that were extracted were tested and confirmed to be topologically correct with no self-intersections. To estimate the surface extraction error, a distance map was computed for a segmented boundary (Maurer et al., 2003), and distance

values were linearly interpolated at the mesh vertex locations. The average distance error for a vertex was 0.106 mm (± 0.006) (see Fig. 30 for surface contour examples).

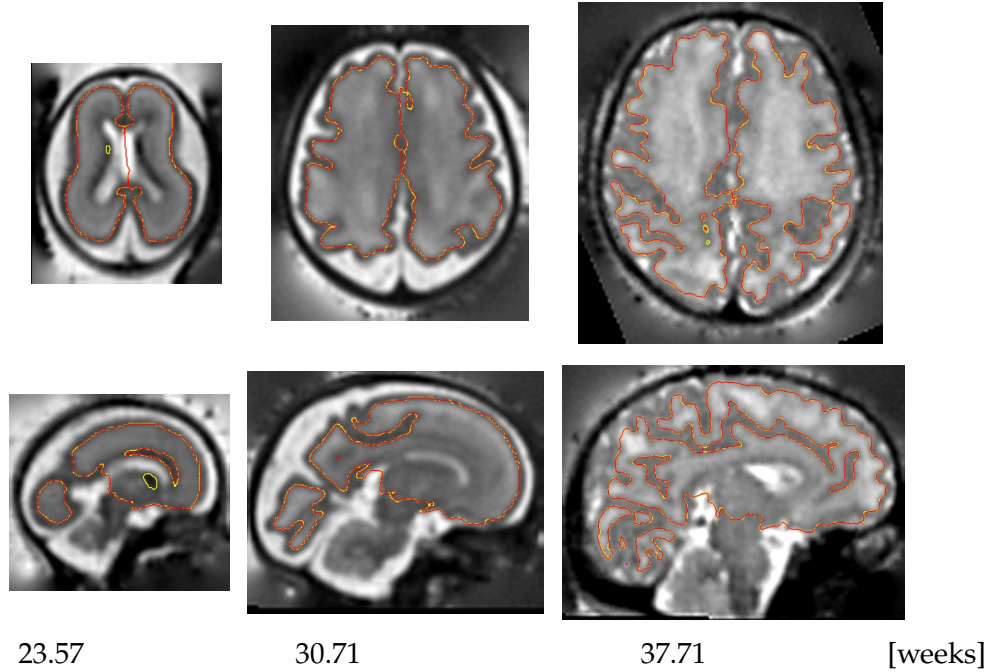


Figure 30: **Cortical Surface Extraction.** Cortical surface models for a selection of gestational ages are overlaid on their corresponding MR image volumes. The isosurface of the sub-cortical tissue probabilities is shown as a yellow contour while the cross-section of the cortical surface model is overlaid as a red contour.

4.2.2 Atlas construction overview

Given the temporal diversity of anatomy within the cohort, a spatio-temporal framework was adopted, building an average template of cortical shape for each week of gestation across the cohort (Fig. 31). In particular, the approach avoids the difficult task of aligning very disparate cortical anatomies and trying to find accurate correspondences between them. Each template was built from a group of cortical hemisphere surfaces ($N = 19.6 \pm 3.1$) within a narrow time-window

of ± 1 week, where cortical features are similar. Both hemisphere surfaces were used to generate the atlas templates (mirroring the right hemisphere), effectively doubling the sample size, thus allowing a more reliable estimate of the population average surface boundary. By choosing overlapping windows, correspondences between successive embeddings can still be established (Section 4.2.7). For each template, the contributing surfaces were first affine-aligned by registering all pairs of surfaces, and deforming each surface using the computed average affine transformation (Seghers et al., 2004). This improved matching accuracy by removing differences in global shape and size. Spectral matching was then performed between all pairs of surfaces, yielding initial vertex-wise correspondences. Surfaces were connected using these preliminary correspondences and a joint spectral analysis gave an embedding in which all the surfaces were co-aligned. The average surface position was then estimated using kernel regression at each point within the embedding, allowing a mean surface to be reconstructed. The following describes these steps in further detail.

4.2.3 Spectral Matching

Spectral matching is a basic concept that has been used by several authors (Jain and Zhang, 2006; Mateus et al., 2008; Lombaert et al., 2011; 2013a;b; Orasanu et al., 2014) to find correspondences between two surfaces from their spectra. To recap, for a surface model $S = \{\mathbf{V}, \mathbf{E}\}$, its $|\mathbf{V}| \times |\mathbf{V}|$ weighted adjacency matrix \mathbf{W} is defined, $w_{ij} = e_{ij} \|\mathbf{x}_i - \mathbf{x}_j\|^{-2}$. The diagonal degree matrix $\mathbf{D} = \text{diag}(d_1, d_2, \dots, d_{|\mathbf{V}|})$ is then given by summing all edge weights for each node, $d_i = \sum_j w_{ij}$. The $|\mathbf{V}| \times |\mathbf{V}|$ general graph Laplacian $\mathbf{L} = \mathbf{D} - \mathbf{W}$ can now be constructed from \mathbf{W} and \mathbf{D} .

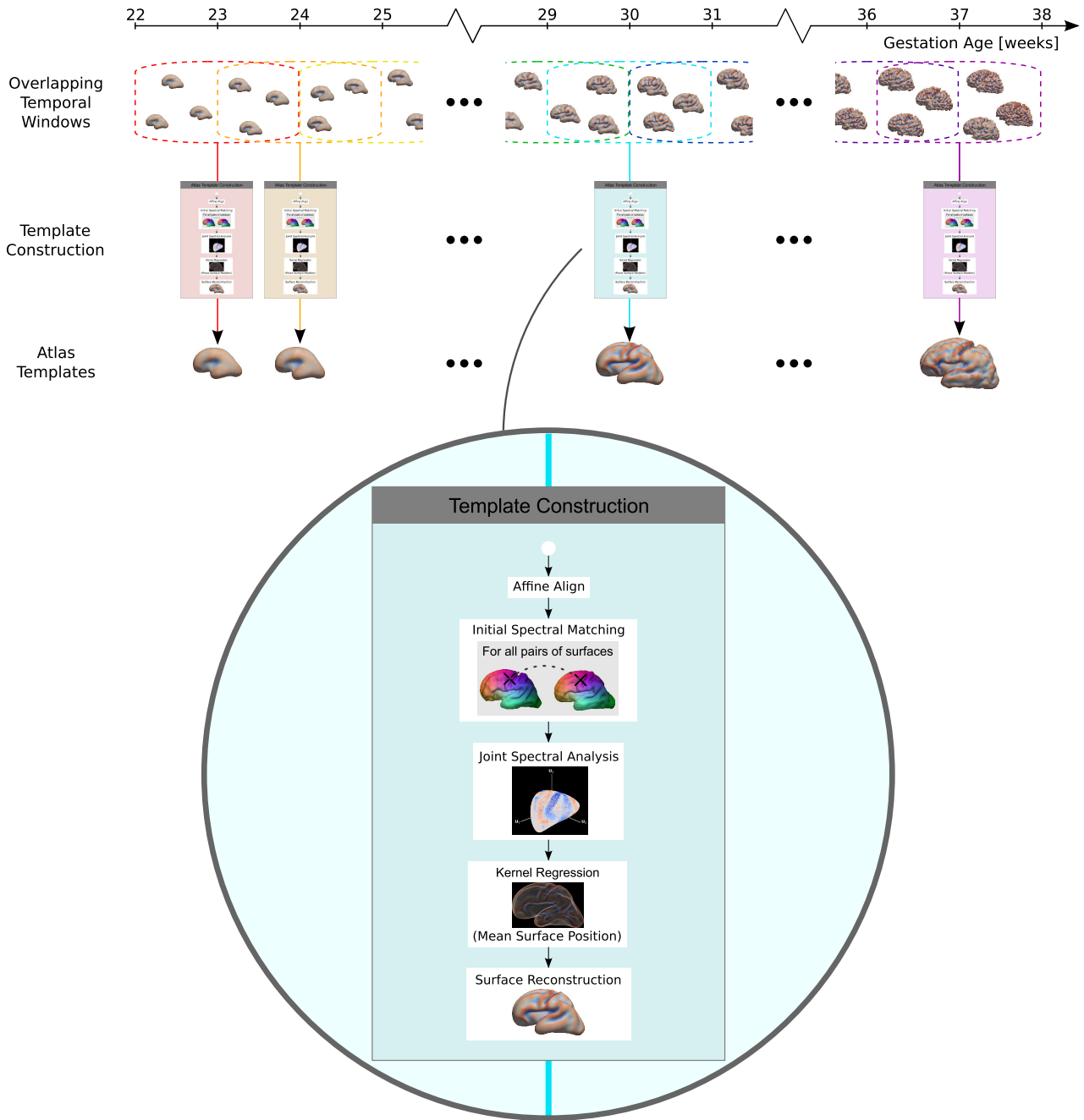


Figure 31: **Framework.** An average cortical surface template is constructed for each week of gestation, with all subjects within a week of the target age contributing to the output. A spectral analysis yields spatial correspondences between a group of cortical surfaces, allowing the average surface position to be computed and a template surface constructed. Note the red/blue surface colour mapping depicts the mean curvature of the surfaces.

Solving the eigenvector problem $\mathbf{L} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^{-1}$ gives the diagonal eigenvalue matrix, $\mathbf{\Lambda} = \text{diag}(\lambda_0, \lambda_1, \dots, \lambda_{|V|})$ and the corresponding eigenvector matrix $\mathbf{U} = (\mathbf{u}^0, \mathbf{u}^1, \dots, \mathbf{u}^{|V|})$, where each eigenvector $\mathbf{u}^i = (u_{0i}, u_{1i}, \dots, u_{|V|i})^T$ is a column of \mathbf{U} . Each of the eigenvectors or eigenmodes, \mathbf{u}^i , can be viewed as a pattern of vibration for the surface model S (Fig. 32) associated with the resonant frequency λ_i . The first eigenvector is not a vibration mode as $\lambda_0 = 0$ and is thus discarded. These eigenmodes give a useful parameterisation which can be used to match points on another similarly shaped surface. For correct matching, eigenmodes may need to be corrected for their sign ambiguity, multiplicity, and perturbation in isometry (Lombaert et al., 2011; 2013a).

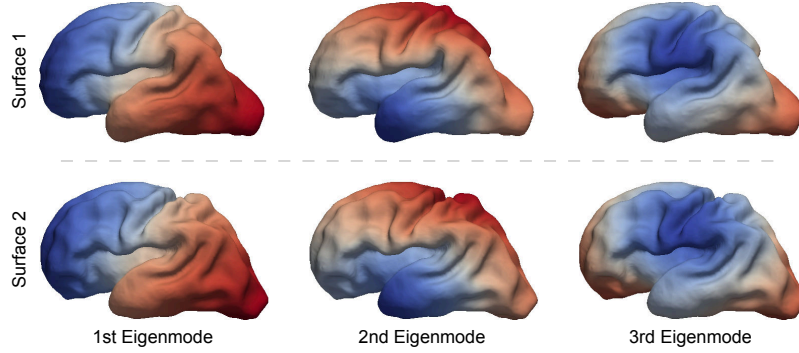


Figure 32: **Surface Eigenmodes.** The 1st three vibration modes of two surfaces are shown. Each mode depicts a pattern of displacement for a resonant frequency of the surface between poles, colour mapped to blue and red respectively. Figure adapted from Lombaert et al. (2013b).

To match points between surfaces the concept of a spectral representation is useful (Fig. 33). The spectral representation of S is denoted by a k -dimensional embedding, where the i th vertex has the spectral coordinates $\mathbf{u}'_i = (u_{i1}, u_{i2}, \dots, u_{ik})^T$, which is obtained from a row of the truncated matrix $\mathbf{U}_k = (\mathbf{u}^1, \mathbf{u}^2, \dots, \mathbf{u}^k)$. A correspondence map c gives the point-wise correspondences between two surfaces S_I and S_J . For each vertex $\mathbf{x}_i \in S_I$, with spectral coordinates \mathbf{u}'_i , the corresponding vertex $\mathbf{y}_j \in S_J$, with spectral coordinates \mathbf{v}'_j , is determined by the shortest Euclidean distance in the spectral domain, $c(i) = j = \operatorname{argmin}_k \|\mathbf{u}'_i - \mathbf{v}'_k\|$. To achieve

a more accurate match, surface descriptors such as mean curvature or the surface normal direction may be added to the embedding as extra dimensions (Lombaert et al., 2011). This ensures that a sulcus (or gyrus) from one surface matches with a nearby sulcus (or gyrus) from another surface. The values for each descriptor may be scaled in order to weight their influence when finding correspondences.

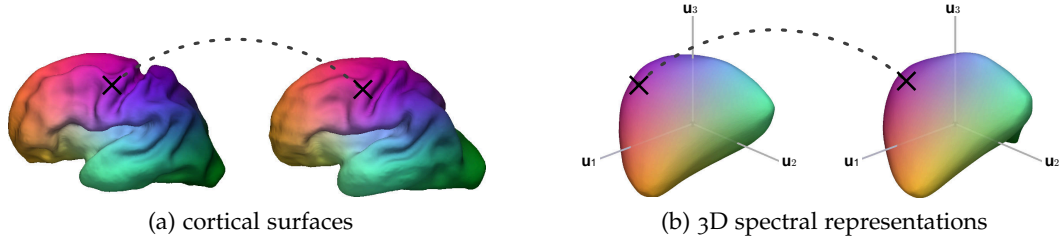


Figure 33: **Spectral Matching.** Vertex-wise correspondences between two surfaces (a) are given by the shortest Euclidean distance in the spectral domain (b). Colour mapping is given by the first three spectral coordinates which are mapped to RGB channels respectively. Figure adapted from Lombaert et al. (2013b).

This nearest neighbour approach to define an initial correspondence map c between two surfaces (Section 4.2.3) lacks regularisation and the matched locations may be incoherent, i.e. two vertices that are close on one surface may match with vertices that are distant on another surface. This can cause irregularities in the embedded surfaces (Section 4.3.4), particularly when incorporating scalar information into the embedding, which is not determined by surface position (e.g. mean curvature), or where scalar data may be noisy.

To address this, edge-based smoothing (Field, 1988) is used in the spectral domain to regularise the connections between two surfaces S_I and S_J , given an initial correspondence map $c = c_{IJ}$ (Algorithm 3). For each vertex $x_i \in S_I$, with corresponding vertex $y_{c(i)} \in S_J$ and corresponding spectral coordinate values $\mathbf{v}'_{c(i)}$, a smooth spectral coordinate $\bar{\mathbf{v}}'_i$ is first obtained. This is achieved by taking a weighted average of the neighbouring spectral coordinates $\mathbf{v}'_{c(j)}$, based on the edge set E_I ,

$\bar{\mathbf{v}}'_i = \sum_{j \in N(i)} w_j \mathbf{v}'_{c(j)}$. A more regular correspondence map can then be defined $c(i) = \operatorname{argmin}_k \|\bar{\mathbf{v}}'_i - \mathbf{v}'_k\|$.

Algorithm 3: Regularise correspondences.

Data: S_I, S_J , correspondence $c = c_{IJ}$

for $i = 1$ **to** $|V_I|$ **do**

$N(i) = \{j : e_{ij} \in E_I\}$
 $\bar{\mathbf{v}}'_i = \sum_{j \in N(i)} w_j \mathbf{v}'_{c(j)}$
 where $w_j \propto 1 / \|\mathbf{v}'_{c(i)} - \mathbf{v}'_{c(j)}\|^2$ and $\sum_{j \in N} w_j = 1$

for $i = 1$ **to** $|V_I|$ **do**

$c(i) = \operatorname{argmin}_k \|\bar{\mathbf{v}}'_i - \mathbf{v}'_k\|$

4.2.4 Joint spectral analysis

In (Lombaert et al., 2013b), a novel, dual-layered surface graph is built from two meshes interconnected by a preliminary correspondence map generated by spectral matching (Section 4.2.3). A joint spectral analysis of the connected surfaces produces a set of shared eigenmodes as opposed to two independent sets of eigenmodes. This produces a single parameterisation of the surfaces.

In this work, this approach is extended to k subjects by building a combined graph for surface graphs S_1, \dots, S_k , allowing a joint spectral analysis of all surfaces simultaneously. A multi-layered surface graph $S_c = \{\mathbf{V}_c, \mathbf{E}_c\}$ is formed where $\mathbf{V}_c = \mathbf{V}_1 \cup \mathbf{V}_2 \dots \cup \mathbf{V}_k$ is a concatenation of all mesh vertices and \mathbf{E}_c is the combined connectivity matrix. \mathbf{E}_c preserves the intra-surface vertex connectivity defined in surface graphs S_1, \dots, S_k as well as incorporating additional inter-surface connections. For each pair of surfaces S_I and S_J , an edge is formed between each vertex $\mathbf{x}_i \in S^I$ and its corresponding vertex $\mathbf{y}_{c(i)} \in S_J$, where $c = c_{IJ}$ is a preliminary correspondence map between S_I and S_J . As surfaces are linked pairwise, the

number of inter-surface edges may be far greater than the number of intra-surface edges. To compensate for this, a parameter ϕ is introduced to control the relative weighting of intra and inter-mesh edges. Intra-mesh edges are weighted as before, by the inverse squared Euclidean distance, $w = \|\mathbf{x}_i - \mathbf{x}_j\|^{-2}$. For inter-surface connections, the weighting $w = \frac{3\phi}{k-1} \|\mathbf{x}_i - \mathbf{y}_{c(i)}\|^{-2}$ was chosen. Note, $\phi = 1$ gives an equal weighting between intra-surface and inter-surface edges. For a triangular surface mesh S , with V vertices, F triangles, E edges and e connections to $k - 1$ other surfaces, the ratio $E : e$ is given:

1. $2E = 3F$ (number of half edges in the graph)
2. $V - E + F = 2$ (Euler's characteristic for a zero genus polyhedron)
3. $e = V(k - 1)$ (one connection for each vertex to every other surface)

$$V - E + \frac{2}{3}E \approx 0 \quad (\text{given that } E > V \gg 2)$$

$$V - \frac{1}{3}E \approx 0$$

$$E \approx 3V$$

$$E \approx \frac{3}{k-1}e$$

The spectral decomposition of the graph Laplacian $\mathbf{L}_c = \mathbf{U}_c \mathbf{\Lambda}_c \mathbf{U}_c^{-1}$, built from the multi-layered surface graph S_c , gives a set of shared eigenmodes \mathbf{U}_c and an orthonormal basis where surfaces S_1, \dots, S_k are co-aligned (Fig. 34).

4.2.5 Kernel Regression

The coordinate system provided by the shared eigenmodes \mathbf{U}_c (obtained from a joint spectral analysis of a group of surfaces) is referred to as the spectral domain,

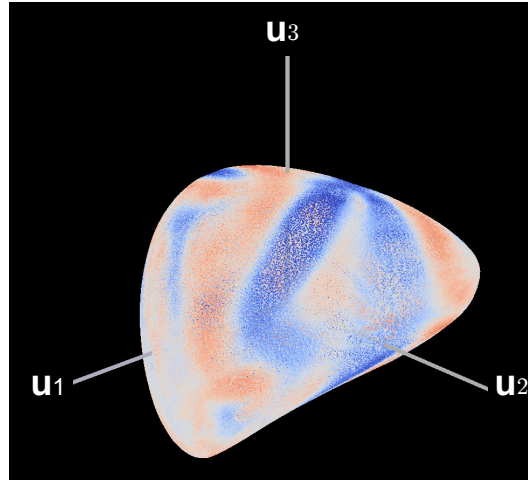


Figure 34: **3D Spectral Embedding of a Group of Surfaces.** Embedded vertices of 18 cortical surfaces are shown (29 – 31 weeks GA). Colour mapping depicts the mean curvature at each point on the original surfaces (red, convex; blue, concave).

or spectral embedding (Fig. 34). This coordinate system allows the value of a scalar variable associated with the surfaces to be estimated at matched locations on the cortex. For example, each embedded vertex may have associated shape descriptors (e.g. mean curvature, sulcal depth) or regional anatomical labels (e.g. Brodmann areas) as well as spatial location.

Kernel regression (Nadaraya, 1964) is a non-parametric technique that can be used to estimate the value of one of these scalar variables y , at any given spectral location, by averaging the values of y associated with nearby vertices, using a kernel based weighting function w . Let \bar{y}_i denote the expected value of y at spectral location \mathbf{u}'_i and $N(i)$ a set of vertices close to \mathbf{u}'_i . Then \bar{y}_i is given by a weighted average of y for all vertices in $N(i)$, i.e. $\bar{y}_i = \sum_{j \in N(i)} y_j w(i, j) / \sum_{j \in N(i)} w(i, j)$. The vertices of each surface are evenly sampled in the spatial domain, however this is not the case in the spectral domain. Therefore, $N(i)$ was defined as the k nearest vertices to \mathbf{u}'_i in order to achieve uniform spatial smoothness.

The weighting function $w(i, j) = w_u(\mathbf{u}'_i, \mathbf{u}'_j) w_t(t_j) / \mathcal{C}(j)$ was chosen which has three terms: a spatial weighting w_u , a temporal weighting w_t and a count func-

tion \mathcal{C} . $\mathcal{C}(j)$ simply counts the number of vertices in $N(i)$ that originate from same surface mesh as j . This removes the influence of different vertex densities for each surface at different locations within the embedding. A Gaussian kernel was used for the spatial weighting $w_u(\mathbf{u}'_i, \mathbf{u}'_j) = \exp(-\|\mathbf{u}'_i - \mathbf{u}'_j\|^2 / 2\sigma_u^2)$ and the temporal weighting $w_t(t_j) = \exp(-(t' - t_j)^2 / 2\sigma_t^2)$, where t' is a target [GA](#). The temporal smoothness and spatial smoothness of \bar{y}_i are controlled by σ_t and σ_u respectively. σ_u can be chosen adaptively, so that the average distance of the k vertices is equal to the full width at half maximum of the kernel, $\sigma_u = \frac{1}{2\sqrt{2\ln 2}} \frac{1}{k} \sum_{j \in N(i)} \|\mathbf{u}'_i - \mathbf{u}'_j\|$. Note, the parameter k now effectively controls the spatial smoothness. Temporal smoothness, σ_t , can be fixed for all time-points or adjusted within regions where the temporal sampling is lower ([Serag et al., 2012a](#)). This will sacrifice temporal sharpness when estimating \bar{y}_i , in order to smooth out variations in cortical anatomy by incorporating more subjects.

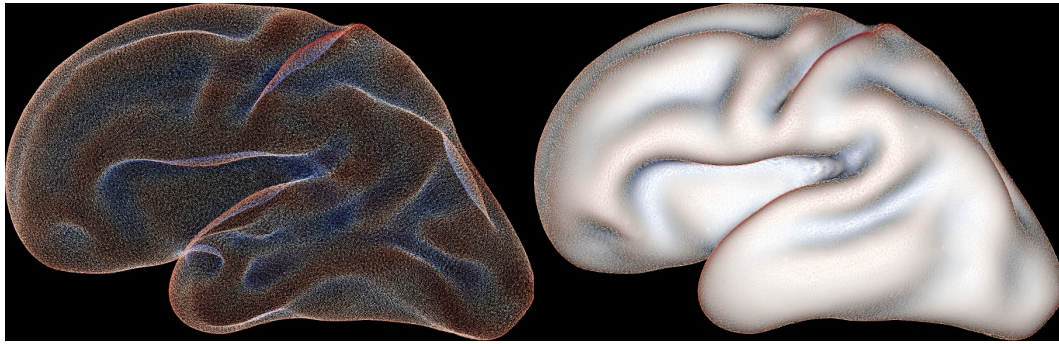
4.2.6 Surface Reconstruction

For each embedded vertex i (Fig. [34](#)), the average cortical surface position, $\bar{\mathbf{x}}_i$, can be estimated using kernel regression, given its spectral location \mathbf{u}'_i and a target gestational age t' . This is achieved by treating each spatial dimension as a separate scalar field. The estimated location for each vertex, $\bar{\mathbf{x}}_i$, gives a smooth point sampling of the average surface (Fig. [35a](#)). Reconstructing surfaces from point data is a well-studied problem driven by the need for measurement and visualisation of real world objects that are geometrically sampled, for example, with 3D range scanning technology. For this work, the popular screened Poisson surface reconstruction algorithm ([Kazhdan and Hoppe, 2013](#)) was used to reconstruct an average surface template from $\bar{\mathbf{x}}$ (Fig. [35b](#)). This technique also requires an estimate

of the average surface normal direction, $\bar{\mathbf{N}}_i$, for each point sample of the average surface boundary, $\bar{\mathbf{x}}_i$.

Averaging orientations is a non-trivial problem given the periodicity of the domain. At a corresponding location on the cortical surface, it is not biologically plausible to have large variations in orientation between subjects. By assuming that the maximum deviation from the average is bounded by π , the difference in orientation can be taken to be the shortest path on the unit sphere, ignoring any issues due to “wrapping”.

A widely used but naive approach to this problem is the “vector sum algorithm” which averages each x , y , z component of the surface normal vector, which is equivalent to minimising the squared chord length on the unit sphere associated with each observation (Olson, 2011). Assuming Gaussian noise, minimising the arc length is a more accurate method since the length of the arc is proportional to the difference in orientation. However, this level of accuracy is unnecessary as the chosen method is robust to inaccuracies in normal direction (Kazhdan, 2005). $\bar{\mathbf{N}}_i$ is therefore estimated from the original surface normal directions associated with each point, weighted using kernel regression, treating each dimension as a separate scalar value, and normalising the results.



(a) point sampling of the average cortical boundary

(b) reconstructed surface

Figure 35: **Average cortical surface template for a 30 week fetus.** Kernel regression in the spectral domain is used to find the average surface position. This gives a point sampling of the average cortical surface (a). Poisson surface reconstruction, is then used to extract a closed surface (b), from (a).

4.2.7 Temporal Correspondence

Establishing temporal correspondences between embeddings is important as it permits a longitudinal analysis of cortical growth. For each embedding, a number of the contributing surfaces will also contribute to a neighbouring embedding, as a spectral analysis was performed across several overlapping time windows (Fig. 31). For the surfaces that are not shared with a particular neighbouring embedding, the position of each of their vertices can be estimated. This is achieved by treating the neighbouring embedding coordinates as a labelling associated with each surface vertex within the current embedding. The neighbouring coordinates can then be estimated at unlabelled vertex locations using kernel regression (Section 4.2.5) by restricting the k nearest vertices to the set of vertices already labelled.

This procedure is the equivalent of multi-atlas label propagation (Rohlfing et al., 2003; Heckemann et al., 2006), adapted for a set of embedding coordinates. This can also be iterated to propagate a set of coordinates to every surface within the

cohort (Fig. 36). This ultimately enables all surfaces to be embedded in a single coordinate system and for correspondences to be found between subjects of different ages, even if these differ significantly in their cortical surface geometry (Fig. 37b). Additionally, the propagation of nineteen anatomical labels derived from a neonatal atlas set (Gousias et al., 2012) is also demonstrated. The MRI volumes for fetuses whose age was greater than 36 weeks GA were first parcellated by registering the neonatal atlas set to each volume and fusing the atlas labels. Each surface voxel was then assigned the anatomical label of its nearest voxel from the corresponding MR volume. This formed a surface atlas set for stepwise propagation to all other cortical surfaces (Fig. 37a).

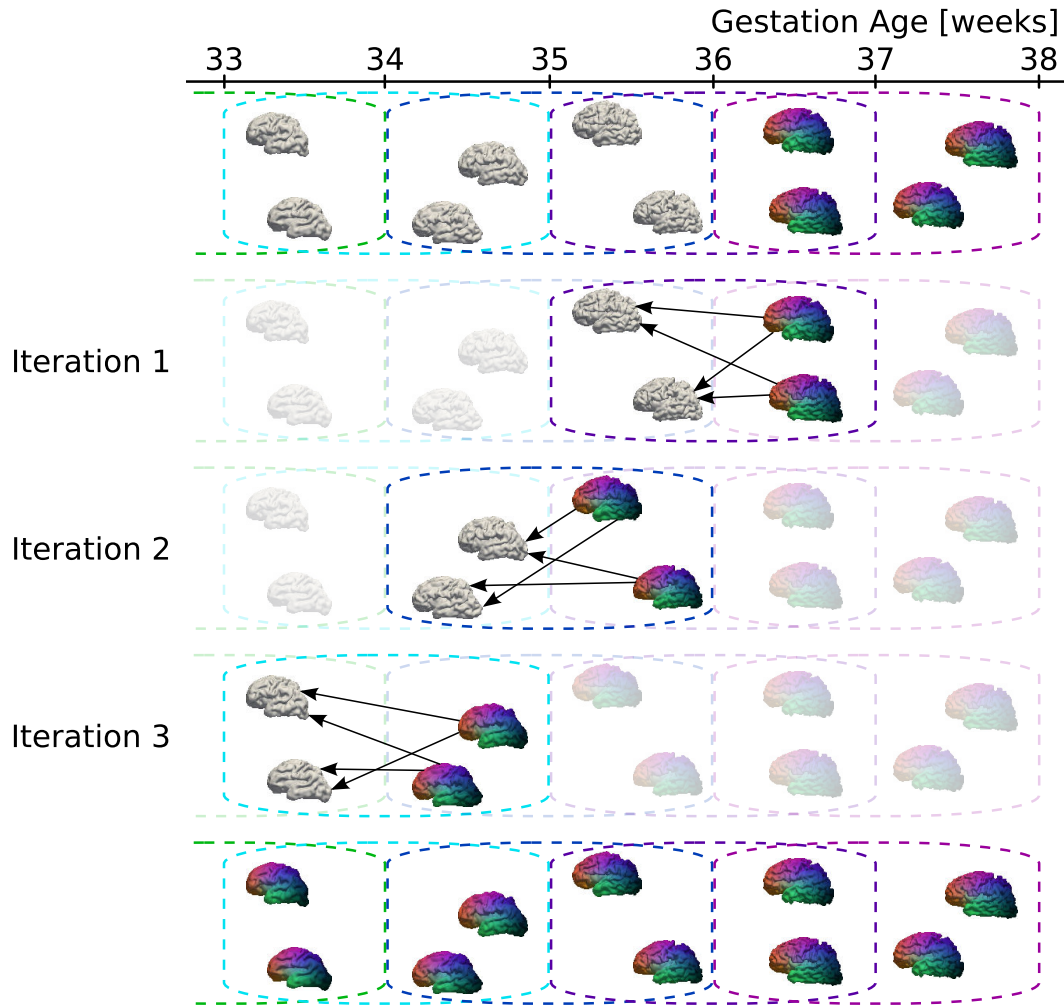


Figure 36: **Stepwise Embedding Coordinate Propagation.** This diagram illustrates the propagation of a set of embedding coordinates ($t' = 37$) to the cortical surfaces of younger subjects. Surface colour mapping shows the first 3 embedding coordinates encoded as RGB values. These coordinates can be viewed as a surface labelling, which may be propagated to neighbouring surfaces via surfaces that are shared across neighbouring embeddings. At each iteration, labelled subjects are treated as atlases, whose labels are propagated to any unlabelled surfaces within the same embedding using kernel regression. This process can be repeated iteratively until all surfaces are labelled, yielding a shared parameterisation for all surfaces. The dashed lines enclose the surfaces used to create each of the embeddings.

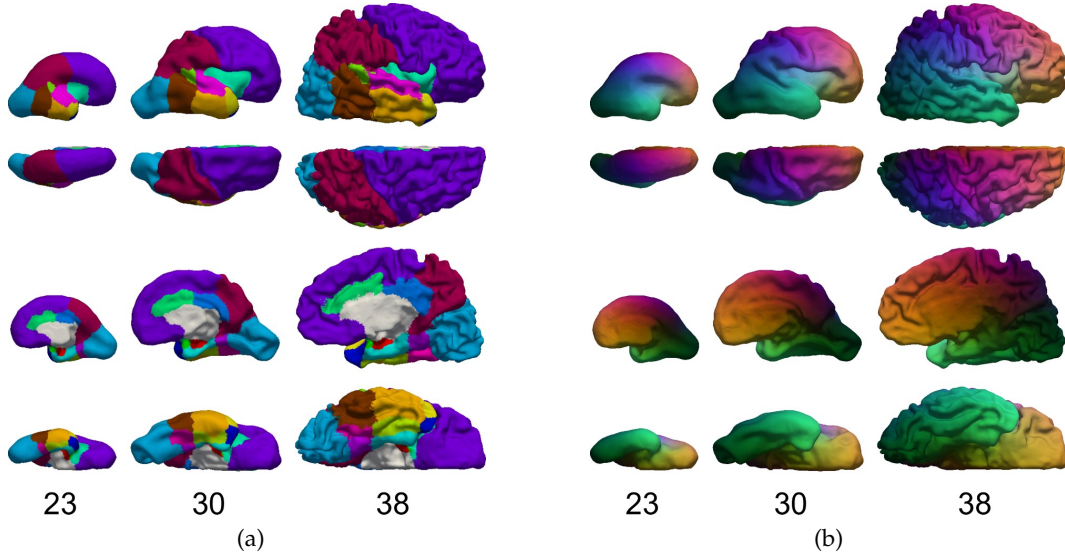


Figure 37: **Multi-atlas Label Propagation.** Three examples of surfaces automatically labelled using kernel regression. (a) A 19 region parcellation. (b) The first embedding coordinates (for $t' = 37$) mapped to RGB channels respectively (See Fig. 36).

4.3 RESULTS & DISCUSSION

4.3.1 Model Parameters

Graph Laplacian Edge Weighting

The parameter ϕ was introduced to control the weighting of inter-mesh edges relative to intra-mesh edges when constructing the graph Laplacian for a multi-layered surface graph ($\phi = 1$, gives an equal total weighting to both types of edges). The value $\phi = 0.1$ was chosen empirically to achieve a compromise between embedding regularity and sulcal alignment accuracy, determined using the validation methods discussed later in this section, and also by visual inspection of the 3D

representations of the final embeddings. This value could be tuned further to increase sulcal alignment. However, increasing the value of ϕ substantially can have a detrimental effect on embedding regularity.

Kernel Parameters

The kernel parameters, $\sigma_t = 1$ and $k = 100$ were chosen by observing the smoothness of the final atlas templates. As the fetal subjects are fairly evenly distributed with age (Section 1.4), a constant temporal smoothness value $\sigma_t = 1$ was chosen. Both parameters are dependent on the size of the dataset. Additionally, the value chosen for k will depend on the required resolution of the surface meshes.

4.3.2 Sulcal Alignment Accuracy

To determine the precision of the sulcal alignment a common anatomical feature was manually delineated for each subject; the central sulcus. A “polyline” (a line defined by a successive sequence of vertices) was extracted by traversing each surface along the centre of the sulcus. These lines were then mapped pairwise between subjects within each embedding using the established spectral correspondence (Fig. 38). The similarity between the central sulci after alignment was then assessed using two established measures: the discrete Fréchet distance (F_d) (Alt and Godau, 1995) and the average Fréchet distance (F_a) (Brakatsoulas et al., 2005). The central sulcus was selected as it is a well-defined feature that can be easily delineated. It is the one of the first sulci to form, and is therefore detectable at an early GA (25 weeks) (Garel et al., 2001) and was observed in most of the subjects within the cohort (> 75).

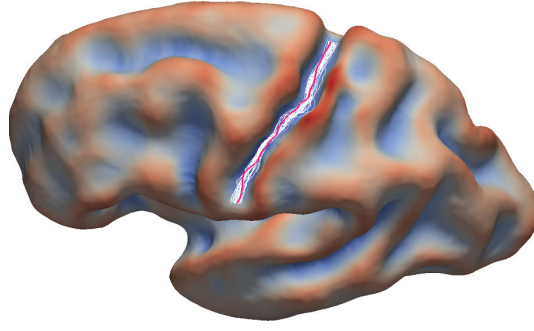


Figure 38: **Aligned Central Sulci.** An example of a manual delineation of the central sulcus (depicted in red), along with 38 central sulci delineations mapped from other subjects, using the established spectral correspondence (shown in white).

The Fréchet distance was chosen as it takes into account the location and ordering of the points along a curve, unlike the popular Hausdorff distance, and is therefore more sensitive to discontinuities in mapped polylines. Furthermore, the Fréchet distance has a well defined average (Brakatsoulas et al., 2005) while the Hausdorff distance is a maximum distance measure, increasing its sensitivity to outliers. The Fréchet distance is often referred to as the “dog walker’s distance” as it can be described intuitively using the scenario of an owner walking its dog. It is given by the minimum leash length needed to connect a dog and its owner as they traverse two separate paths. The owner and dog may proceed at any speed, and even stop, but may not backtrack. Formally, let S be a metric space, then a curve A in S is a continuous mapping from the unit interval $A : [0, 1] \rightarrow S$. A reparameterisation α is a continuous monotonic mapping $\alpha : [0, 1] \rightarrow [0, 1]$. If A and B are two curves in S , then their Fréchet distance $F(A, B)$ is defined as the infimum, over all reparameterisations α and β , of the maximum distance in S between $A(\alpha(t))$ and $B(\beta(t))$, for all $t \in [0, 1]$.

$$F(A, B) = \inf_{\alpha, \beta} \max_{t \in [0, 1]} \{d(A(\alpha(t)), B(\beta(t)))\}.$$

Here d is a distance function of S . The sulcal alignment error was compared with that given by spherical demons (Yeo et al., 2010) and two further baseline methods which register surfaces using transformation in 3D: Iterative Closest Point (ICP) registration using affine transformations and ICP using FFDs (with a control point spacing of 2.5 mm) (Rueckert et al., 1999). Alignment error was first compared within each embedding, i.e. between subjects of a similar age. The error between spectral matching and spherical demons was similar, with both outperforming the baseline methods, especially at older ages. The spectral alignment accuracy improved from 25 weeks onwards, with the lowest mean error, $F_a \approx 0.4$ mm, at 30 weeks (Fig. 39). This trend was replicated with spherical demons and may be explained by the central sulcus becoming more distinct as it transforms from a shallow depression into a deep sulcus, which is more easily matched. After this, the mean matching accuracy dropped for both methods as the cortical complexity increased, with the largest average error seen around 36 weeks ($F_a \approx 1$ mm).

The sulcal mapping accuracy was also compared between subjects of different ages, by first embedding both of their surfaces in the same coordinate system (Section 4.2.7). Accuracy was then compared when mapping sulcal delineations from younger subjects (≈ 27 weeks) to progressively older subjects (Fig. 40a) and also when mapping sulci for older subjects (≈ 37 weeks) to progressively younger ones (Fig. 40b). Again, performance of the two methods was similar and superior to the baseline methods.

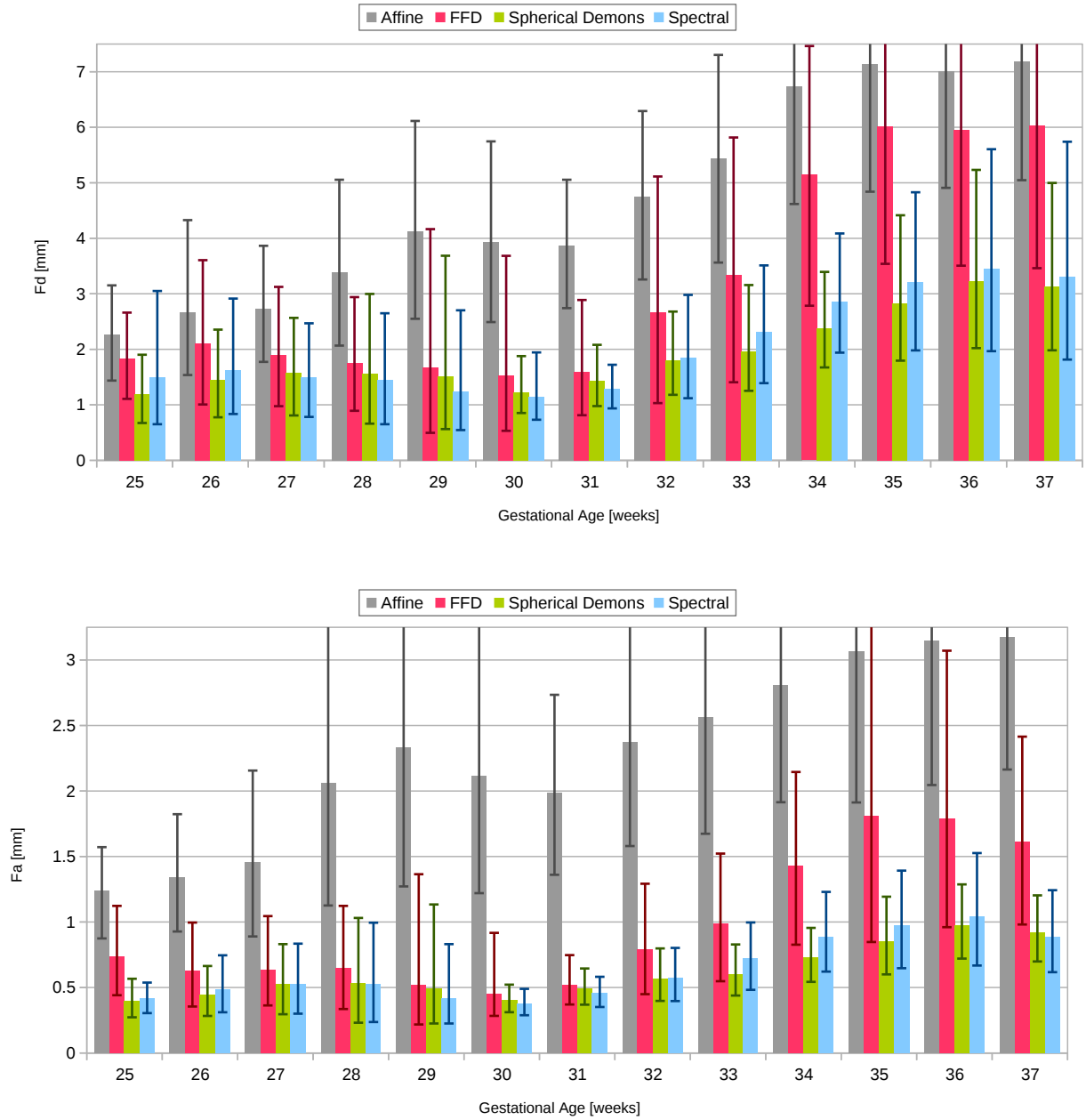


Figure 39: **Intra-embedding sulcal mapping error.** For each embedding, the sulcal delineations were mapped pairwise for all surfaces. The average alignment error is shown for four methods, quantified by the Fréchet distance, F_d (top), and the average Fréchet distance, F_a (bottom). Error bars show the standard deviation of the alignment error.

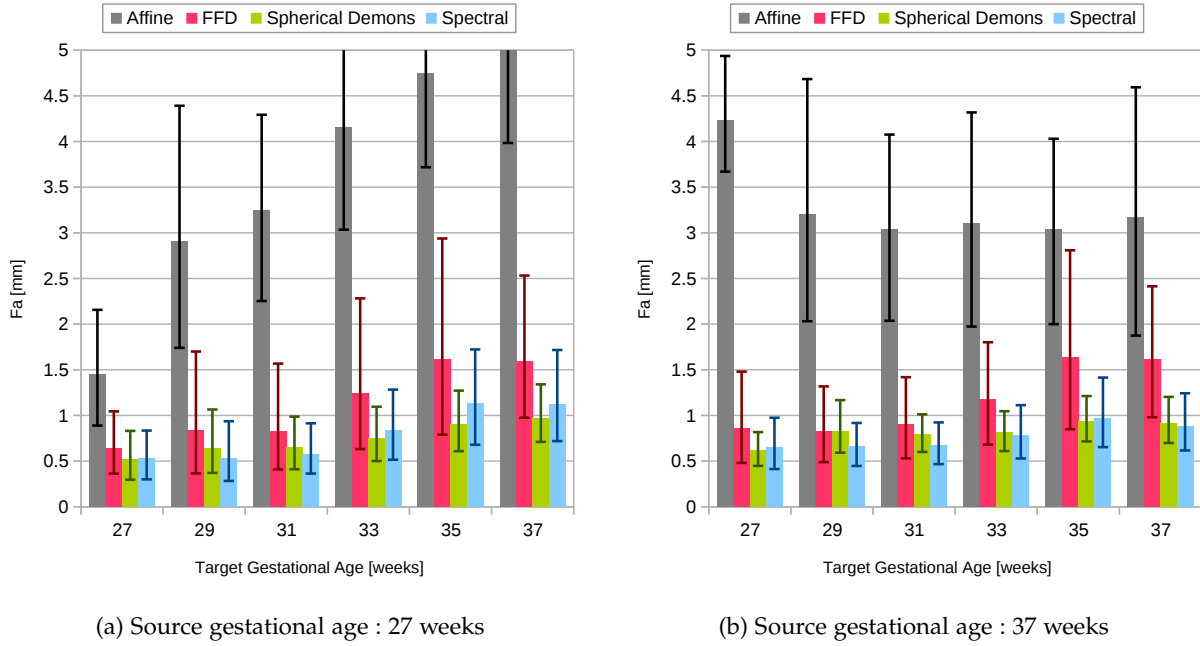


Figure 40: **Inter-embedding sulcal mapping error.** Central sulcus delineations were mapped between surfaces that contributed to different embeddings by first establishing temporal correspondences (Section 4.2.7). The average alignment error, quantified by the average Fréchet distance, F_a , is shown for four methods (a) when mapping delineations for younger source subjects (≈ 27 weeks) to progressively older target subjects and (b) when mapping delineations for older source subjects (≈ 37 weeks) to progressively younger target subjects. Error bars show the standard deviation of the alignment error.

4.3.3 Regional Overlaps

To assess the alignment accuracy over the whole brain surface, the overlap between different regions of the brain was also evaluated. Each surface was parcellated into 6 regions using a publicly available neonatal MR atlas database (Gousias et al., 2012) (Fig. 41). MR atlases were first registered to each subject's MR image volume using the FFD model (Rueckert et al., 1999), and the atlas labels were fused to create a volumetric parcellation for the subject. This parcellation was then refined using

the method of [Makropoulos et al. \(2014\)](#). Each surface vertex was then assigned the anatomical label of its nearest voxel from the corresponding [MR](#) image volume. Labels were propagated between each pair of surfaces within an embedding using the spectral (or spherical) correspondences, and their overlap measured using the Dice coefficient ([Sørensen, 1948](#)).

The results were again similar between the spectral method and spherical demons over all six regions (Table 6). As expected, higher Dice scores were observed in the lobar regions that have a larger surface area. The lowest Dice scores were observed for the insula, which is likely due to the absence of a clearly defined inferior anatomical boundary.

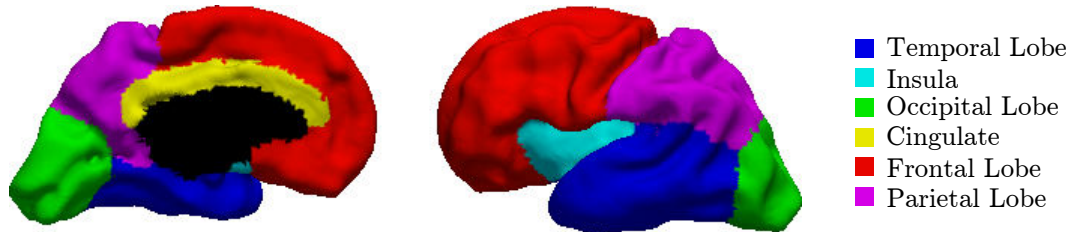


Figure 41: **Surface Parcellation.** An example of a cortical surface model parcellated into 6 regions.

| Region | Spectral | Spherical Demons |
|----------------|-----------------------|-----------------------|
| Temporal lobe | 0.937 (± 0.015) | 0.931 (± 0.015) |
| Insula | 0.747 (± 0.088) | 0.746 (± 0.084) |
| Occipital Lobe | 0.929 (± 0.024) | 0.914 (± 0.025) |
| Cingulate | 0.829 (± 0.045) | 0.823 (± 0.050) |
| Frontal Lobe | 0.964 (± 0.009) | 0.962 (± 0.011) |
| Parietal Lobe | 0.922 (± 0.020) | 0.913 (± 0.025) |

Table 6: **Cortical region overlap.** Region labels were propagated between pairs of cortical surfaces using the established spectral or spherical correspondences, before measuring their overlap using the Dice coefficient. Average Dice scores are summarised.

4.3.4 Embedding Regularity

The n th eigenvector of the Laplacian has the property of being smooth and monotonous over a surface between, at most, n poles (Lombaert et al., 2013b). This gives a diffeomorphic mapping between a surface and its spectral representation. The local structure of a surface graph is preserved in the embedding, i.e. an embedded vertex is central to its connected vertices. A simple experiment was conducted to ascertain whether the local structure of each surface graph is preserved when a group of surfaces are embedded using the methodology outlined in this chapter.

For each surface vertex i with spatial location \mathbf{x}_i and embedded position \mathbf{u}'_i , edge-based smoothing (Field, 1988) was used to find the spectral centroid, $\bar{\mathbf{u}}'_i$, of its neighbouring vertices j , i.e. $\bar{\mathbf{u}}'_i = \sum_{j \in N(i)} w_j \mathbf{u}'_j$, where $w_j \propto 1 / \|\mathbf{x}_i - \mathbf{x}_j\|$ and $\sum_{j \in N(i)} w_j = 1$. If the local structure is preserved, \mathbf{u}'_i and $\bar{\mathbf{u}}'_i$ should be close. The closest vertex to $\bar{\mathbf{u}}'_i$ was found in the spectral domain, with spatial position $\bar{\mathbf{x}}_i$, and the distance $\|\bar{\mathbf{x}}_i - \mathbf{x}_i\|$ was measured. This distance should be small if the vertex i is embedded close to its neighbouring mesh vertices.

This experiment was repeated for spectral matching with and without regularisation of the surface links (Section 4.2.3, final paragraph) and also for spherical demons replacing the spectral embedding coordinates with the method's associated spherical coordinates. For each graph, the percentage of voxels where $\|\bar{\mathbf{x}}_i - \mathbf{x}_i\| = 0$ was computed, i.e. the percentage of voxels where the local structure of the graph was preserved. The average percentage was high for spherical demons (99.57%) and for both variations of spectral matching (with regularisation, 99.91%, and without, 98.83%) (Table 7). For points where $\|\bar{\mathbf{x}}_i - \mathbf{x}_i\| > 0$, the mean distance was also computed and this was found to be, on average, around the mesh resolu-

tion for all methods (mesh edge length ≈ 0.5 mm). The average maximum distance was relatively high for spectral matching with irregular surface links (4.15 mm), however this decreased substantially with regularisation (0.724 mm) in line with spherical demons (0.949 mm).

| Measure | Spectral Matching | | Spherical Demons |
|------------|----------------------|----------------------|----------------------|
| | irregular | regularised | |
| Percentage | 98.35 (± 1.12) | 99.91 (± 0.09) | 99.57 (± 0.34) |
| Mean[mm] | 0.65 (± 0.14) | 0.40 (± 0.03) | 0.41 (± 0.032) |
| Max[mm] | 4.15 (± 2.55) | 0.72 (± 0.23) | 0.95 (± 0.38) |

Table 7: **Embedding regularity.** Percentage of mesh vertices where structure is preserved (see text).

These results show that the local structure is largely preserved for all methods and that most irregularities are small. However, there was a small number of large irregularities with spectral matching without regularisation, particularly at older gestational ages as the cortical surfaces become more complex (Fig. 42).

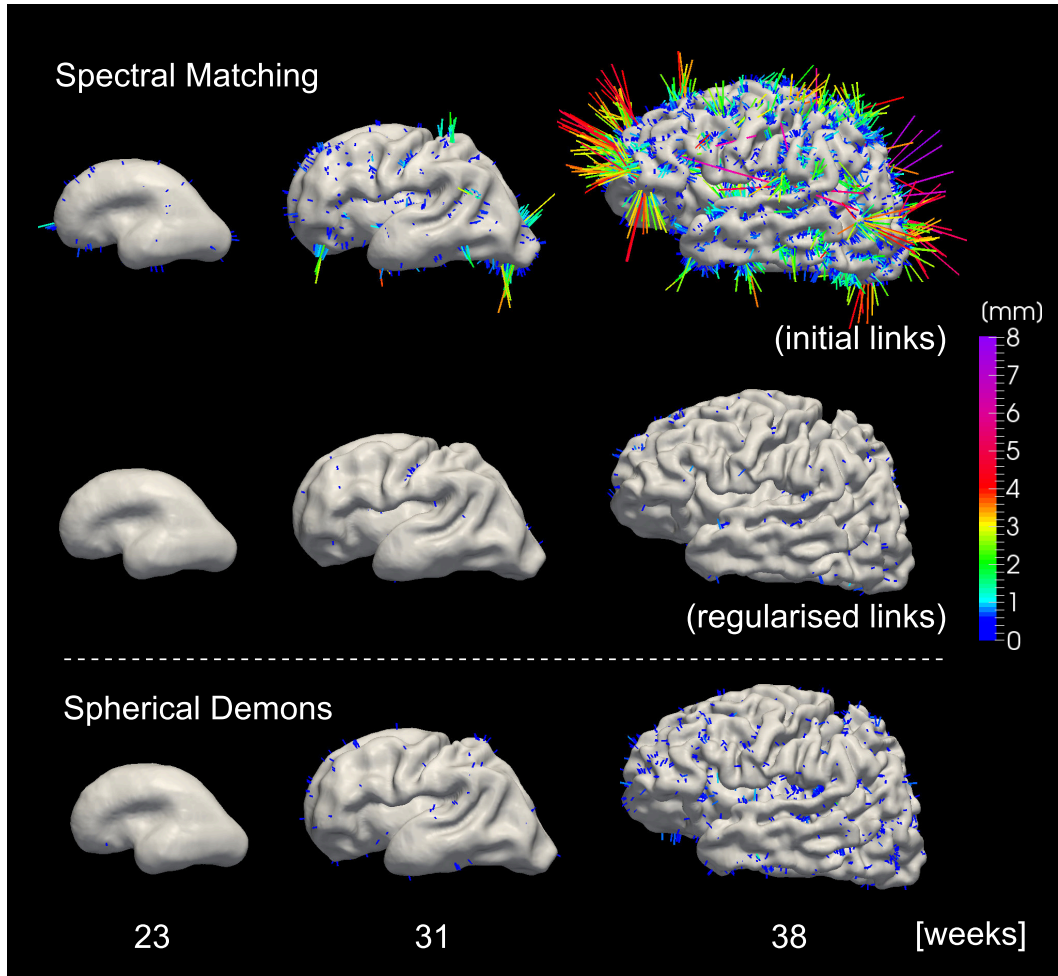


Figure 42: **Embedding regularity visualisation.** Each line signifies a vertex where the local structure of the surface was not preserved in the embedding. The length of the line is proportional to the distance that the embedded vertex is from the centroid of its neighbouring vertices, when both are mapped back to the spatial domain (the exact distance in mm is given by the colour mapping). Note the irregularities seen when the initial surface links are not regularised.

4.3.5 Average Cortical Surfaces

Average cortical surfaces were generated for each week of gestation from the spectral embeddings as well as via spherical demons (Fig. 43). Generating an average surface from a spectral embedding was discussed in Section 4.2.6, and for spherical

demonstrates a similar approach was taken. A regularly sampled spherical mesh gives the correct topology and provides the vertex connectivity for an average surface. This was readily generated by subdividing a platonic solid mesh and projecting the vertices back onto the sphere. Kernel regression (Section 4.2.5) can then be used to find an average surface position at each vertex. Transforming the spherical mesh back to the spatial domain, using the average surface position at each vertex, gives an average cortical surface.

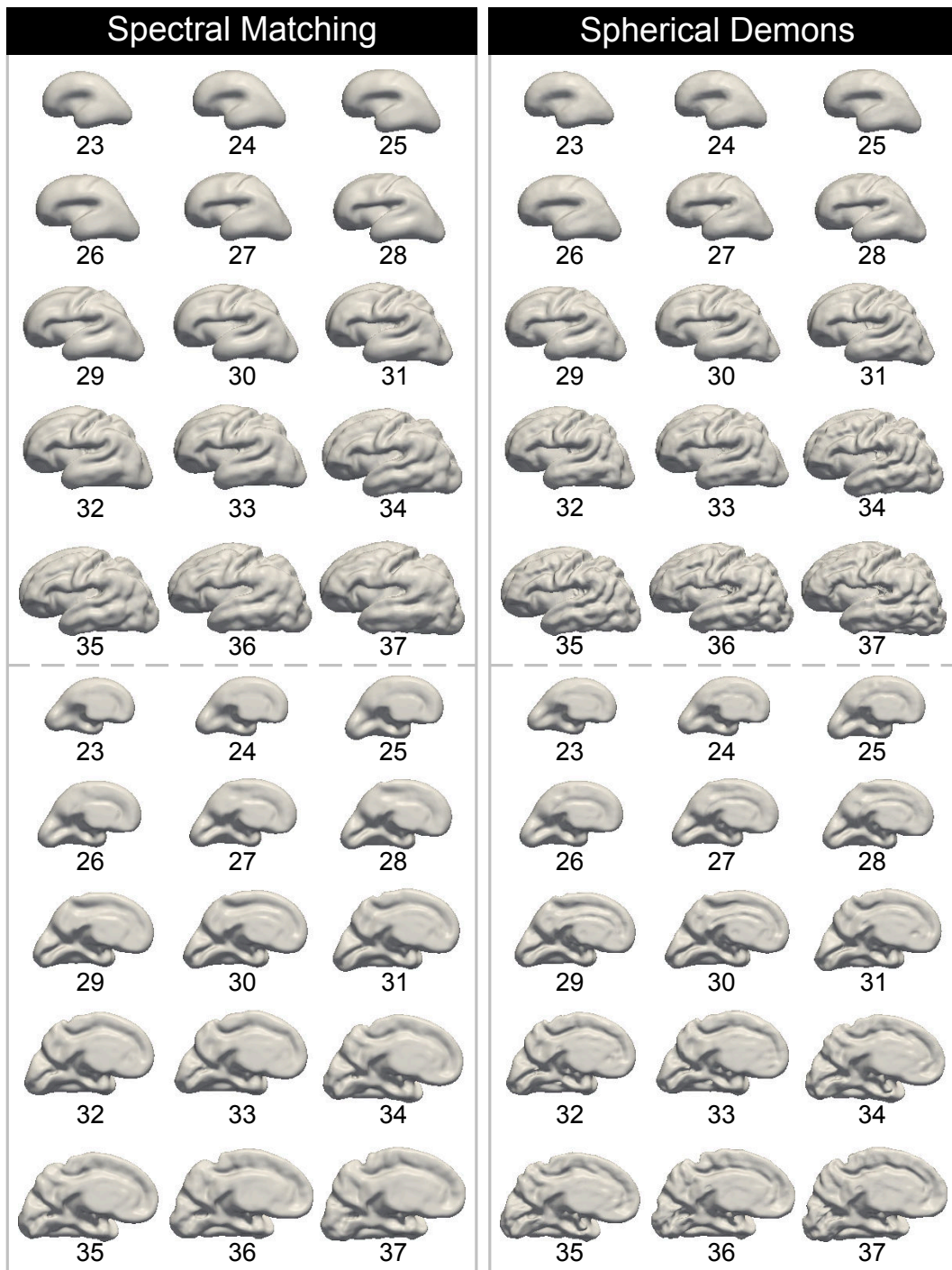


Figure 43: **Average cortical surface templates.** Cortical surface templates were constructed for each week of gestation, for both spectral embeddings and spherical demons. Both methods produced visually similar templates, capturing an estimate of the average growth for the cohort.

For both methods, the timing of the sulcation observed in the average surfaces matched previous observations from MR images (Garel et al., 2001). At 23 weeks, the global shape of the brain has already formed and the sylvian, calcarine, parieto-occipital and hippocampic fissures, as well as the cingulate, are all well defined on the average surfaces. The first sulcus to form is the central sulcus, which is detectable from 24 weeks and very prominent after 27 weeks. By 29 weeks the precentral and postcentral sulci, frontal sulci, superior temporal sulcus and intra-parietal sulcus are all clearly visible on the cortical surfaces. Visually the atlas templates are similar to those previously constructed by Habas et al. (2012) and Clouchoux et al. (2013).

Atlas templates constructed using the spherical demons pipeline produced “sharper” templates at the expense of regularity; deeper sulcal pits were observed as well as discontinuities at the posterior end of the Sylvian fissure. This suggests that the spectral pipeline has a stronger regularising effect, as indicated by the experiments conducted in Section 4.3.4. A trade-off exists between regularity and smoothness and these differences may easily be accounted for by the chosen parameter values for each of the pipelines. With the spectral method, the initial regularisation of surface links (Algorithm 3) and the intra-inter edge weighting (ϕ) will affect the smoothness of the templates, as well as the choice of kernel parameters (σ_t , k). These parameters were not fine tuned extensively and more experimentation is needed to ascertain whether a finer sulcal alignment may be achievable while preserving regularity.

4.3.6 Variability of average cortical surfaces

To estimate the variability of generated average surface templates, given a sample size N , two distinct templates were computed from two disjoint subsets, both of size N , of the cortical surface models within a temporal window, and then the average distance between them was measured. This was performed for a range of sample sizes (2-11) and 4 target ages (24, 28, 32 and 36 weeks). For each sample size and target age, 10 iterations were performed using randomly generated disjoint pairs of subsets from the available data. The average of these 10 iterations is shown as a point in Fig. 44.

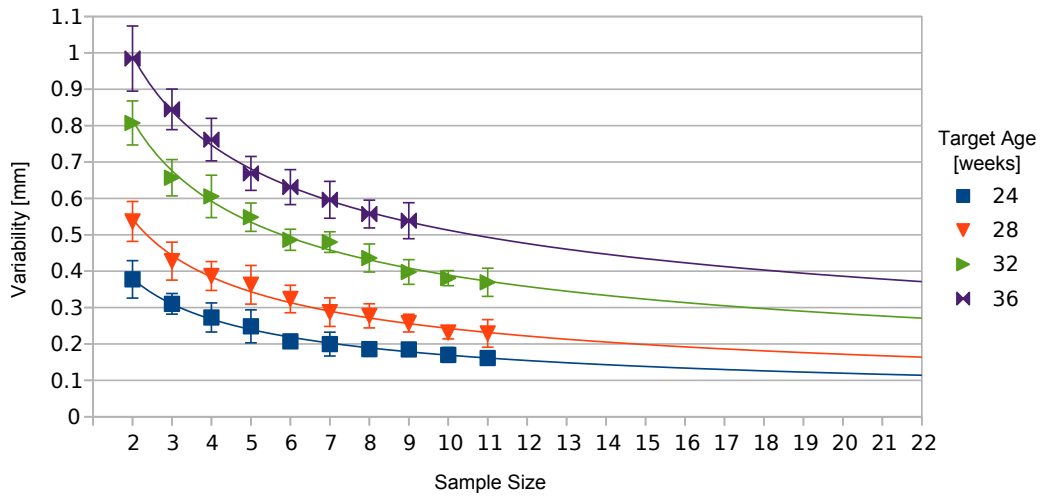


Figure 44: **Effect of sample size on the variability of generated atlas templates.** The average distance between distinct templates generated from disjoint subsets of the cortices was computed to give an estimate of the variability of generated surface templates. The points plotted show an average of 10 iterations for a particular sample size and age, with the error bars showing the standard deviation of the 10 iterations. The lines show a power law fit of the data for each target age.

At 36 weeks, where anatomical variation is the largest (for the target ages examined), the average surface generated was consistent to around $0.54 \text{ mm} \pm 0.04$, for $N = 9$. The relationship between sample size and the average distance be-

tween generated average surfaces is well-modelled by a power law curve ($R^2 = 0.992 \pm 0.003$) for sample sizes up to 11. Extrapolating this relationship to sample sizes above 11, the consistency of the average surfaces generated from the complete dataset was predicted (Table 8). At 36 weeks, for the full sample size ($N = 18$), it is therefore estimated that the atlas templates generated are consistent to around 0.40 mm, with smaller values for ages less than 36 weeks.

| Target Age (t) [weeks] | 24 | 26 | 28 | 30 | 32 | 34 | 36 |
|----------------------------|-------|-------|-------|-------|-------|-------|-------|
| Sample Size (N) | 22 | 16 | 22 | 22 | 22 | 16 | 18 |
| Variability [mm] | 0.114 | 0.159 | 0.164 | 0.237 | 0.271 | 0.396 | 0.403 |

Table 8: **Atlas template variability prediction.** By extrapolating the power law fit of the variability measured to sample sizes greater than 11 (see Fig. 44), it is possible to estimate the variability of surface templates generated using the full dataset size.

4.4 CONCLUSION

In this chapter, a spatio-temporal surface atlas of the developing fetal brain was constructed. Establishing a set of preliminary correspondences between cortical surfaces allowed a group of surface graphs to be linked. This permitted an analysis of their graph Laplacian spectrum as a single entity, yielding an orthonormal basis for modes of an entire group of surfaces. This gave accurate and unbiased correspondences between surfaces in a fraction of the time required by spherical registration methods.

Average cortical surface templates were constructed for each week of gestation by averaging the surface position over the spectral embeddings. These accurately capture the global shape and the timing of sulcal development over the cortex.

Furthermore, the prominence of visible sulcation suggested a good alignment between embedded surfaces. The average surfaces generated were reproducible, with an average boundary variability of around 0.54 mm, for disjoint subsamples of the dataset ($N = 9$) and a target age of 36 weeks, where sulcal variation is high.

The sulcal alignment was found to be significantly better than 3D deformation methods and comparable to spherical demons, a state of the art 2D surface registration technique. The regularity of the embedded surfaces was also examined, which is necessary for a diffeomorphic mapping across surfaces, and this was also found to be comparable with that provided by spherical demons. A key advantage of spectral matching over spherical demons is the ability to do group-wise atlasing without bias to a reference. [Lombaert et al. \(2013b\)](#) have demonstrated that their spectral method has minimal bias when compared to spherical demons and the group-wise extension presented in this chapter further removes bias.

A spatio-temporal atlas is an extremely useful tool for analysing cortical shape and, in the next chapter, the potential of this methodology for comparing cortical anatomy will be explored. For this, surface templates will be constructed for both fetuses and preterm neonates, and the difference in mean shapes will be quantified by computing displacements for matched locations.

A COMPARISON OF CORTICAL SHAPE FOR FETUSES AND PRETERMS

5.1 INTRODUCTION

At the time of writing, very little work has been done to investigate the differences in cortical geometry between fetuses and preterms at equivalent ages. A first attempt was made by [Clouchoux et al. \(2012\)](#) who measured a reduction in cortical plate surface area for preterms. More recently, [Lefèvre et al. \(2015\)](#) quantified cortical folding for fetuses ($N = 14$) and also preterms ($N = 27$, imaged shortly after birth) and found differences for a number of metrics, including global folding measures, lateral surface area and gyrification index (age range overlap: 25-35 weeks). This suggests that the process of birth itself may have some effect on the shape of the brain or that folding characteristics may diverge in preterms due to their exposure to an *ex utero* environment after birth.

Meaningful comparison of brain development between fetuses and neonates is a difficult task. Inescapable differences in image acquisition such as native voxel resolution, signal-to-noise ratio and the additional post-processing image reconstruction step for addressing movement in fetal data, may systematically influence image segmentation, mesh extraction and ultimately any quantification of shape or gyrification between the two groups. Therefore, great care must be taken when drawing conclusions from any quantitative comparison.

The fetal images used in this thesis were created by first acquiring 2D slices with a resolution of $1.25 \times 1.25 \times 2.5$ mm, (disoriented due to fetal motion) which were then reconstructed into a coherent 3D volume with an isotropic resolution of $1.18 \times 1.18 \times 1.18$ mm, using SVR registration method (Jiang et al., 2007). In comparison, the neonatal image datasets have a resolution of $0.86 \times 0.86 \times 1$ mm. Due to the relatively thick slices of the fetal acquisition and the non-super resolution reconstruction, extracted fetal surfaces are smoother than comparable neonatal surfaces. Thus a comparison of fine sulcal geometry using mean curvature or sulcal depth is confounded. This chapter therefore focusses on quantifying global shape differences between the two cohorts.

Note the aim of this thesis is develop methods for quantifying folding and when higher quality data become available (for example, through the [Developing Human Connectome Project](#)), the methodology developed in this chapter and the previous chapter may be readily applied to a quantification of folding.

5.1.0.1 Overview

In this chapter, surface atlas templates are constructed for both fetuses and neonates for the postmenstrual age range 28 - 36 weeks and a comparison of mean cortical shapes across this age range is presented. Note, for the purpose of this thesis, the term postmenstrual age is defined as gestational age plus chronological age, thus postmenstrual age and gestational age are equivalent for a fetus. The cortical shape of the preterm brain is likely to depend not only on PMA at scan but also GA at birth, i.e. prematurity. Thus, preterm surface templates were constructed for groups of subjects with a similar PMA at scan and also GA at birth where sufficient imaging data were available. This allowed changes in cortical shape due to birth and also changes that occur due to a longer exposure to an ex

utero environment to be investigated. Differences in shape between preterm templates and corresponding fetal templates (constructed for the same target [PMA](#)) are quantified through displacements measured between matched locations over the surfaces. In the next section, full details of atlas construction are given including a weighting scheme that explicitly tackles unevenness in the distribution of ages at which imaging data are available. In the following section, preliminary results from this analysis are presented.

5.2 ATLAS CONSTRUCTION

Image segmentation, mesh extraction and atlas template construction were performed using the same protocols for both fetuses and preterms (as outlined Chapter 4). Left and right cortical surface meshes were extracted for 78 fetal subjects and 196 neonatal subjects, with both hemispheres contributing to the construction of symmetric templates of mean cortical shape. Nine fetal and 27 preterm templates were constructed in total for the postmenstrual age range 28-36 weeks (the overlap in age range for the two groups). For fetuses, a template was constructed for each week of gestation, with a temporal window width of 3 weeks. For preterms, templates were constructed for groups of subjects with both a similar [GA](#) at birth and [PMA](#). For this, circular windows (diameter = 3 weeks) were defined in the 2D age space (Fig. 45) with centroids at integer values of [PMA](#) and [GA](#) at birth, and atlases were constructed if enough subjects were available. The sample size for each template constructed is detailed in Table 9.

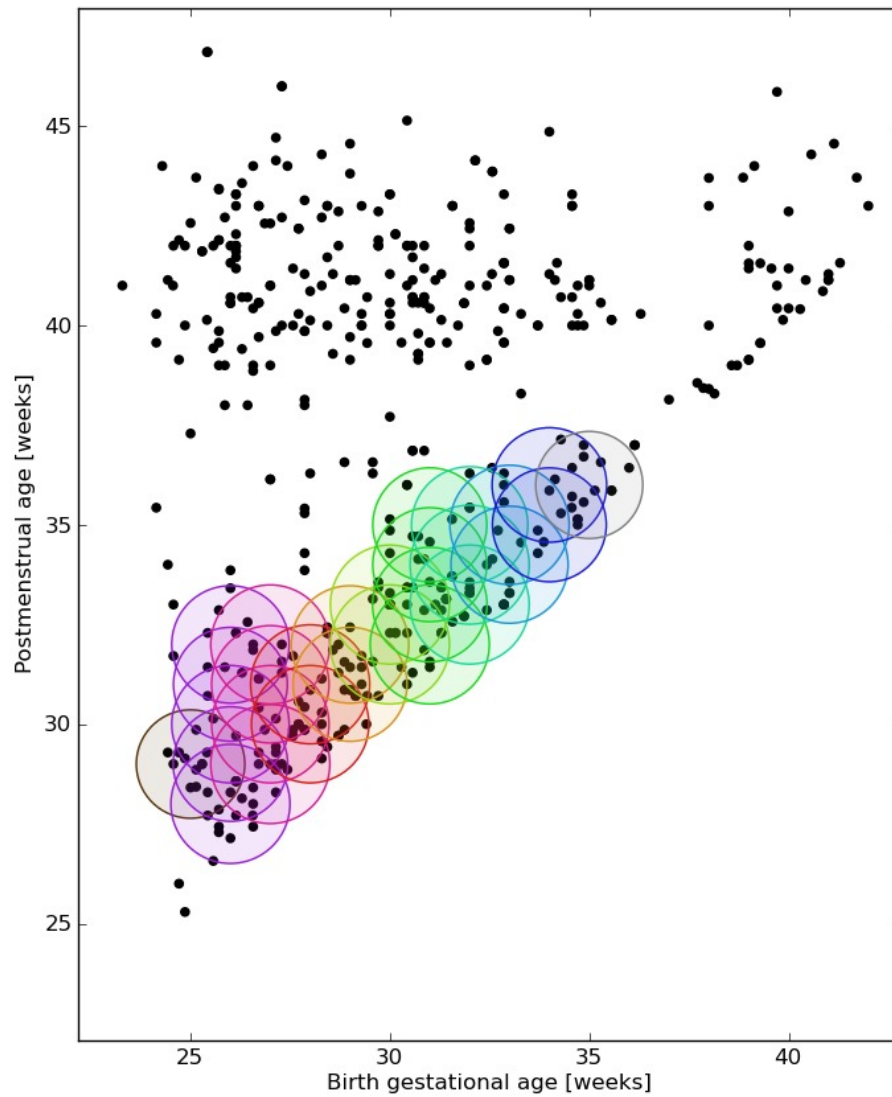


Figure 45: **Preterm atlas construction using temporal windows.** Each circle depicts a window for atlas construction, where the surfaces for subjects contained within are averaged to generate a mean template. Colours correspond to developmental trajectories for babies born at different gestational ages. For example, purple correspondences to babies born at 26 weeks GA ± 1.5 .

| GA at birth [weeks] | Sample size | | | | | | | | | |
|---------------------|-------------|----|----|----|----|----|----|----|----|--|
| fetal (N/A) | 17 | 17 | 15 | 20 | 13 | 14 | 8 | 14 | 13 | |
| 25 | – | 20 | – | – | – | – | – | – | – | |
| 26 | 28 | 35 | 25 | 19 | 18 | – | – | – | – | |
| 27 | – | 31 | 30 | 24 | 17 | – | – | – | – | |
| 28 | – | – | 37 | 42 | – | – | – | – | – | |
| 29 | – | – | – | 32 | 34 | – | – | – | – | |
| 30 | – | – | – | – | 36 | 35 | – | – | – | |
| 31 | – | – | – | – | – | 42 | 36 | 16 | – | |
| 32 | – | – | – | – | – | 31 | 36 | 17 | – | |
| 33 | – | – | – | – | – | – | 21 | 17 | – | |
| 34 | – | – | – | – | – | – | – | 16 | 20 | |
| 35 | – | – | – | – | – | – | – | – | 18 | |
| PMA [weeks] | 28 | 29 | 30 | 31 | 32 | 33 | 34 | 35 | 36 | |

Table 9: **Sample sizes for atlas construction.** The number of subjects that contributed to the construction of each template is shown. Note that symmetric templates were constructed from both the left and right hemisphere of each brain, thus the number of surfaces averaged to form a template was double the sample size above.

5.2.1 Accurate mean surface estimation from irregularly distributed samples

5.2.1.1 Fetal subjects

Previous kernel-based methods for constructing atlas templates (Kuklisova-Murgasova et al., 2011; Serag et al., 2012a) for a target age have not addressed the issue of unevenness in the distribution of ages for subjects within a temporal window. Note the method of Serag et al. is aimed to adapt to low or high sample density within a window as a whole and does not address the potentially asymmetric distribution of samples. To illustrate this issue, if we have twice as many subjects that are younger than the target age, the constructed template would have lower cerebral volume and the cortical surface would be less convoluted than expected for the target age. This can lead to inconsistent increases in brain size and

convolution between templates that are constructed for regular time-intervals. Removing this bias is important for an atlas-based comparison of cortical shape. In order to address this, a scheme for initialising a set of weights and subsequently adjusting them (reweighting) is described below.

Let t_i denote the GA of the i -th subject and t_{target} denote the target age, i.e. the age for which we wish to generate an average anatomy for. Assuming that the average surface position is linearly related to time within a window, then an ideal set of weights to reconstruct the average surface must satisfy:

$$t_{target} = \sum_i w_i t_i,$$

i.e. the target age is the a weighted mean (or centroid) of the ages. If weights are first initialised using a Gaussian distribution (Serag et al., 2012a), they can then be reweighted to satisfy this requirement by scaling the weights for all subjects younger (or older) than the target age. Let t'_i denote the difference in age of subject i from the target age, i.e. $t'_i = t_i - t_{target}$, then the weighting for the younger subjects may be recomputed by multiplying them by following scale factor:

$$\frac{\sum_{t'_i > 0} w_i |t'_i|}{\sum_{t'_i < 0} w_i |t'_i|}. \quad (15)$$

Alternatively, the inverse scale factor may be used to reweight the older subjects. Note the weights must now be renormalised so that they sum to one. Unfortunately, if the data within a window is highly non-uniformly distributed, this simple approach can lead to very high weights for a few individual subjects. For example, if only only a few subjects are older than the target with a large number of younger subjects, then this reweighting would introduce a strong bias towards the individual anatomy of the few older subjects. To mitigate this problem, a limit

was placed on the maximum weight attributed to any single subject. A reweighting factor was computed iteratively starting at 1 and incrementing up to the optimum value determined by Eqn. 15, while all weights remain below the maximum allowed value (a value 0.1 was chosen for this purpose).

5.2.1.2 Preterm subjects

For preterm subjects, templates were constructed for circular windows in the 2D age space (Fig 45). Again, if subjects are not well distributed within a window, the mean shape will not be representative for a target PMA and GA at birth. Let a point in the 2D age-domain be denoted as a vector $\mathbf{t} = [t_1, t_2]$, representing the GA at birth and PMA respectively, and let a target point be denoted \mathbf{t}_{target} . An ideal weighting is given when the weighted mean of the subject age vectors is equal to the target age vector:

$$\mathbf{t}_{target} = \sum_i w_i \mathbf{t}_i. \quad (16)$$

In some cases, a set of positive weights that satisfy this equation may not exist, for example, if all points are located within a sector of the window with an arc length of less than 180 degrees. In this case, it is not possible to construct the average shape for a target age vector at the centre of the window.

To automatically determine if a window was suitable for constructing an average template, the angular density was estimated along numerous directions using a Gaussian kernel (Rosenblatt, 1956; Parzen, 1962). If all the densities computed were greater than a certain threshold, then an atlas was constructed for the window in question. For subject i with age vector \mathbf{t}_i inside a window centred at \mathbf{t}_{target} , with

Euclidean displacement from the centre $\mathbf{t}'_i = \mathbf{t}_i - \mathbf{t}_{target}$, its angular displacement θ_i is given:

$$\theta_i = \text{angle}(\mathbf{t}'_i),$$

$$\text{angle}(\mathbf{t}) = \arctan\left(\frac{|t_1|}{|t_2|}\right).$$

Then for a group of k subjects with angular displacements $\Theta = [\theta_1, \theta_2, \dots, \theta_k]$, the density may be estimated by a Gaussian kernel G for a particular direction ϕ :

$$\text{density}(\Theta, \phi) = \sum_i G(\text{wrap}(\theta_i - \phi)),$$

$$G(\theta) = \exp\left(\frac{-\theta^2}{2\sigma^2}\right), \quad \text{wrap}(\theta) = \begin{cases} \theta - 2\pi & \text{if } \theta > \pi \\ \theta + 2\pi & \text{if } \theta < -\pi \\ \theta & \text{else} \end{cases}.$$

Here G is a Gaussian weighting function and “wrap” is a function that ensures the angular displacement $\theta_i - \phi$ is given by the shortest route around the circle. Densities were computed for several uniformly spaced directions $\Phi = [\phi_1, \phi_2, \dots, \phi_l]$, $\phi_i = \frac{2\pi i}{l}$. If the estimated density was greater than a threshold τ for all directions ϕ_i , then an atlas was constructed for \mathbf{t}_{target} . Parameter values $\sigma = \pi\sqrt{2 \times \ln(2)}$ (i.e. FWHM = $\frac{\pi}{2}$) and $\tau = 1.5$ were chosen through experimentation, by observing the windows selected in the 2D age domain.

After selecting windows for atlas construction, subject weights were again initialised with a Gaussian weighting (Serag et al., 2012a) and iteratively reweighted to gain a better estimate of the average anatomy for a target PMA/GA. When averaging a group of cortical surfaces, each subject exerts a “pull” away from the desired

average anatomy towards an anatomy more similar to itself. The exact relationship between the average surface position and both PMA and GA is unknown. However, if we assume locally linear relationships within a window, this “pull” is given by the product of its weight and its displacement from the centre of the window, and balancing this for all subjects is equivalent to satisfying Eqn. 16. Therefore, to gain a better estimate of the average surface for a target GA/PMA, each subject’s weight was normalised by the inverse of the density of the “pull” for its particular angular displacement. This procedure was carried out iteratively while the subjects’ weights did not exceed a threshold (see Alg. 4 for a technical description). Using this reweighting approach, the average estimation error $\epsilon = \|\mathbf{t}_{target} - \sum w_i \mathbf{t}_i\|_2$ was reduced by a factor of approximately 10 from 0.212 ± 0.097 weeks to 0.028 ± 0.030 weeks.

Algorithm 4: Subject reweighting scheme. The intuition here is that all subjects exert a “pull” away from the desired average anatomy. The magnitude of the “pull” is given by the product of its weight and its distance away from the target age in the 2D age domain. Reweighting is achieved by normalizing this “pull” for all angular directions. An assumption here is that there are no large angular displacements between subjects within the window.

Data:

subject age vectors: $\mathbf{t}_1, \dots, \mathbf{t}_k$
 target age vector: \mathbf{t}_{target}
 initial weighting: $\mathbf{w} = [w_1, \dots, w_k]$
 angular displacements: $\Theta = [\theta_1, \dots, \theta_k]$

Initialise Parameters:

$\epsilon_{best} = \infty, \quad \epsilon_{min} = 0.01, \quad w_{max} = 0.1, \quad \text{max_iterations} = 10$

for $i = 1$ **to** max_iterations **do**

```

// compute centriod error  $\epsilon$ 
 $\epsilon = \|\mathbf{t}_{target} - \sum w_i \mathbf{t}_i\|_2$ 

// update best weights if error is lower than previous iterations
if  $\epsilon < \epsilon_{best}$  &  $\forall w_i. w_i < w_{max}$  then
     $\epsilon_{best} = \epsilon$ 
     $\mathbf{w}_{best} = \mathbf{w}$ 
    // exit if error is acceptable
    if  $\epsilon < \epsilon_{min}$  then
        break

// copy old weights
for  $i = 1$  to  $k$  do
     $v_i = w_i$ 

// reweight by the inverse weighted density
for  $i = 1$  to  $k$  do
     $\text{density}_i = \sum_{j=1}^k \|\mathbf{t}_i - \mathbf{t}_{target}\|_2 v_j G(\text{wrap}(\theta_j - \theta_i))$ 
     $w_i = v_i / \text{density}_i$ 

// normalise weights
 $w_{total} = \sum_i w_i$ 
 $\mathbf{w} = \mathbf{w} / w_{total}$ 

```

return \mathbf{w}_{best}

5.3 RESULTS

Fig. 47 shows average cortical templates constructed for extremely preterm subjects (25-29 weeks), while Fig. 48 shows templates for very preterm (30-32 weeks) and moderately preterm subjects (33-35 weeks). Fetal templates with equivalent postmenstrual ages are also shown for comparison. Note the colour mapping on these templates is to aid visual perception of the surface geometry; convex and concave regions are shown as blue and red respectively. More views are presented in Appendix A.1. Volumes were also estimated for each template using the divergence theorem algorithm (Alyassin et al., 1994) and can be seen in Fig. 49.

The volumes for preterm templates were generally lower compared to fetal templates with the same PMA, with a few exceptions at later PMAs. Furthermore, for extremely preterm babies born around 26-27 weeks GA (where enough data was available to construct several templates along a developmental trajectory), a non-linear increase in volume was observed compared to a linear increase for fetal templates. Initially, a slow increase in volume was observed followed by a fast increase in volume. This suggests that an adjustment period may occur after birth as the brain adapts to its new environment, where growth is adversely affected, and after this point development may accelerate to compensate.

For extremely preterm babies, the most striking difference in shape between the surface templates, which can be observed visually, is the shift of the occipital lobe anteriorly (Fig. 47). For fetal templates, the extremity of the occipital lobe is noticeably more angular and posterior. Additionally, the depth of the inferior and superior frontal sulci as well as the superior temporal sulcus appeared reduced. For preterm babies born later (30-36 weeks), visible differences were not as appar-

ent (Fig. 48). In the next section, these differences in shape will be quantified by computing displacements for matched surface locations.

5.3.1 *Surface displacement*

In order to quantify differences in cortical shape, for each preterm template, displacements were computed for matched locations on the corresponding fetal template that was constructed for the same target PMA. For this, the preterm template mesh was first rigidly aligned with its corresponding fetal template mesh using ICP registration. The pair of surfaces were then embedded in the spectral domain via a joint spectral decomposition and corresponding fetal template spatial coordinates were estimated for each preterm template vertex using kernel regression in the spectral domain (see Chapter 4). For each preterm template, the scalar projection of the displacement along the outward surface normal N was computed by taking the dot product. This allows displacements to be characterised as positive (outward) and negative (inward) relative to the surface, and can be visualised as a colour mapping over the surface. For the complete set of surfaces with the scalar projection of the displacement colour mapped see Appendix A.2.

A common pattern of displacements was observed for templates constructed for extremely preterm babies relative to templates constructed for fetal subjects (see Fig. 50 & 51). The occipital lobe and temporal lobes were shifted anteriorly relative to the parietal lobe, which was shifted superiorly. Additionally, the superior edge of the anterior temporal lobe appeared to be displaced superiorly. Negative displacements were also seen around the middle temporal gyrus and the middle frontal gyrus. For the same premenstrual age, this pattern of displacements was intensified for babies born more prematurely (see Appendix A.2).

For templates constructed for moderate and very preterm babies, again, negative displacements were seen around the middle temporal gyrus and the superior edge of the anterior temporal lobe appeared to be displaced superiorly. However, a different pattern of displacements was observed elsewhere. Positive displacements were seen at the anterior and posterior extremities of the brain, resulting in an elongated head shape. Once again, for the same premenstrual age, the pattern of displacements was more pronounced in babies born more prematurely.

5.4 DISCUSSION

For the fetal images used in this chapter, the relatively thick slices acquired ($1.25 \times 1.25 \times 2.5$ mm) and non-super resolution SVR reconstruction (Jiang et al., 2007) results in significant blurring of fine scale structures. As a result, extracted cortical surface models are noticeably smoother at later gestations and segmentation errors are common in highly convoluted regions such as the calcarine and parieto-occipital fissures. This may explain the relatively large negative displacement in these regions for templates constructed for 33-36 weeks PMA (Fig 63, Appendix A.2). For an example of typical segmentation errors, please see Figure 46. Analysis in this chapter has, therefore, been restricted to macroscopic shape differences rather than sulcal geometry. However, with better quality data, this methodology is readily applied to an analysis of sulcal geometry.

Fetal imaging is a relatively new area of research and image quality is still being improved, for example, through advances in acquisition such as 3T MRI (Victoria et al., 2016). The increased magnetic field strength over 1.5T may be used to decrease noise, increase resolution or speed up acquisition to reduce motion blurring

artefacts. Acquiring images at 3T is a challenging task, however, due to increased inhomogeneity in the magnetic field.

Post-processing techniques such as super-resolution reconstruction ([Rousseau et al., 2010](#); [Kuklisova-Murgasova et al., 2012](#); [Kainz et al., 2015](#)) have also been used to enhance the resolution of images. These techniques exploit overlapping slices to infer boundary locations with sub-voxel accuracy by observing the graded intensity response of adjacent voxels, similarly to how the human eye can discern thousands of hues from just three different cone types. Instead of simply interpolating 2D slice voxel intensities, an image may be reconstructed by estimating an unknown high resolution image volume (e.g. 0.75 mm^3) whose voxel intensities best explain the acquired 2D slices, after applying a Point Spread Function (PSF) (the impulse response of the imaging system to a point source). For an SSFSE imaging sequence, a 3D Gaussian kernel is a good approximation of the PSF. In practice, the unknown 3D volume is estimated by minimising the SSD between acquired slice voxel intensities and the estimated intensities after the 3D Gaussian kernel is applied.

A large dataset of high quality fetal brain images is currently being acquired on the [Developing Human Connectome Project](#) (which makes use of both 3T acquisition and super-resolution reconstruction). Surface models extracted from these data are greatly improved compared to those used in this chapter, with fewer segmentation errors, deeper sulci and better reproduction of fine scale detail (Fig. 46). When enough datasets become available, it will be possible to investigate the findings in this chapter further and also compare sulcal geometry between groups.

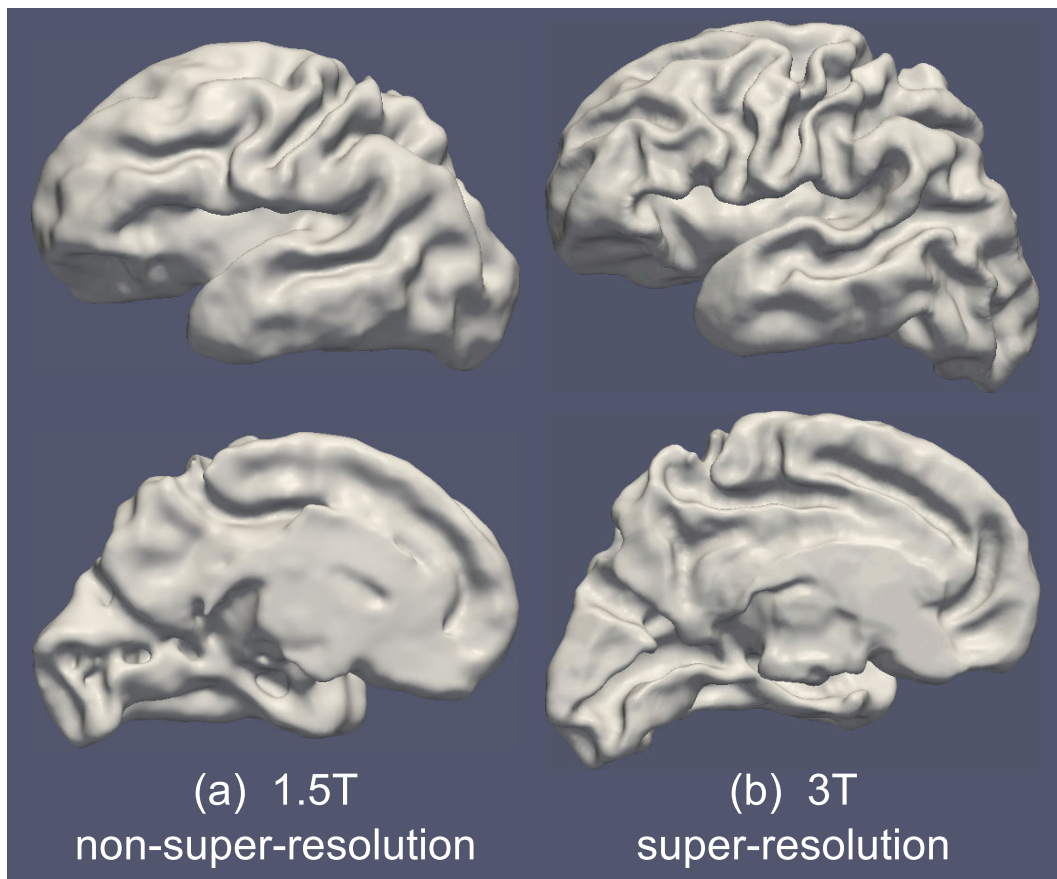


Figure 46: **Influence of imaging quality on extracted cortical surface models.** Typical cortical surface models automatically extracted for two subjects aged 32 weeks GA, from a 1.5T non-super resolution image (a) and a 3T super-resolution image (b). Note the reduced sulcal depth in (a) and also the implausible geometry around the calcarine fissure due to segmentation error.

5.5 CONCLUSION

In this chapter, an atlas-based comparison of mean shape between preterms and fetuses was undertaken. Nine fetal and 27 preterm templates were constructed in total over the postmenstrual age range 28-36 weeks. The problem of irregularly distributed subjects within a temporal window was highlighted, and a weighting scheme was proposed which contrasts with that of (Serag et al., 2012a) and (Kuklisova-Murgasova et al., 2011). Surface displacements were computed between preterm templates and corresponding fetal templates for matched locations over the surfaces, allowing an analysis of shape between the two cohorts. Patterns of displacement were uncovered that were intensified in more premature babies for the same PMA. Due to the low resolution of the reconstructed fetal images, a comparison of fine scale cortical structure, such as folding, was prohibited. However, when high quality data becomes available in the future (e.g. through the [Developing Human Connectome Project](#)), the methodology developed in this chapter will enable such a comparison.

There is a lot of scope for further analysis with this methodology. For example, geometric features such as mean curvature may be compared across a surface template to discover differences in sulcation between the two groups. Moreover, changes in curvature over time can be analysed in order to determine the rate of development along different developmental trajectories. This will be discussed in more detail in the next chapter, where the contributions of this thesis are reviewed and future work is explored.

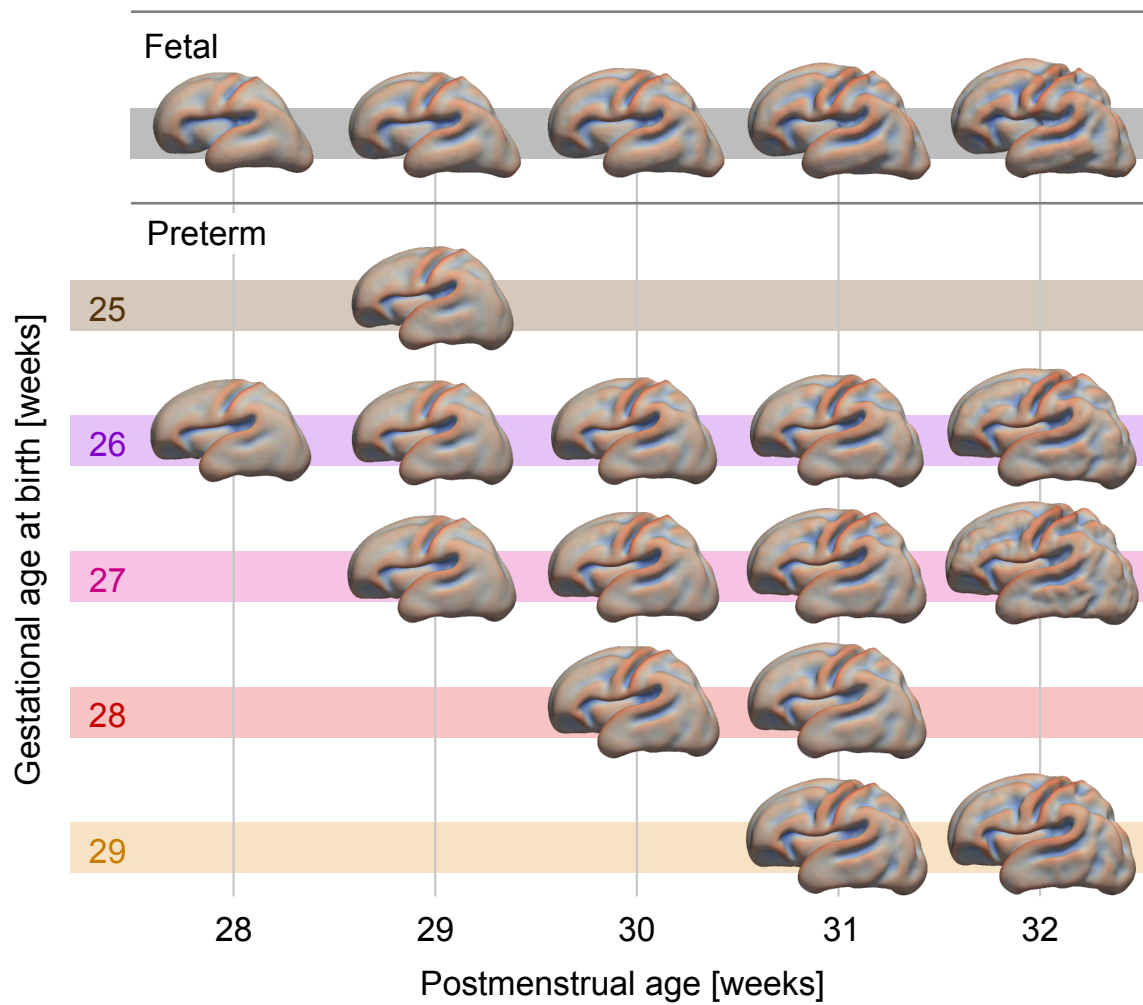


Figure 47: **Atlas templates (GA at birth: 25-29 weeks).** For the preterm templates, each row shows a developmental trajectory for subjects born around the same age. Note the background colour for each birth age corresponds with Fig. 45 & 49. The colour mapping shows mean curvature (red, convex; blue, concave), which enhances visual perception of the template geometries.

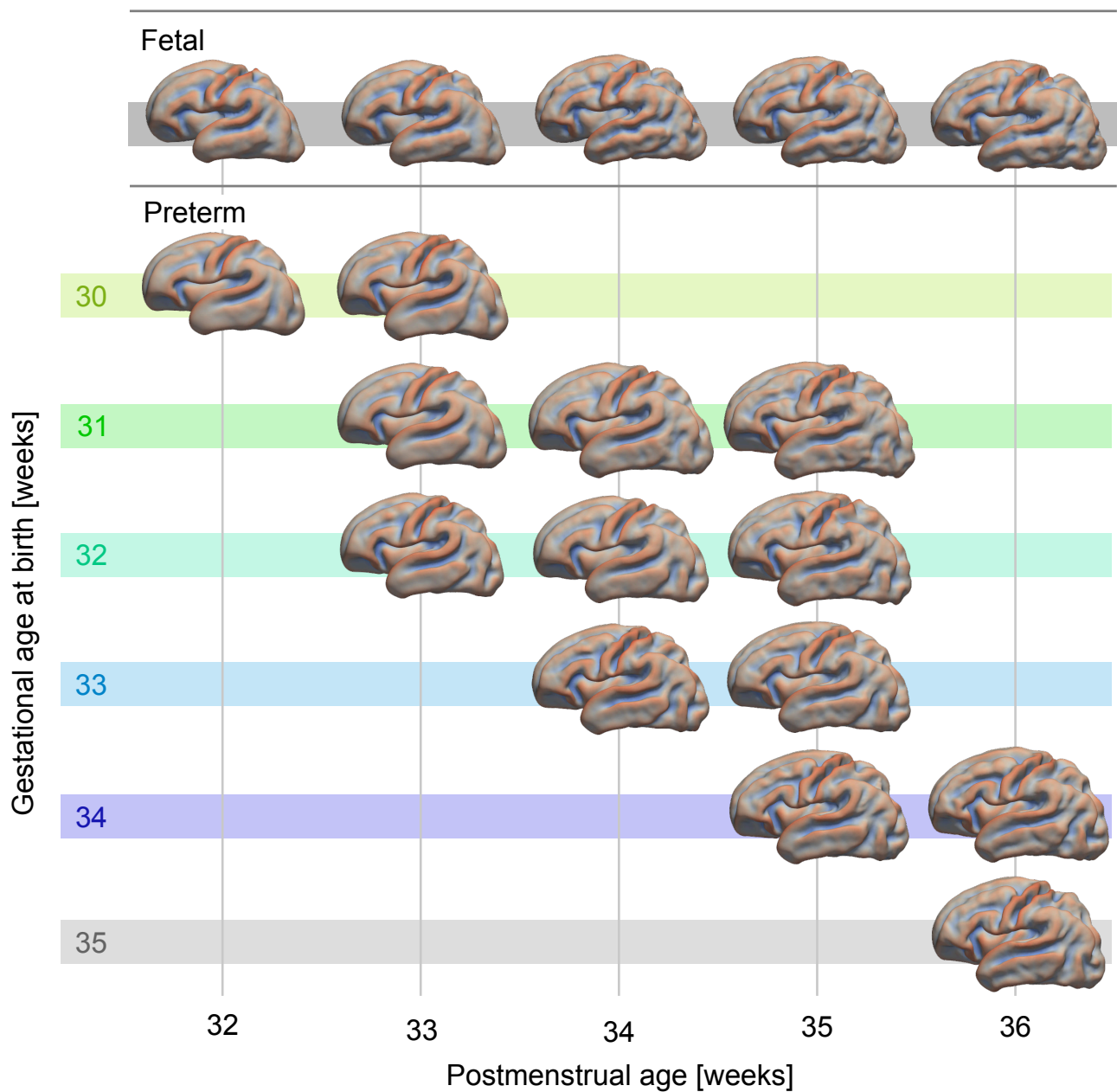


Figure 48: **Atlas templates (GA at birth: 30-36 weeks)** For the preterm templates, each row shows a developmental trajectory for subjects born around the same age. Note the background colour for each birth age corresponds with Fig. 45 & 49. The colour mapping shows mean curvature (red, convex; blue, concave), which enhances visual perception of the template geometries.

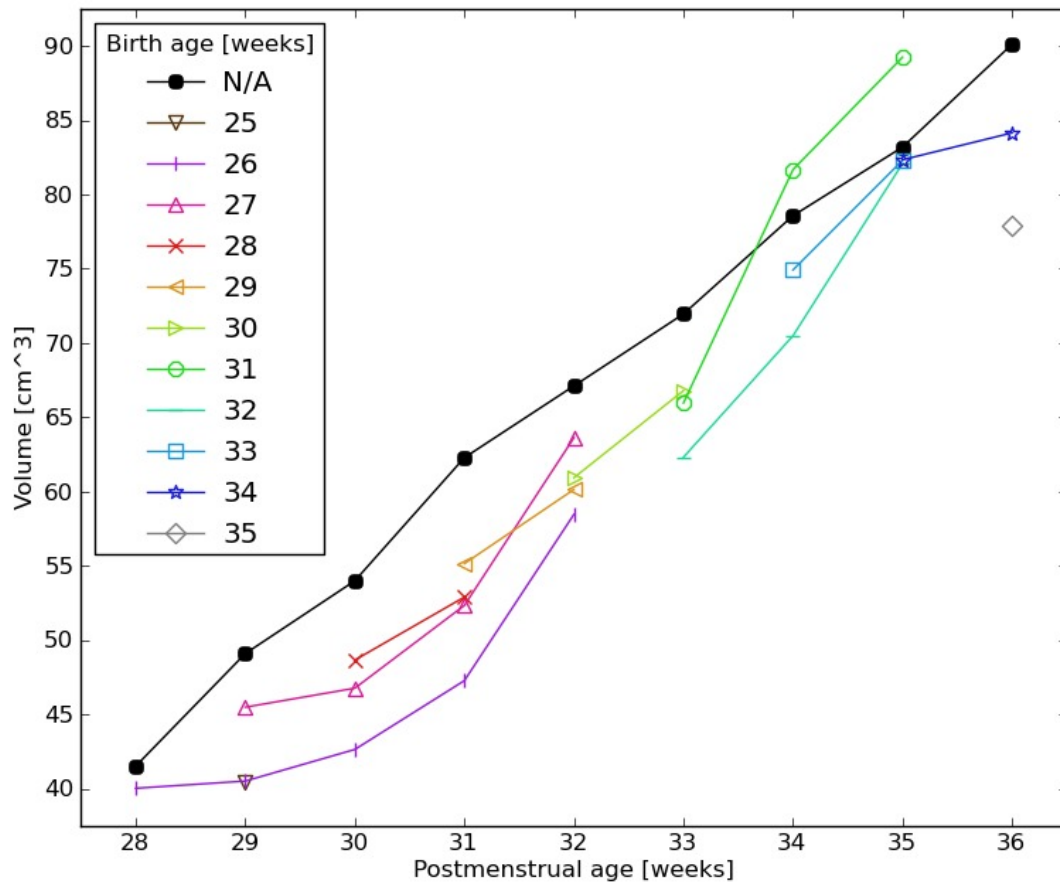


Figure 49: **Template volumes.** Each colour depicts a developmental trajectories for subjects born at a particular age. These colours correspond to those in Fig. 45, 47 & 48.

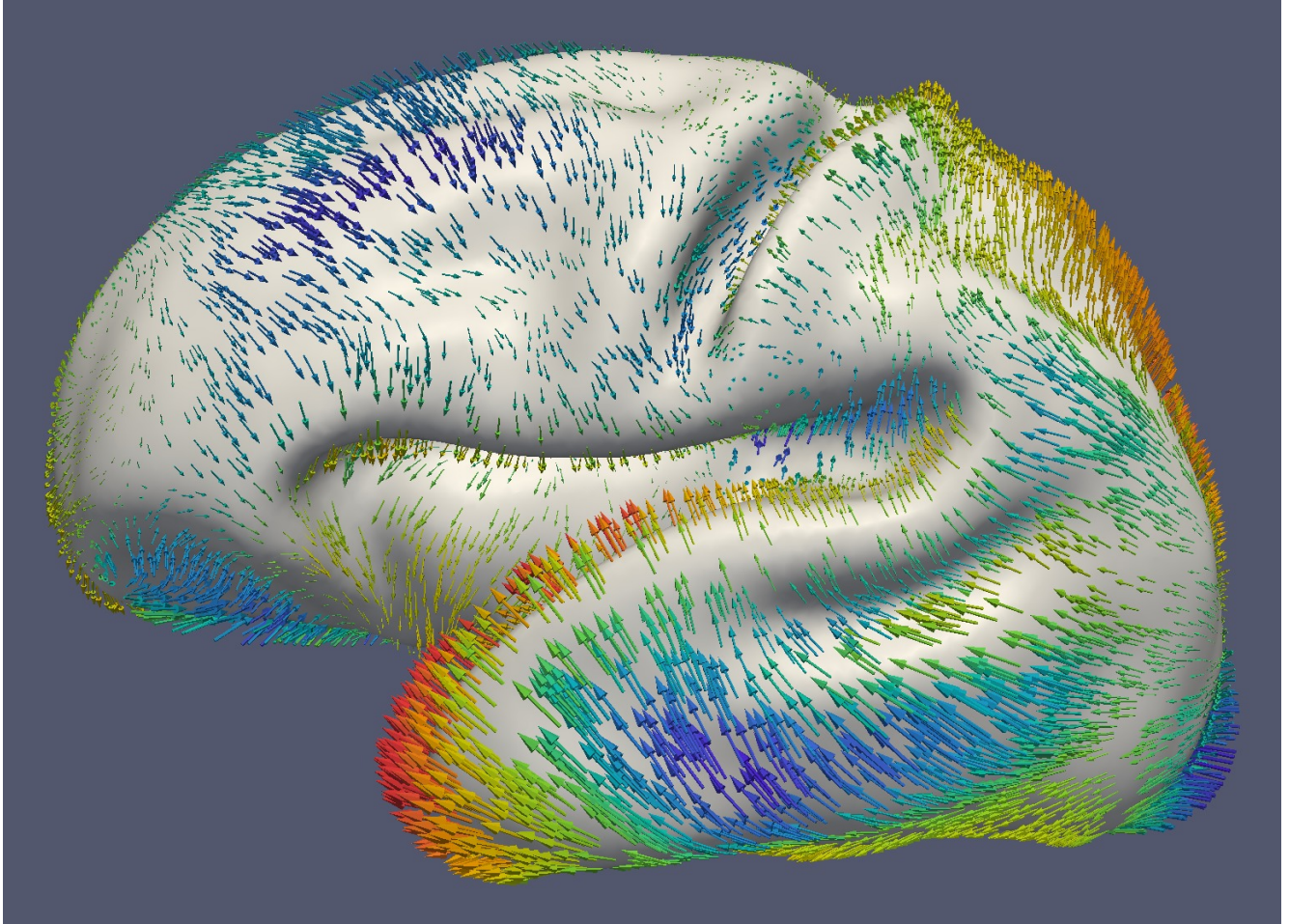


Figure 50: **Typical template displacement for extremely preterm babies (lateral view).**

The template shown was constructed for babies born at 27 ± 1.5 weeks GA and scanned at 29 ± 1.5 weeks PMA. Arrows depict displacements from matched locations on the corresponding fetal template constructed for subjects with a GA of 29 ± 1.5 weeks. The length of the arrows show the magnitude of the displacement with a one-to-one scaling. Colour mapping depicts the scalar projection of the displacement along the surface normal (red: +3 mm, blue: -3 mm). Warmer colours show positive (outward) displacements and cooler colours show negative (inward) displacements.

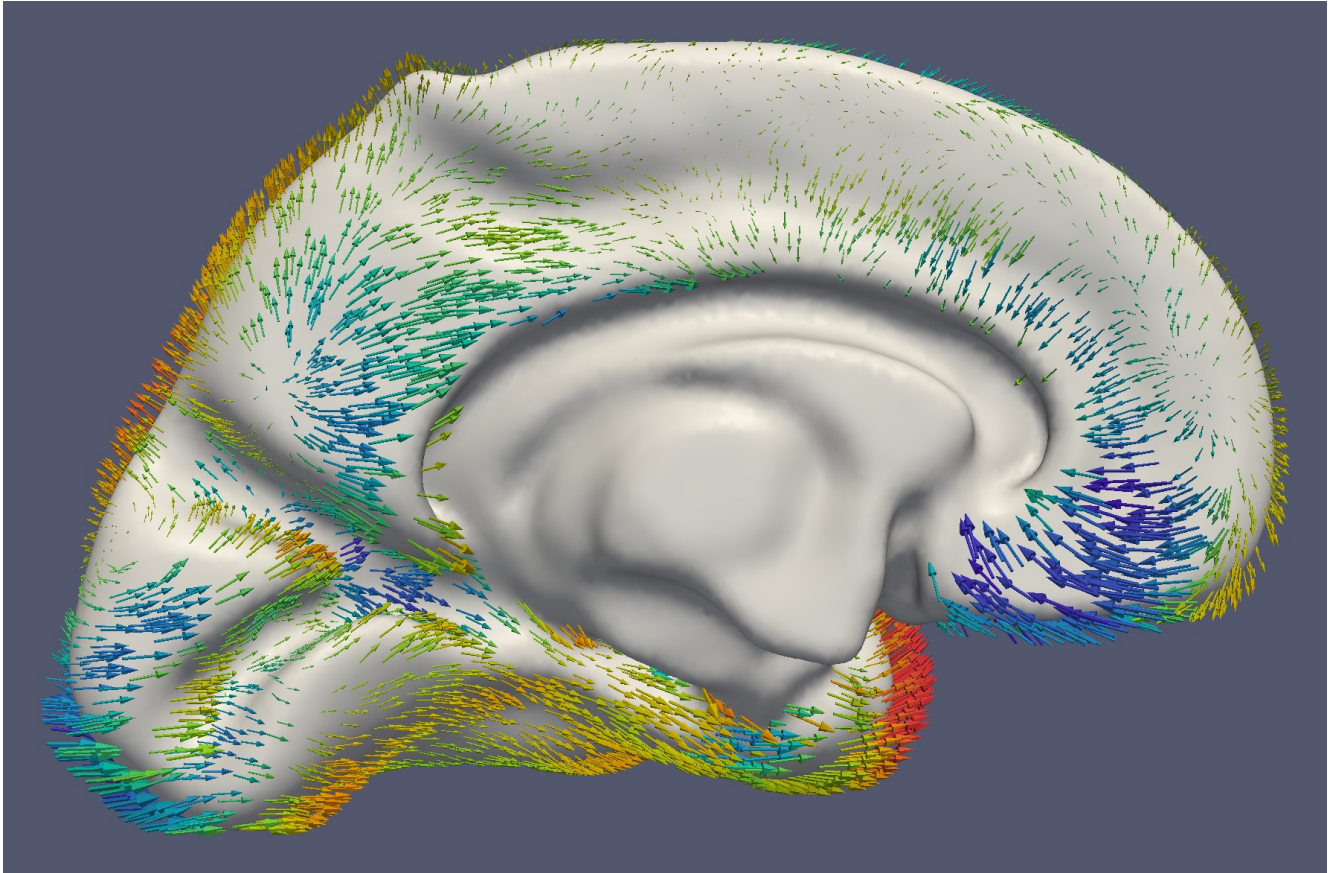


Figure 51: **Typical template displacement for extremely preterm babies (medial view).** The template shown was constructed for babies born at 27 ± 1.5 weeks GA and scanned at 29 ± 1.5 weeks PMA. Arrows depict displacements from matched locations on the corresponding fetal template constructed for subjects with a GA of 29 ± 1.5 weeks. The length of the arrows show the magnitude of the displacement with a one-to-one scaling. Colour mapping depicts the scalar projection of the displacement along the surface normal (red: +3 mm, blue: -3 mm). Warmer colours show positive (outward) displacements and cooler colours show negative (inward) displacements.

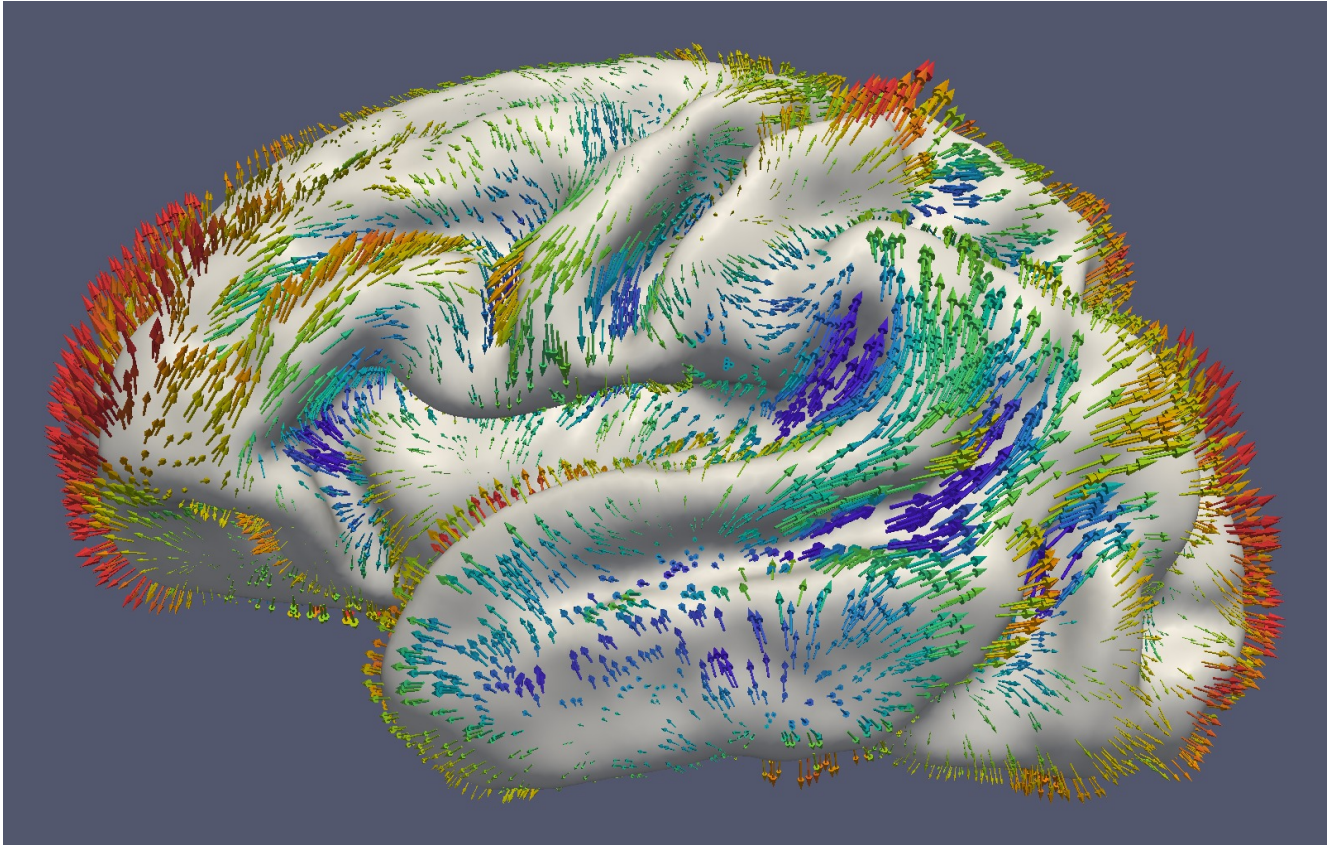


Figure 52: **Typical template displacement for moderately preterm babies (lateral view).** The template shown was constructed for babies born at 33 ± 1.5 weeks GA and scanned at 35 ± 1.5 weeks PMA. Arrows depict displacements from matched locations on the corresponding fetal template constructed for subjects with a GA of 35 ± 1.5 weeks. The length of the arrows show the magnitude of the displacement with a one-to-one scaling. Colour mapping depicts the scalar projection of the displacement along the surface normal (red: +3 mm, blue: -3 mm). Warmer colours show positive (outward) displacements and cooler colours show negative (inward) displacements.

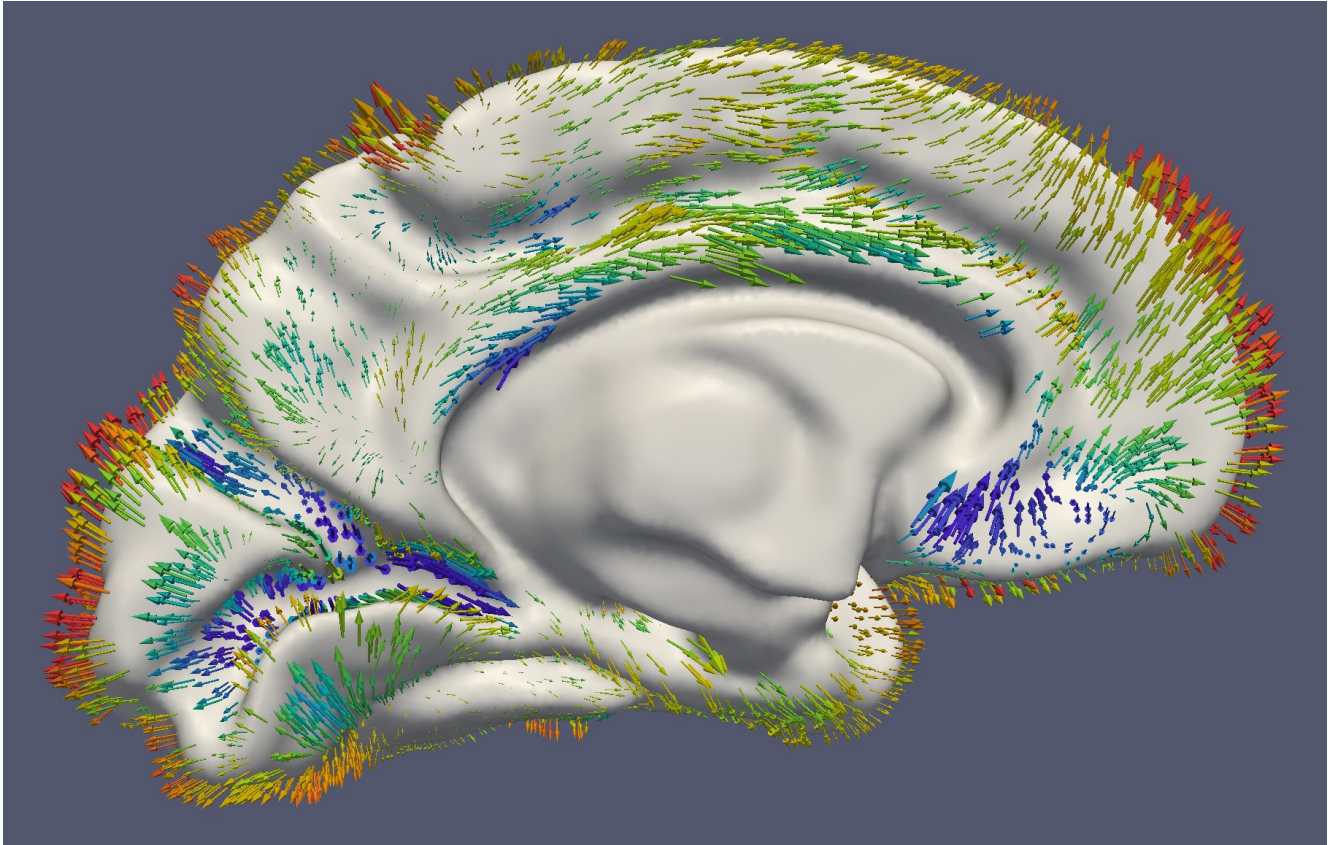


Figure 53: **Typical template displacement for moderately preterm babies (medial view).** The template shown was constructed for babies born at 33 ± 1.5 weeks GA and scanned at 35 ± 1.5 weeks PMA. Arrows depict displacements from matched locations on the corresponding fetal template constructed for subjects with a GA of 35 ± 1.5 weeks. The length of the arrows show the magnitude of the displacement with a one-to-one scaling. Colour mapping depicts the scalar projection of the displacement along the surface normal (red: +3 mm, blue: -3 mm). Warmer colours show positive (outward) displacements and cooler colours show negative (inward) displacements.

CONCLUSION

This thesis has presented work on quantifying cortical folding patterns for fetuses and preterm neonates. Techniques have been developed for quantifying global and regional gyrification, and also for constructing a spatio-temporal surface atlas of cortical morphology. These techniques have been used to investigate and compare folding patterns for both fetuses and preterms. In the following section, the contributions of this thesis will be reviewed for each chapter and then, in Section 6.2, avenues for future work will be explored.

6.1 CONTRIBUTIONS

In Chapter 3, an image analysis framework for quantifying gyrification was presented. Eight curvature-based folding measures from the literature were adapted and applied to a normal fetal dataset ($N = 80$) over a large gestational age range (21.7 – 38.9 weeks), extending our understanding of how the cortex folds through this critical developmental period. The relationship between folding measures and GA reflected the rate of gyrification that has been observed visually with sulcation accelerating rapidly over this time-period, suggesting that these measures provide a good characterisation of gyrification. Moreover, folding measures were found to be better correlated with age than cerebral volume, suggesting they may be a

better biomarker for assessing the neural health of fetuses. A Gompertz function was found to be a good fit for this relationship, which subsequently allowed an accurate prediction of physiological age. The concept of physiological age allows a developmental delay to be characterised by an offset in weeks, which may be of more meaningful in a clinical setting. Although this framework is well-suited for detecting gross discrepancies in gyrification, a limitation is that localised disturbances may remain undetected when averaging curvatures over a relatively large region.

Chapter 4 proposed a framework for constructing a spatio-temporal surface atlas, which allows folding to be analysed on a localised scale. Cortical surface models were extracted for fetal subjects by deforming a mesh with spherical topology towards a segmented cortical boundary. Correspondences were then established for age-matched groups of surfaces through a joint spectral analysis. Fifteen surface templates were constructed at weekly intervals using kernels to gain a spatially and temporally weighted estimate of the average surface position. Templates constructed from disjoint samples of the available data were found to be reproducible with a small amount of deviation measured between them. This methodology performs favourably in comparison to a state of the art spherical registration algorithm, spherical demons. Comparable sulcal alignment accuracy was achieved with a reduced computational burden. Furthermore, a key advantage is the ability to perform group-wise atlasing without bias to an initial reference.

In Chapter 5, an atlas-based comparison of cortical shape for fetuses and preterms was presented. The problem of unevenly distributed subjects within a temporal window was highlighted and subject weighting schemes were proposed to compensate for this, allowing a more accurate estimation of the average surface for a particular target age. For preterms, surface templates were constructed for groups of subjects with both a similar GA at birth and PMA, allowing the effects

of premature birth to be investigated along different developmental trajectories. Mean shape was compared between each preterm template and its corresponding fetal template (constructed for the same [PMA](#)). For this, correspondences were established via a joint spectral analysis and displacements were computed for matched surface locations. Patterns of displacement were characterised for extremely preterm babies and also for very and moderately preterm babies. Moreover, these patterns appeared intensified in more premature babies for the same [PMA](#). It is possible that the differences in imaging data used for this comparison had some influence over the results, therefore more work is needed to assess the reliability of these preliminary findings. This will be possible when large, high-quality datasets become available through the ([Developing Human Connectome Project](#)).

6.2 LIMITATIONS AND FUTURE WORK

A spatio-temporal surface atlas is an extremely useful tool for comparing cortical anatomy between cohorts. In this thesis, due to the limited size of the fetal dataset used, cortical shape was compared using symmetric templates only. However, with enough data, separate hemispheric templates could be generated, thus allowing lateralisation of the cortical surface to be investigated and to determine how prematurity effects this. Another interesting pathway that was not explored using this methodology is the temporal evolution of the cortical surface. This could be accomplished by comparing surface position or curvature between temporally adjacent atlas templates. Alternatively, all subjects may be embedded in a common reference space as in [Section 4.2.7](#) and regression models constructed to model cortical evolution for various locations over a template as in ([Habas et al., 2012](#)).

A direct comparison of average surface position through surface templates (as in Chapter 5) has a weakness: the reliability of displacements measured is not quantified, i.e. there is no control over type 1 errors. In Section 4.3.6 the consistency for a fetal surface template was found to be around 0.4 mm on average for a sample size of 20 and a target GA of 36 weeks. Thus, a displacement measured from a preterm template to a matched location on a fetal template of this magnitude could easily be due to random variation. A surface atlas does, however, facilitate a more robust comparison through Surface-Based Morphometry (SBM), which is a surface-based analogue to VBM. For this, the surface locations of the individual subjects are first mapped onto an average surface template. A statistical test is then performed at each template vertex location, to determine if the difference in mean surface position is significant between the two groups. SBM is not limited to discovering significant displacements between groups and can also be applied to any vertex data accompanying the individual cortical surface models. For example, geometric features such as mean curvature or sulcal depth may be compared, as can neurological parameters that can be projected onto the surface, such as cortical thickness.

In this thesis, techniques to quantify cortical folding have been applied to normal fetuses and preterms, however, these techniques could also be applied to investigate folding in other populations with congenital anomalies. Cohorts of particular interest include fetuses and infants with mild ventricular enlargement (ventriculomegaly) and congenital heart disease. Ventriculomegaly is one of the most common congenital anomalies, with a prevalence of around 1%. In most cases, prognosis is good (Gaglioti et al., 2005), however, it is associated with a greater risk of cognitive, language, and behavioural impairments in infancy. It is thought that disturbances in cortical development may be indicative of poor outcome in these babies (Kyriakopoulou et al., 2014). Congenital Heart Disease (CHD) is an-

other common birth defect, with around 0.6% of live births having moderate or severe cases (Hoffman and Kaplan, 2002). There is increasing evidence that brain development is altered *in utero*. For example, Limperopoulos et al. (2010) found evidence of impaired neuroaxonal development and metabolism for fetuses with CHD in the third trimester and also smaller brain volumes compared to normal fetuses. Furthermore, reduced gyrification has been observed *in utero* (Clouchoux et al., 2013) and also before surgical intervention (Ortinou et al., 2013).

BIBLIOGRAPHY

- M. Ajayi-Obe, N. Saeed, F. Cowan, M. Rutherford, and A. Edwards. Reduced development of cerebral cortex in extremely preterm infants. *The Lancet*, 356(9236):1162 – 1163, 2000. doi:10.1016/S0140-6736(00)02761-6.
- P. Aljabar, R. Heckemann, A. Hammers, J. Hajnal, and D. Rueckert. Multi-atlas based segmentation of brain images: Atlas selection and its effect on accuracy. *NeuroImage*, 46(3):726 – 738, 2009. ISSN 1053-8119. doi:10.1016/j.neuroimage.2009.02.018.
- H. Alt and M. Godau. Computing the fr chet distance between two polygonal curves. *International Journal of Computational Geometry & Applications*, 11(9):75–91, 1995. doi:10.1007/BF02439643.
- A. M. Alyassin, J. L. Lancaster, J. H. Downs, and P. T. Fox. Evaluation of new algorithms for the interactive measurement of surface area and volume. *Medical Physics*, 21(6):741–752, 1994. doi:10.1118/1.597333.
- X. Artaechevarria, A. Munoz-Barrutia, and C. Ortiz-de Solorzano. Combination strategies in multi-atlas image segmentation: Application to brain mr data. *Medical Imaging, IEEE Transactions on*, 28(8):1266–1277, 2009. ISSN 0278-0062. doi:10.1109/TMI.2009.2014372.
- J. Ashburner and K. J. Friston. Voxel-based morphometry-the methods. *NeuroImage*, 11(6):805 – 821, 2000. doi:10.1006/nimg.2000.0582.

- S. P. Awate, L. Win, P. Yushkevich, R. T. Schultz, and J. C. Gee. 3d cerebral cortical morphometry in autism: Increased folding in children and adolescents in frontal, parietal, and temporal lobes. In *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2008*, volume 5241 of *Lecture Notes in Computer Science*, pages 559–567. Springer Berlin Heidelberg, 2008. ISBN 978-3-540-85987-1. doi:10.1007/978-3-540-85988-8_67.
- M. Bach Cuadra, M. Schaer, A. Andre, L. Guibaud, S. Eliez, and J.-P. Thiran. Brain tissue segmentation of fetal MR images. In *Workshop on Image Analysis for Developing Brain, in 12th International Conference on Medical Image Computing and Computer Assisted Intervention*. 2009.
- G. Ball, J. P. Boardman, D. Rueckert, P. Aljabar, T. Arichi, N. Merchant, I. S. Gousias, A. D. Edwards, and S. J. Counsell. The effect of preterm birth on thalamic and cortical development. *Cerebral Cortex*, 22(5):1016–1024, 2012. doi:10.1093/cercor/bhr176.
- P. Batchelor, A. Castellano Smith, D. Hill, D. Hawkes, T. Cox, and A. Dean. Measures of folding applied to the development of the human fetal brain. *Medical Imaging, IEEE Transactions on*, 21(8):953–965, 2002. ISSN 0278-0062. doi:10.1109/TMI.2002.803108.
- J. Boardman, C. Craven, S. Valappil, S. Counsell, L. Dyet, D. Rueckert, P. Aljabar, M. Rutherford, A. Chew, J. Allsop, F. Cowan, and A. Edwards. A common neonatal image phenotype predicts adverse neurodevelopmental outcome in children born preterm. *NeuroImage*, 52(2):409–414, 2010. doi:10.1016/j.neuroimage.2010.04.261.
- J. P. Boardman, S. J. Counsell, D. Rueckert, O. Kapellou, K. K. Bhatia, P. Aljabar, J. Hajnal, J. M. Allsop, M. A. Rutherford, and A. D. Edwards. Abnormal deep

- grey matter development following preterm birth detected using deformation-based morphometry. *NeuroImage*, 32(1):70 – 78, 2006. ISSN 1053-8119. doi:10.1016/j.neuroimage.2006.03.029.
- F. Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(6):567–585, 1989. ISSN 0162-8828. doi:10.1109/34.24792.
- BrainMaps.org. Nissl stained coronal slice of a primate cortex, accessed via [screen-shot webpage](http://brainmaps.org/treewidget.jpg). <http://brainmaps.org/treewidget.jpg>. Used under CC BY 3.0 / cropped and removed background widgets and text.
- S. Brakatsoulas, D. Pfoser, R. Salas, and C. Wenk. On map-matching vehicle tracking data. In *Proceedings of the 31st International Conference on Very Large Data Bases*, VLDB '05, pages 853–864. VLDB Endowment, 2005.
- K. Brodmann. Vergleichende lokalisationslehre der groshirnrinde. *Leipzig: Barth*, 1909.
- B. Caldalairou, N. Passat, P. Habas, C. Studholme, M. Koob, J. Dietemann, and F. Rousseau. Segmentation of the cortex in fetal MRI using a topological model. In *IEEE International Symposium on Biomedical Imaging: ISBI*, pages 2045–2048. 2011. doi:10.1109/ISBI.2011.5872814.
- M. J. Cardoso, A. Melbourne, G. S. Kendall, M. Modat, C. F. Hagmann, N. J. Robertson, N. Marlow, and S. Ourselin. Adaptive neonate brain segmentation. In *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2011*, volume 6893 of *Lecture Notes in Computer Science*, pages 378–386. Springer Berlin Heidelberg, 2011. ISBN 978-3-642-23625-9. doi:10.1007/978-3-642-23626-6_47.
- M. J. Cardoso, A. Melbourne, G. S. Kendall, M. Modat, N. J. Robertson, N. Marlow, and S. Ourselin. Adapt: An adaptive preterm segmentation algorithm

- for neonatal brain MRI. *NeuroImage*, 65:97 – 108, 2013. ISSN 1053-8119. doi: 10.1016/j.neuroimage.2012.08.009.
- C.-H. Chang, C.-H. Yu, F.-M. Chang, H.-C. Ko, and H.-Y. Chen. The assessment of normal fetal brain volume by 3-d ultrasound. *Ultrasound in Medicine & Biology*, 29(9):1267 – 1272, 2003. ISSN 0301-5629. doi:10.1016/S0301-5629(03)00989-X.
- F.-M. Chang, K.-F. Hsu, H.-C. Ko, B.-L. Yao, C.-H. Chang, C.-H. Yu, R.-I. Liang, and H.-Y. Chen. Fetal heart volume assessment by three-dimensional ultrasound. *Ultrasound in Obstetrics and Gynecology*, 9(1):42–48, 1997. ISSN 1469-0705. doi: 10.1046/j.1469-0705.1997.09010042.x.
- G. Christensen, R. Rabbitt, and M. Miller. Deformable templates using large deformation kinematics. *Image Processing, IEEE Transactions on*, 5(10):1435–1447, 1996. ISSN 1057-7149. doi:10.1109/83.536892.
- F. Chung. *Spectral Graph Theory*. AMS, 1997.
- M. Chung, K. Dalton, and R. Davidson. Tensor-based cortical surface morphometry via weighted spherical harmonic representation. *Medical Imaging, IEEE Transactions on*, 27(8):1143–1151, 2008. doi:10.1109/TMI.2008.918338.
- C. Clouchoux, A. du Plessis, M. Bouyssi-Kobar, W. Tworetzky, D. McElhinney, D. Brown, A. Gholipour, D. Kudelski, S. Warfield, R. McCarter, R. Robertson, A. Evans, J. Newburger, and C. Limperopoulos. Delayed cortical development in fetuses with complex congenital heart disease. *Cerebral Cortex*, 23(12):2932–2943, 2013. doi:10.1093/cercor/bhs281.
- C. Clouchoux, D. Kudelski, A. Gholipour, S. Warfield, S. Viseur, M. Bouyssi-Kobar, J. Mari, A. Evans, A. du Plessis, and C. Limperopoulos. Quantitative in vivo mri measurement of cortical development in the fetus. *Brain Structure and Function*, 217(1):127–139, 2012. doi:10.1007/s00429-011-0325-x.

- A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Marchal. Automated multi-modality image registration based on information theory. In *Information Processing in Medical Imaging*, pages 263–274. Kluwer Academic Publishers., 1995.
- P. Coupé, J. Manjón, V. Fonov, J. Pruessner, M. Robles, and D. Collins. Patch-based segmentation using expert priors: Application to hippocampus and ventricle segmentation. *NeuroImage*, 54(2):940–954, 2011. doi:10.1016/j.neuroimage.2010.09.018.
- A. M. Dale and M. I. Sereno. Improved Localizadon of Cortical Activity by Combining EEG and MEG with MRI Cortical Surface Reconstruction: A Linear Approach. *Journal of Cognitive Neuroscience*, 5(2):162–176, 1993. ISSN 0898-929X. doi:10.1162/jocn.1993.5.2.162.
- C. Davatzikos and R. Bryan. Using a deformable surface model to obtain a shape representation of the cortex. *Medical Imaging, IEEE Transactions on*, 15(6):785–795, 1996. ISSN 0278-0062. doi:10.1109/42.544496.
- Developing Human Connectome Project. <http://www.developingconnectome.org>.
- N. Dilmen. Ultrasound image showing measurement of head circumference and biparietal diameter for a fetus. https://commons.wikimedia.org/wiki/File:Ultrasound_Scan_ND_0103155247_1553320.png, 2001. Used under CC BY-SA 3.0 / cropped and some text removed.
- E. Dittrich, T. Riklin Raviv, G. Kasprian, R. Donner, P. C. Brugger, D. Prayer, and G. Langs. A spatio-temporal latent atlas for semi-supervised learning of fetal brain segmentations and morphological age estimation. *Medical Image Analysis*, 18(1):9–21, 2014. doi:10.1016/j.media.2013.08.004.

- J. Dubois, M. Benders, C. Borradori-Tolsa, A. Cachia, F. Lazeyras, R. Ha-Vinh Leuchter, S. V. Sizonenko, S. K. Warfield, J. F. Mangin, and P. S. Hüppi. Primary cortical folding in the human newborn: an early marker of later functional development. *Brain*, 131(8):2028–2041, 2008. doi:10.1093/brain/awn137.
- A. Elad, Y. Keller, and R. Kimmel. Texture mapping via spherical multi-dimensional scaling. In *Scale Space and PDE Methods in Computer Vision*, volume 3459 of *Lecture Notes in Computer Science*, pages 443–455. Springer Berlin Heidelberg, 2005. ISBN 978-3-540-25547-5. doi:10.1007/11408031_38.
- S. Eskildsen, P. Coupé, V. Fonov, J. Manjón, K. Leung, N. Guizard, S. Wassef, L. O. Stergaard, and D. Collins. BEaST: Brain extraction based on nonlocal segmentation technique. *NeuroImage*, 59(3):2362–2373, 2012. doi:10.1016/j.neuroimage.2011.09.012.
- D. Field. Laplacian smoothing and delaunay triangulations. *Communications in Applied Numerical Methods*, 4(6):709–712, 1988. doi:10.1002/cnm.1630040603.
- B. Fischl, A. Liu, and A. Dale. Automated manifold surgery: constructing geometrically accurate and topologically correct models of the human cerebral cortex. *Medical Imaging, IEEE Transactions on*, 20(1):70–80, 2001. ISSN 0278-0062. doi:10.1109/42.906426.
- B. Fischl, M. Sereno, R. Tootell, and A. Dale. High-resolution intersubject averaging and a coordinate system for the cortical surface. *Human Brain Mapping*, 8(4):272–284, 1999a. doi:10.1002/(SICI)1097-0193(1999)8:4<272::AID-HBM10>3.0.CO;2-4.
- B. Fischl, M. I. Sereno, and A. M. Dale. Cortical surface-based analysis: I: Inflation, flattening, and a surface-based coordinate system. *NeuroImage*, 9(2):195 – 207, 1999b. ISSN 1053-8119. doi:10.1006/nimg.1998.0396.

- P. Gaglioti, D. Danelon, S. Bontempo, M. Mombrò, S. Cardaropoli, and T. Todros. Fetal cerebral ventriculomegaly: outcome in 176 cases. *Ultrasound in Obstetrics and Gynecology*, 25(4):372–377, 2005. doi:10.1002/uog.1857.
- C. Garel, E. Chantrel, H. Brisse, M. Elmaleh, D. Luton, J. Oury, G. Sebag, and M. Hassan. Fetal cerebral cortex: Normal gestational landmarks identified using prenatal mr imaging. *American Journal of Neuroradiology*, 22(1):184–189, 2001.
- R. Geva, R. Eshel, Y. Leitner, A. Valevski, and S. Harel. Neuropsychological outcome of children with intrauterine growth restriction: A 9-year prospective study. *Pediatrics*, 118(1):91–100, 2006. doi:10.1542/peds.2005-2343.
- A. Gholipour, A. Akhondi-Asl, J. Estroff, and S. Warfield. Multi-atlas multi-shape segmentation of fetal brain MRI for volumetric and morphometric analysis of ventriculomegaly. *NeuroImage*, 60(3):1819–1831, 2012. doi:10.1016/j.neuroimage.2012.01.128.
- A. Gholipour, J. Estroff, and S. Warfield. Robust super-resolution volume reconstruction from slice acquisitions: Application to fetal brain MRI. *IEEE Transactions on Medical Imaging*, 29(10):1739–1758, 2010. doi:10.1109/tmi.2010.2051680.
- S. F. Gibson. Constrained elastic surface nets: Generating smooth surfaces from binary segmented data. In *Medical Image Computing and Computer-Assisted Intervention —MICCAI’ 98*, volume 1496 of *Lecture Notes in Computer Science*, pages 888–898. Springer Berlin Heidelberg, 1998. ISBN 978-3-540-65136-9. doi:10.1007/BFb0056277.
- C. Glastonbury and A. Kennedy. Ultrafast MRI of the fetus. *Australasian Radiology*, 46(1):22–32, 2002. doi:10.1046/j.1440-1673.2001.00990.x.
- I. Gousias, A. Edwards, M. Rutherford, S. Counsell, J. Hajnal, D. Rueckert, and A. Hammers. Magnetic resonance imaging of the newborn brain: Manual seg-

- mentation of labelled atlases in term-born and preterm infants. *NeuroImage*, 62(3):1499–1509, 2012. doi:10.1016/j.neuroimage.2012.05.083.
- I. Guskov and Z. J. Wood. Topological noise removal. In *Proceedings of Graphics Interface 2001*, GI '01, pages 19–26. Canadian Information Processing Society, Toronto, Ont., Canada, 2001. ISBN 0-9688808-0-0.
- P. Habas, K. Kim, F. Rousseau, O. Glenn, A. Barkovich, and C. Studholme. Atlas-based segmentation of developing tissues in the human brain with quantitative validation in young fetuses. *Human Brain Mapping*, 31(9):1348–1358, 2010. doi:10.1002/hbm.20935.
- P. Habas, J. Scott, A. Roosta, V. Rajagopalan, K. Kim, F. Rousseau, A. Barkovich, O. Glenn, and C. Studholme. Early folding patterns and asymmetries of the normal human brain detected from in utero MRI. *Cerebral Cortex*, 22(1):13–25, 2012. doi:10.1093/cercor/bhr053.
- R. Heckemann, J. Hajnal, P. Aljabar, D. Rueckert, and A. Hammers. Automatic anatomical brain MRI segmentation combining label propagation and decision fusion. *NeuroImage*, 33(1):115 – 126, 2006. doi:10.1016/j.neuroimage.2006.05.061.
- P. Hellier, C. Barillot, E. Memin, and P. Perez. Hierarchical estimation of a dense deformation field for 3-d robust registration. *Medical Imaging, IEEE Transactions on*, 20(5):388–402, 2001. ISSN 0278-0062. doi:10.1109/42.925292.
- J. Hill, D. Dierker, J. Neil, T. Inder, A. Knutsen, J. Harwell, T. Coalson, and D. Van Essen. A surface-based analysis of hemispheric asymmetries and folding of cerebral cortex in term-born human infants. *The Journal of Neuroscience*, 30(6):2268–2276, 2010. doi:10.1523/JNEUROSCI.4682-09.2010.

- C. Hillenbrand and G. Huisman. *MRI of the Newborn, pt. 1*. Magnetic Resonance Imaging Clinics. Elsevier - Health Sciences Division, fourth edition, 2012. ISBN 9781437726923.
- J. I. Hoffman and S. Kaplan. The incidence of congenital heart disease. *Journal of the American College of Cardiology*, 39(12):1890–1900, 2002. doi:10.1016/S0735-1097(02)01886-7.
- I. Hosny and H. Elghawabi. Ultrafast MRI of the fetus: an increasingly important tool in prenatal diagnosis of congenital anomalies. *Magnetic Resonance Imaging*, 28(10):1431–1439, 2010. doi:10.1016/j.mri.2010.06.024.
- T. E. Inder, P. S. Huppi, S. Warfield, R. Kikinis, G. P. Zientara, P. D. Barnes, F. A. Jolesz, and J. J. Volpe. Periventricular white matter injury in the premature infant is followed by reduced cerebral cortical gray matter volume at term. *Annals of neurology*, 46(5):755–760, 1999.
- F. Jacob, P. Habas, K. Kim, J. Corbett-Detig, D. Xu, C. Studholme, and O. Glenn. Fetal hippocampal development: analysis by magnetic resonance imaging volumetry. *Pediatric Research*, 69(5 pt 1):425–429, 2011. doi:10.1203/pdr.obo13e318211dd7f.
- V. Jain and H. Zhang. Robust 3d shape correspondence in the spectral domain. In *Shape Modeling and Applications, 2006. SMI 2006. IEEE International Conference on*, pages 19–19. 2006. doi:10.1109/SMI.2006.31.
- S. Jiang, H. Xue, A. Glover, M. Rutherford, D. Rueckert, and J. Hajnal. MRI of moving subjects using multislice snapshot images with volume reconstruction (SVR): Application to fetal, neonatal, and adult brain studies. *IEEE Transactions on Medical Imaging*, 26(7):967–980, 2007. doi:10.1109/tmi.2007.895456.

- S. Johnson, E. Hennessy, R. Smith, R. Trikic, D. Wolke, and N. Marlow. Academic attainment and special educational needs in extremely preterm children at 11 years of age: the EPICure study. *Archives of Disease in Childhood - Fetal and Neonatal Edition*, 94(4):F283–F289, 2009. doi:10.1136/adc.2008.152793.
- S. E. Jones, B. R. Buchbinder, and I. Aharon. Three-dimensional mapping of cortical thickness using laplace’s equation. *Human Brain Mapping*, 11(1):12–32, 2000. ISSN 1097-0193. doi:10.1002/1097-0193(200009)11:1<12::AID-HBM20>3.0.CO;2-K.
- M. Joyce A. Martin, P. Brady E. Hamilton, M. Michelle J.K. Osterman, M. Sally C. Curtin, and M. T.J. Matthews. Births: Final data for 2013. *National Vital. Statistics Reports.*, 64(1), 2015.
- T. Ju, F. Losasso, S. Schaefer, and J. Warren. Dual contouring of hermite data. *ACM Trans. Graph.*, 21(3):339–346, 2002. ISSN 0730-0301. doi:10.1145/566654.566586.
- B. Kainz, M. Steinberger, W. Wein, M. Kuklisova-Murgasova, C. Malamate-niou, K. Keraudren, T. Torsney-Weir, M. Rutherford, P. Aljabar, J. Hajnal, and D. Rueckert. Fast volume reconstruction from motion corrupted stacks of 2d slices. *Medical Imaging, IEEE Transactions on*, 34(9):1901–1913, 2015. ISSN 0278-0062. doi:10.1109/TMI.2015.2415453.
- E. Kandel. *Principles of Neural Science, Fifth Edition*. Principles of Neural Science. McGraw-Hill Education, 2013. ISBN 9780071390118.
- O. Kapellou, S. Counsell, N. Kennea, L. Dyet, N. Saeed, J. Stark, E. Maalouf, P. Duggan, M. Ajayi-Obe, J. Hajnal, J. Allsop, J. Boardman, M. Rutherford, F. Cowan, and A. Edwards. Abnormal cortical development after premature birth shown by altered allometric scaling of brain growth. *PLoS Med*, 3(8):e265, 2006. doi:10.1371/journal.pmed.0030265.

- G. Karas, P. Scheltens, S. Rombouts, P. Visser, R. van Schijndel, N. Fox, and F. Barkhof. Global and local gray matter loss in mild cognitive impairment and alzheimer's disease. *NeuroImage*, 23(2):708 – 716, 2004. ISSN 1053-8119. doi:10.1016/j.neuroimage.2004.07.006.
- M. Kazhdan. Reconstruction of solid models from oriented point sets. In *Proceedings of the Third Eurographics Symposium on Geometry Processing*, SGP '05. Eurographics Association, 2005.
- M. Kazhdan and H. Hoppe. Screened poisson surface reconstruction. *ACM Trans. Graph.*, 32(3):29:1–29:13, 2013. doi:10.1145/2487228.2487237.
- S. S. Keller, M. Wilke, U. C. Wieshmann, V. A. Sluming, and N. Roberts. Comparison of standard and optimized voxel-based morphometry for analysis of brain changes associated with temporal lobe epilepsy. *NeuroImage*, 23(3):860 – 868, 2004. ISSN 1053-8119. doi:10.1016/j.neuroimage.2004.07.030.
- J. S. Kim, V. Singh, J. K. Lee, J. Lerch, Y. Ad-Dab'bagh, D. MacDonald, J. M. Lee, S. I. Kim, and A. C. Evans. Automated 3-d extraction and evaluation of the inner and outer cortical surfaces using a laplacian map and partial volume effect classification. *NeuroImage*, 27(1):210 – 221, 2005. doi:10.1016/j.neuroimage.2005.03.036.
- K. Kim, P. Habas, F. Rousseau, O. Glenn, A. Barkovich, and C. Studholme. Intersection based motion correction of multislice MRI for 3-D in utero fetal brain image formation. *IEEE Transactions on Medical Imaging*, 29(1):146–158, 2010. doi:10.1109/tmi.2009.2030679.
- J. J. Koenderink and A. J. van Doorn. Surface shape and curvature scales. *Image and Vision Computing*, 10(8):557 – 564, 1992. ISSN 0262-8856. doi:10.1016/0262-8856(92)90076-F.

- M. Kuklisova-Murgasova, P. Aljabar, L. Srinivasan, S. Counsell, V. Doria, A. Serag, I. Gousias, J. Boardman, M. Rutherford, A. Edwards, J. Hajnal, and D. Rueckert. A dynamic 4D probabilistic atlas of the developing brain. *NeuroImage*, 54(4):2750–2763, 2011. doi:10.1016/j.neuroimage.2010.10.019.
- M. Kuklisova-Murgasova, G. Quaghebeur, M. Rutherford, J. Hajnal, and J. Schnabel. Reconstruction of fetal brain MRI with intensity matching and complete outlier removal. *Medical Image Analysis*, 16(8):1550–1564, 2012. doi:10.1016/j.media.2012.07.004.
- V. Kyriakopoulou, D. Vatansever, S. Elkommos, S. Dawson, A. McGuinness, J. Allsop, Z. Molnár, J. Hajnal, and M. Rutherford. Cortical overgrowth in fetuses with isolated ventriculomegaly. *Cerebral Cortex*, 24(8):2141–2150, 2014. doi:10.1093/cercor/bht062.
- A. Laird. Dynamics of tumour growth. *British journal of cancer*, 18(3):490–502, 1964.
- C. Ledig, P. Aljabar, R. Wright, A. Serag, and D. Rueckert. Neonatal brain segmentation using second order neighborhood information. In *Workshop on Perinatal and Paediatric Imaging: PaPI, Medical Image Computing and Computer-Assisted Intervention: MICCAI*, pages 33–40. 2012.
- J. Lefèvre, D. Germanaud, J. Dubois, F. Rousseau, I. de Macedo Santos, H. Angleys, J.-F. Mangin, P. S. Hüppi, N. Girard, and F. De Guio. Are developmental trajectories of cortical folding comparable between cross-sectional datasets of fetuses and preterm newborns? *Cerebral Cortex*, 2015. doi:10.1093/cercor/bhv123.
- G. Li, L. Wang, F. Shi, W. Lin, and D. Shen. Constructing 4d infant cortical surface atlases based on dynamic developmental trajectories of the cortex. In *Medical Image Computing and Computer-Assisted Intervention*, volume 8675 of *Lecture Notes*

- in *Computer Science*, pages 89–96. Springer International Publishing, 2014. doi:10.1007/978-3-319-10443-0_12.
- C. Limperopoulos, W. Tworetzky, D. B. McElhinney, J. W. Newburger, D. W. Brown, R. L. Robertson, N. Guizard, E. McGrath, J. Geva, D. Annese, C. Dunbar-Masterson, B. Trainor, P. C. Laussen, and A. J. du Plessis. Brain volume and metabolism in fetuses with congenital heart disease: Evaluation with quantitative magnetic resonance imaging and spectroscopy. *Circulation*, 121(1):26–33, 2010. doi:10.1161/CIRCULATIONAHA.109.865568.
- H. Lombaert, L. Grady, J. Polimeni, and F. Cheriet. Fast brain matching with spectral correspondence. In *Information Processing in Medical Imaging*, volume 6801 of *Lecture Notes in Computer Science*, pages 660–673. Springer Berlin Heidelberg, 2011. doi:10.1007/978-3-642-22092-0_54.
- H. Lombaert, L. Grady, J. Polimeni, and F. Cheriet. Focusr: Feature oriented correspondence using spectral regularization—a method for precise surface matching. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 35(9):2143–2160, 2013a. doi:10.1109/TPAMI.2012.276.
- H. Lombaert, J. Sporring, and K. Siddiqi. Diffeomorphic spectral matching of cortical surfaces. In *Information Processing in Medical Imaging (IPMI)*, volume 7917 of *Lecture Notes in Computer Science*, pages 376–389. Springer, Springer, 2013b. doi:10.1007/978-3-642-38868-2_32.
- J. G. Long, J. F. Lucey, and A. G. S. Philip. Noise and hypoxemia in the intensive care nursery. *Pediatrics*, 65(1):143–145, 1980.
- W. Lorensen and H. Cline. Marching cubes: A high resolution 3d surface construction algorithm. In *ACM Siggraph Computer Graphics*, volume 21, pages 163–169. 1987.

- J. Lui, D. Hansen, and A. Kriegstein. Development and evolution of the human neocortex. *Cell*, 146(1):18–36, 2011. doi:10.1016/j.cell.2011.06.030.
- O. Lyttelton, M. Boucher, S. Robbins, and A. Evans. An unbiased iterative group registration template for cortical surface analysis. *NeuroImage*, 34(4):1535 – 1544, 2007. doi:10.1016/j.neuroimage.2006.10.041.
- D. MacDonald, N. Kabani, D. Avis, and A. C. Evans. Automated 3-d extraction of inner and outer surfaces of cerebral cortex from MRI. *NeuroImage*, 12(3):340 – 356, 2000. doi:10.1006/nimg.1999.0534.
- R. Maciunas, R. Kessler, C. Maurer, V. Mandava, G. Watt, and G. Smith. Positron emission tomography imaging-directed stereotactic neurosurgery. *Stereotactic and Functional Neurosurgery*, 58(1-4):134–140, 1992. ISSN 1011-6125. doi:10.1159/000098986.
- V. A. Magnotta, N. C. Andreasen, S. K. Schultz, G. Harris, T. Cizadlo, D. Heckel, P. Nopoulos, and M. Flaum. Quantitative in vivo measurement of gyrification in the human brain: Changes associated with aging. *Cerebral Cortex*, 9(2):151–160, 1999. doi:10.1093/cercor/9.2.151.
- A. Makropoulos. *Automatic MRI segmentation of the developing neonatal brain*. Ph.D. thesis, Imperial College London, 2014.
- A. Makropoulos, I. Gousias, C. Ledig, P. Aljabar, A. Serag, J. Hajnal, A. Edwards, S. Counsell, and D. Rueckert. Automatic whole brain mri segmentation of the developing neonatal brain. *Medical Imaging, IEEE Transactions on*, 33(9):1818–1831, 2014. doi:10.1109/TMI.2014.2322280.
- D. Mateus, R. Horaud, D. Knossow, F. Cuzzolin, and E. Boyer. Articulated shape matching using laplacian eigenfunctions and unsupervised point registration.

- In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. 2008. doi:10.1109/CVPR.2008.4587538.
- MATLAB and S. Toolbox. *version 7.13.0.564 (R2011b)*. The MathWorks Inc., Natick, Massachusetts, United States, 2011.
- J. Maurer, C.R., R. Qi, and V. Raghavan. A linear time algorithm for computing exact euclidean distance transforms of binary images in arbitrary dimensions. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 25(2):265–270, 2003. doi:10.1109/TPAMI.2003.1177156.
- A. Melbourne, G. Kendall, M. Cardoso, R. Gunney, N. Robertson, N. Marlow, and S. Ourselin. Radial structure in the preterm cortex; persistence of the preterm phenotype at term equivalent age? In *Medical Image Computing and Computer-Assisted Intervention –MICCAI 2012*, volume 7512 of *Lecture Notes in Computer Science*, pages 256–263. Springer Berlin Heidelberg, 2012. ISBN 978-3-642-33453-5. doi:10.1007/978-3-642-33454-2_32.
- A. Melbourne, G. S. Kendall, M. J. Cardoso, R. Gunny, N. J. Robertson, N. Marlow, and S. Ourselin. Preterm birth affects the developmental synergy between cortical folding and cortical connectivity observed on multimodal MRI. *NeuroImage*, 89:23 – 34, 2014. ISSN 1053-8119. doi:10.1016/j.neuroimage.2013.11.048.
- E. Nadaraya. On estimating regression. *Theory of Probability & Its Applications*, 9(1):141–142, 1964. doi:10.1137/1109020.
- G. M. Nielson and B. Hamann. The asymptotic decider: Resolving the ambiguity in marching cubes. In *Proceedings of the 2Nd Conference on Visualization '91, VIS '91*, pages 83–91. IEEE Computer Society Press, Los Alamitos, CA, USA, 1991. ISBN 0-8186-2245-8.

- L. Nyúl and J. Udupa. On standardizing the MR image intensity scale. *Magnetic Resonance in Medicine*, 42(6):1072–1081, 1999. doi:10.1002/(sici)1522-2594(199912)42:6<1072::aid-mrm11>3.0.co;2-m.
- E. Olson. On computing the average orientation of vectors and lines. In *Robotics and Automation (ICRA), 2011 IEEE International Conference on*, pages 3869–3874. 2011. doi:10.1109/ICRA.2011.5979562.
- OpenStax College. Image depicting brodmann areas, anatomy & physiology. openstax cnx. <http://cnx.org/content/col11496/1.6>. Licensed under CC BY 3.0.
- E. Orasanu, A. Melbourne, H. Lombaert, M. Cardoso, S.F.Johnsen, G. Kendall, N. Robertson, N. Marlow, and S. Ourselin. Prefrontal cortical folding of the preterm brain: a longitudinal analysis of preterm-born neonates. In *MICCAI Workshop on Spatiotemporal Image Analysis for Longitudinal and Time-Series Image Data (STIA)*. 2014.
- C. Ortinau, D. Alexopoulos, D. Dierker, D. V. Essen, J. Beca, and T. Inder. Cortical folding is altered before surgery in infants with congenital heart disease. *The Journal of Pediatrics*, 163(5):1507 – 1510, 2013. ISSN 0022-3476. doi:10.1016/j.jpeds.2013.06.045.
- E. Parzen. On estimation of a probability density function and mode. *Ann. Math. Statist.*, 33(3):1065–1076, 1962. doi:10.1214/aoms/1177704472.
- C. F. A. Peralta, P. Cavoretto, B. Csapo, O. Falcon, and K. H. Nicolaides. Lung and heart volumes by three-dimensional ultrasound in normal fetuses at 12–32 weeks’ gestation. *Ultrasound in Obstetrics and Gynecology*, 27(2):128–133, 2006. ISSN 1469-0705. doi:10.1002/uog.2670.

- U. Pöhls and A. Rempen. Fetal lung volumetry by three-dimensional ultrasound. *Ultrasound in Obstetrics and Gynecology*, 11(1):6–12, 1998. ISSN 1469-0705. doi:10.1046/j.1469-0705.1998.11010006.x.
- M. Prastawa, J. H. Gilmore, W. Lin, and G. Gerig. Automatic segmentation of MR images of the developing newborn brain. *Medical Image Analysis*, 9(5):457 – 466, 2005. ISSN 1361-8415. doi:10.1016/j.media.2005.05.007.
- S. Price, D. Paviour, R. Scahill, J. Stevens, M. Rossor, A. Lees, and N. Fox. Voxel-based morphometry detects patterns of atrophy that help differentiate progressive supranuclear palsy and parkinson’s disease. *NeuroImage*, 23(2):663 – 669, 2004. ISSN 1053-8119. doi:10.1016/j.neuroimage.2004.06.013.
- V. Rajagopalan, J. Scott, P. Habas, K. Kim, J. Corbett-Detig, F. Rousseau, A. Barkovich, O. Glenn, and C. Studholme. Local tissue growth patterns underlying normal fetal human brain gyrification quantified in utero. *The Journal of Neuroscience*, 31(8):2878–2887, 2011. doi:10.1523/jneurosci.5458-10.2011.
- P. Rakic. Evolution of the neocortex: a perspective from developmental biology. *Nat Rev Neurosci*, 10(10):724 – 735, 2009. doi:10.1038/nrn2719.
- D. P. Richman, R. M. Stewart, J. W. Hutchinson, and J. Verne S. Caviness. Mechanical model of brain convolutional development. *Science*, 189(4196):18–21, 1975.
- B. Rieger, F. Timmermans, L. van Vliet, and P. Verbeek. On curvature estimation of iso surfaces in 3D gray-value images and the computation of shape descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(8):1088–1094, 2004. doi:10.1109/tpami.2004.50.
- E. C. Robinson, S. Jbabdi, M. F. Glasser, J. Andersson, G. C. Burgess, M. P. Harms, S. M. Smith, D. C. V. Essen, and M. Jenkinson. Msm: A new flexible framework

- for multimodal surface matching. *NeuroImage*, 100:414 – 426, 2014. ISSN 1053-8119. doi:j.neuroimage.2014.05.069.
- A. Roche, G. Malandain, X. Pennec, and N. Ayache. The correlation ratio as a new similarity measure for multimodal image registration. In *Proceedings of the First International Conference on Medical Image Computing and Computer-Assisted Intervention*, MICCAI '98, pages 1115–1124. Springer-Verlag, London, UK, 1998.
- C. Rodriguez-Carranza, P. Mukherjee, D. Vigneron, J. Barkovich, and C. Studholme. A framework for in vivo quantification of regional brain folding in premature neonates. *NeuroImage*, 41(2):462–478, 2008. doi:10.1016/j.neuroimage.2008.01.008.
- T. Rohlfing, D. Russakoff, and C. Maurer. Expectation maximization strategies for multi-atlas multi-label segmentation. In *Information Processing in Medical Imaging*, volume 2732 of *Lecture Notes in Computer Science*, pages 210–221. Springer Berlin Heidelberg, 2003. doi:10.1007/978-3-540-45087-0_18.
- M. Rosenblatt. Remarks on some nonparametric estimates of a density function. *Ann. Math. Statist.*, 27(3):832–837, 1956. doi:10.1214/aoms/1177728190.
- F. Rousseau, O. Glenn, B. Iordanova, C. Rodriguez-Carranza, D. Vigneron, J. Barkovich, and C. Studholme. Registration-based approach for reconstruction of high-resolution in utero fetal MR brain images. *Academic Radiology*, 13(9):1072–1081, 2006. doi:10.1016/j.acra.2006.05.003.
- F. Rousseau, K. Kim, C. Studholme, M. Koob, and J. Dietemann. On super-resolution for fetal brain MRI. In *Medical Image Computing and Computer-Assisted Intervention: MICCAI*, volume 6362 of *Lecture Notes in Computer Science: LNCS*, pages 355–362. 2010. doi:10.1007/978-3-642-15745-5_44.

- S. Ruan, C. Jaggi, J. Xue, J. Fadili, and D. Bloyet. Brain tissue classification of magnetic resonance images using partial volume modeling. *Medical Imaging, IEEE Transactions on*, 19(12):1179–1187, 2000. ISSN 0278-0062. doi:10.1109/42.897810.
- D. Rueckert, L. Sonoda, C. Hayes, D. Hill, M. Leach, and D. Hawkes. Nonrigid registration using free-form deformations: application to breast MR images. *IEEE Transactions on Medical Imaging*, 18(8):712–721, 1999. doi:10.1109/42.796284.
- M. Rutherford. *MRI of the Neonatal Brain*. W.B. Saunders, fourth edition, 2002. ISBN 9780702025341.
- M. Rutherford, S. Jiang, J. Allsop, L. Perkins, L. Srinivasan, T. Hayat, S. Kumar, and J. Hajnal. MR imaging methods for assessing fetal brain development. *Developmental Neurobiology*, 68(6):700–711, 2008. doi:10.1002/dneu.20614.
- J. Scott, P. Habas, V. Rajagopalan, K. Kim, A. Barkovich, O. Glenn, and C. Studholme. Volumetric and surface-based 3D MRI analyses of fetal isolated mild ventriculomegaly. *Brain Structure and Function*, 218(3):645–655, 2013. doi:10.1007/s00429-012-0418-1.
- J. A. Scott, P. A. Habas, K. Kim, V. Rajagopalan, K. Hamzelou, J. Corbett-Detig, A. Barkovich, O. Glenn, and C. Studholme. Growth trajectories of the human fetal brain tissues estimated from 3D reconstructed in utero MRI. *International Journal of Developmental Neuroscience*, 29(5):529–536, 2011. doi:10.1016/j.ijdevneu.2011.04.001.
- D. Seghers, E. D’Agostino, F. Maes, D. Vandermeulen, and P. Suetens. Construction of a brain template from mr images using state-of-the-art registration and segmentation techniques. In *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2004*, volume 3216 of *Lecture Notes in Computer Science*,

- pages 696–703. Springer Berlin Heidelberg, 2004. ISBN 978-3-540-22976-6. doi:10.1007/978-3-540-30135-6_85.
- F. Ségonne, E. Grimson, and B. Fischl. A genetic algorithm for the topology correction of cortical surfaces. In *Information Processing in Medical Imaging*, volume 3565 of *Lecture Notes in Computer Science*, pages 393–405. Springer Berlin Heidelberg, 2005. doi:10.1007/11505730_33.
- F. Ségonne, J. Pacheco, and B. Fischl. Geometrically accurate topology-correction of cortical surfaces using nonseparating loops. *Medical Imaging, IEEE Transactions on*, 26(4):518–529, 2007. ISSN 0278-0062. doi:10.1109/TMI.2006.887364.
- A. Serag, P. Aljabar, G. Ball, S. Counsell, J. Boardman, M. Rutherford, A. Edwards, J. Hajnal, and D. Rueckert. Construction of a consistent high-definition spatio-temporal atlas of the developing brain using adaptive kernel regression. *NeuroImage*, 59(3):2255–2265, 2012a. doi:10.1016/j.neuroimage.2011.09.062.
- A. Serag, V. Kyriakopoulou, P. Aljabar, S. Counsell, J. Boardman, M. Rutherford, A. Edwards, J. Hajnal, and D. Rueckert. A multi-channel 4D probabilistic atlas of the developing brain: Application to fetuses and neonates. In *Annals of the British Machine Vision Association*, 3, pages 1–14. 2012b.
- P. Shaw, J. Lerch, D. Greenstein, W. Sharp, L. Clasen, A. Evans, J. Giedd, F. Castellanos, and J. Rapoport. Longitudinal mapping of cortical thickness and clinical outcome in children and adolescents with attention-deficit/hyperactivity disorder. *Archives of General Psychiatry*, 63(5):540–549, 2006. doi:10.1001/archpsyc.63.5.540.
- S. Smith. Fast robust automated brain extraction. *Human Brain Mapping*, 17(3):143–155, 2002. doi:10.1002/hbm.10062.

- R. J. M. Snijders and K. H. Nicolaides. Fetal biometry at 14–40 weeks' gestation. *Ultrasound in Obstetrics and Gynecology*, 4(1):34–48, 1994. ISSN 1469-0705. doi: 10.1046/j.1469-0705.1994.04010034.x.
- T. Sørensen. A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons. *Biol. Skr.*, 5:1–34, 1948.
- A. Sotiras, C. Davatzikos, and N. Paragios. Deformable medical image registration: A survey. *Medical Imaging, IEEE Transactions on*, 32(7):1153–1190, 2013. ISSN 0278-0062. doi:10.1109/TMI.2013.2265603.
- C. Studholme, D. Hill, and D. Hawkes. An overlap invariant entropy measure of 3d medical image alignment. *Pattern Recognition*, 32(1):71 – 86, 1999. ISSN 0031-3203. doi:10.1016/S0031-3203(98)00091-0.
- J. Tao. Research on gompertz curve model used for mobile user growth. In *International Conference on Educational and Network Technology*, pages 557–560. 2010. doi:10.1109/ICENT.2010.5532100.
- J. B. Tenenbaum, V. d. Silva, and J. C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500):2319–2323, 2000. doi: 10.1126/science.290.5500.2319.
- G. Treece, R. Prager, and A. Gee. Regularised marching tetrahedra: improved iso-surface extraction. *Computers & Graphics*, 23(4):583 – 598, 1999. ISSN 0097-8493. doi:10.1016/S0097-8493(99)00076-X.
- N. Tustison, B. Avants, P. Cook, Y. Zheng, A. Egan, P. Yushkevich, and J. Gee. N4ITK: Improved N₃ bias correction. *IEEE Transactions on Medical Imaging*, 29(6):1310–1320, 2010. doi:10.1109/tmi.2010.2046908.

- D. Van Essen. A population-average, landmark- and surface-based (pals) atlas of human cerebral cortex. *NeuroImage*, 28(3):635 – 662, 2005. doi:10.1016/j.neuroimage.2005.06.058.
- D. Van Essen and H. Drury. Structural and functional analyses of human cerebral cortex using a surface-based atlas. *The Journal of Neuroscience*, 17(18):7079–7102, 1997.
- K. Van Leemput, F. Maes, D. Vandermeulen, and P. Suetens. Automated model-based tissue classification of MR images of the brain. *IEEE Transactions on Medical Imaging*, 18(10):897–908, 1999. doi:10.1109/42.811270.
- K. Van Leemput, F. Maes, D. Vandermeulen, and P. Suetens. A unifying framework for partial volume segmentation of brain mr images. *Medical Imaging, IEEE Transactions on*, 22(1):105–119, 2003. ISSN 0278-0062. doi:10.1109/TMI.2002.806587.
- L. J. van Vliet and P. W. Verbeek. Curvature and bending energy in digitized 2D and 3D images. In *Scandinavian Conference on Image Analysis: SCIA*, volume 2, pages 1403–1410. 1993.
- T. Victoria, A. M. Johnson, J. C. Edgar, D. M. Zarnow, A. Vossough, and D. Jaramillo. Comparison between 1.5-t and 3-t mri for fetal imaging: Is there an advantage to imaging with a higher field strength? *American Journal of Roentgenology*, 206(1):195 – 201, 2016. doi:10.2214/AJR.14.14205.
- P. Viola and W. Wells. Alignment by maximization of mutual information. In *Computer Vision, 1995. Proceedings., Fifth International Conference on*, pages 16–23. 1995. doi:10.1109/ICCV.1995.466930.
- Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Medical Imaging*, 13(4):600–612, 2004. doi:10.1109/tip.2003.819861.

- S. Warfield, K. Zou, and W. Wells. Simultaneous truth and performance level estimation (staple): an algorithm for the validation of image segmentation. *Medical Imaging, IEEE Transactions on*, 23(7):903–921, 2004. ISSN 0278-0062. doi:10.1109/TMI.2004.828354.
- N. I. Weisenfeld, A. U. Mewes, and S. K. Warfield. Highly accurate segmentation of brain tissue and subcortical gray matter from newborn mri. In *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2006*, volume 4190 of *Lecture Notes in Computer Science*, pages 199–206. Springer Berlin Heidelberg, 2006. ISBN 978-3-540-44707-8. doi:10.1007/11866565_25.
- I. Wells, W.M., W. Grimson, R. Kikinis, and F. Jolesz. Adaptive segmentation of mri data. *Medical Imaging, IEEE Transactions on*, 15(4):429–442, 1996. ISSN 0278-0062. doi:10.1109/42.511747.
- Wikimedia Commons. Image depicting joint visualisation of aligned MR and PET images. <https://commons.wikimedia.org/wiki/File:PET-MR2-Head-Keosys.JPG>. Released into the public domain.
- G. Windreich, N. Kiryati, and G. Lohmann. Voxel-based surface area estimation: from theory to practice. *Pattern Recognition*, 36(11):2531–2541, 2003. doi:10.1016/S0031-3203(03)00173-0.
- R. Wright, D. Vatansever, V. Kyriakopoulou, C. Ledig, R. Wolz, A. Serag, D. Rueckert, M. Rutherford, J. Hajnal, and P. Aljabar. Age dependent fetal MR segmentation using manual and automated approaches. In *Workshop on Perinatal and Paediatric Imaging: PaPI, Medical Image Computing and Computer-Assisted Intervention: MICCAI*, pages 97–104. 2012.
- G. Xu, A. K. Knutsen, K. Dikranian, C. D. Kroenke, P. V. Bayly, and L. A. Taber. Axons pull on the brain, but tension does not drive cortical folding. *Journal of*

- Biomechanical Engineering*, 132(7):071013, 2010. ISSN 0148-0731. doi:10.1115/1.4001683.
- H. Xue, L. Srinivasan, S. Jiang, M. Rutherford, A. Edwards, D. Rueckert, and J. Hajnal. Automatic segmentation and reconstruction of the cortex from neonatal MRI. *NeuroImage*, 38(3):461–477, 2007. doi:10.1016/j.neuroimage.2007.07.030.
- I. Yanovsky, C. Le Guyader, A. Leow, W. Toga, Arthur, M. Thompson, Paul, and L. Vese. Unbiased Volumetric Registration via Nonlinear Elastic Regularization. In *2nd MICCAI Workshop on Mathematical Foundations of Computational Anatomy*. New-York, United States, 2008.
- B. Yeo, M. Sabuncu, T. Vercauteren, N. Ayache, B. Fischl, and P. Golland. Spherical demons: Fast diffeomorphic landmark-free surface registration. *Medical Imaging, IEEE Transactions on*, 29(3):650–668, 2010. doi:10.1109/TMI.2009.2030797.
- R. A. Yotter, R. Dahnke, P. M. Thompson, and C. Gaser. Topological correction of brain surface meshes using spherical harmonics. *Human Brain Mapping*, 32(7):1109–1124, 2011. ISSN 1097-0193. doi:10.1002/hbm.21095.
- K. Zilles, E. Armstrong, A. Schleicher, and H.-J. Kretschmann. The human pattern of gyrification in the cerebral cortex. *Anatomy and Embryology*, 179(2):173–179, 1988. ISSN 0340-2061. doi:10.1007/BF00304699.
- K. Zilles, N. Palomero-Gallagher, and K. Amunts. Development of cortical folding during evolution and ontogeny. *Trends in Neurosciences*, 36(5):275 – 284, 2013. ISSN 0166-2236. doi:10.1016/j.tins.2013.01.006.

APPENDIX

A.1 CORTICAL SURFACE TEMPLATES

Average cortical surface templates constructed for preterm babies and fetuses are shown (see Chapter 5) for multiple different viewpoints (lateral, medial, superior and inferior). The colour mapping shows mean curvature (red, convex; blue, concave), which enhances visual perception of the template geometries.

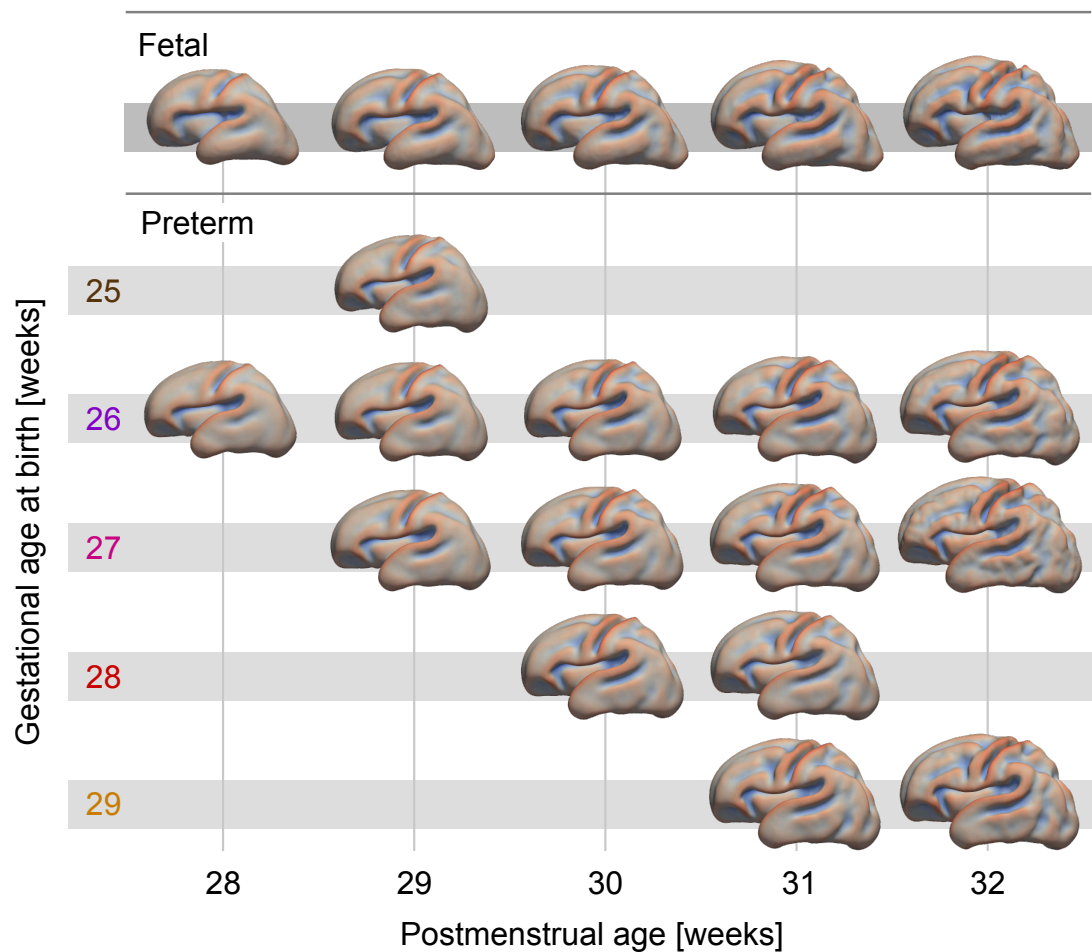


Figure 54: Lateral view, templates for extremely preterm babies (25-29 weeks GA at birth).

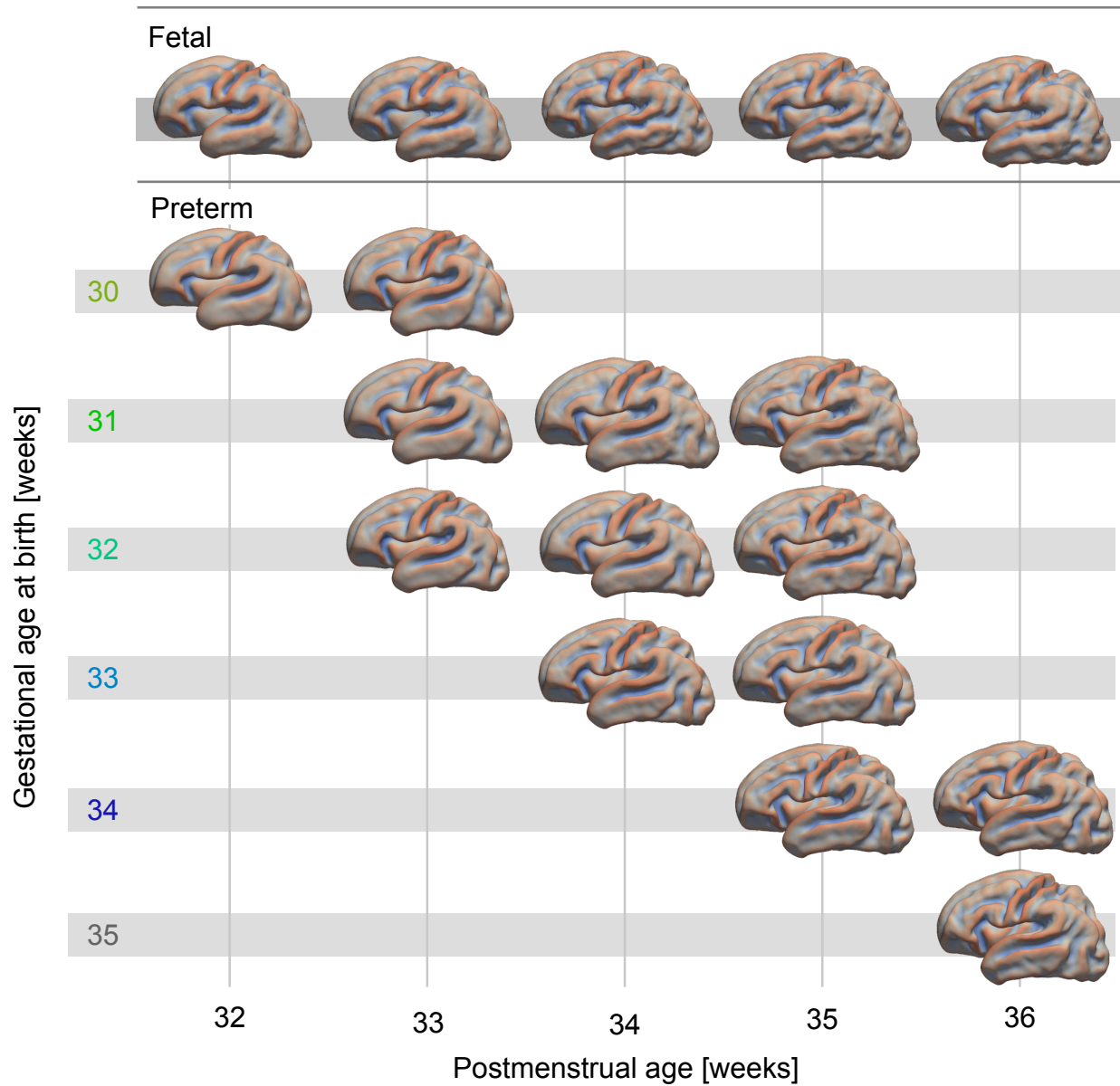


Figure 55: Lateral view, templates for very and moderately preterm babies (30-35 weeks GA at birth).

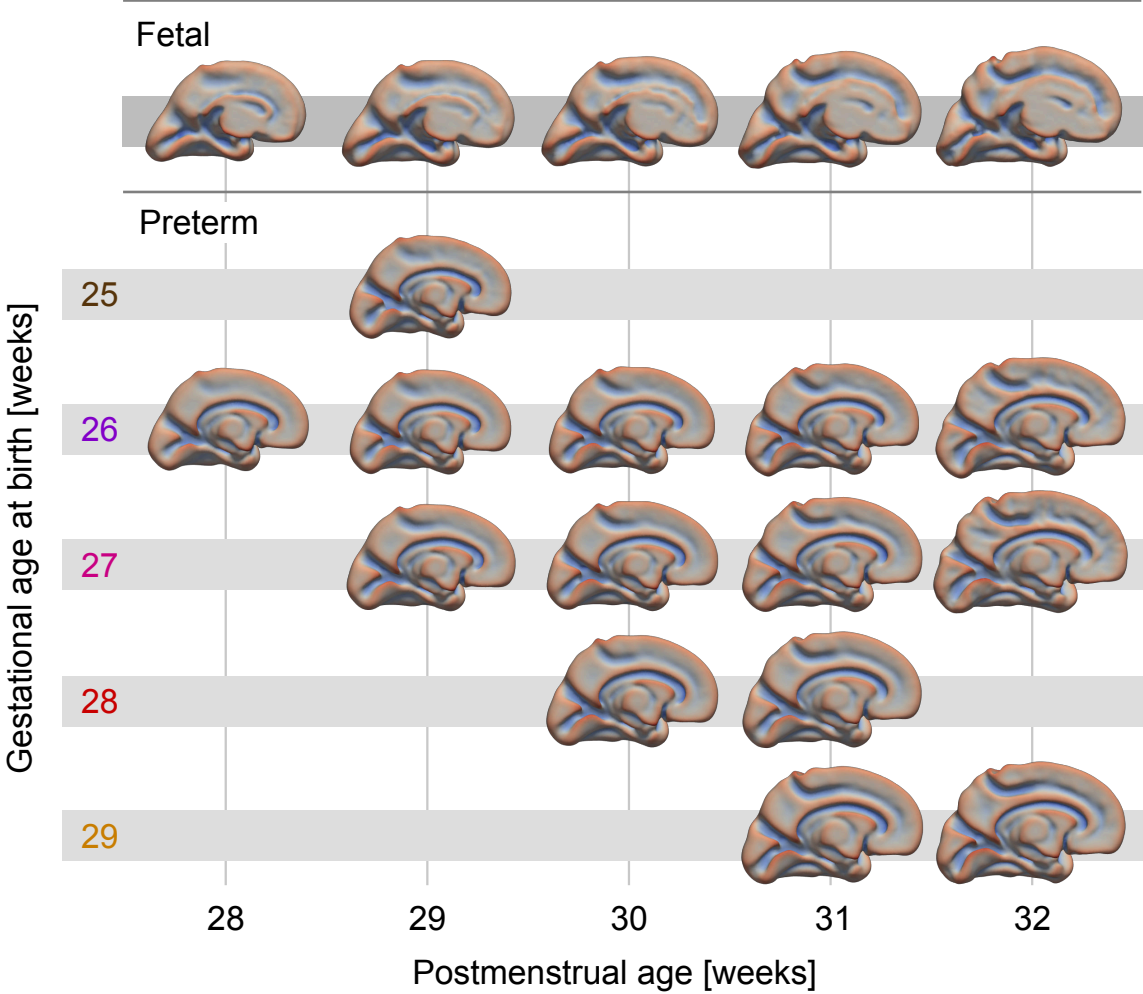


Figure 56: Medial view, templates for extremely preterm babies (25-29 weeks GA at birth).

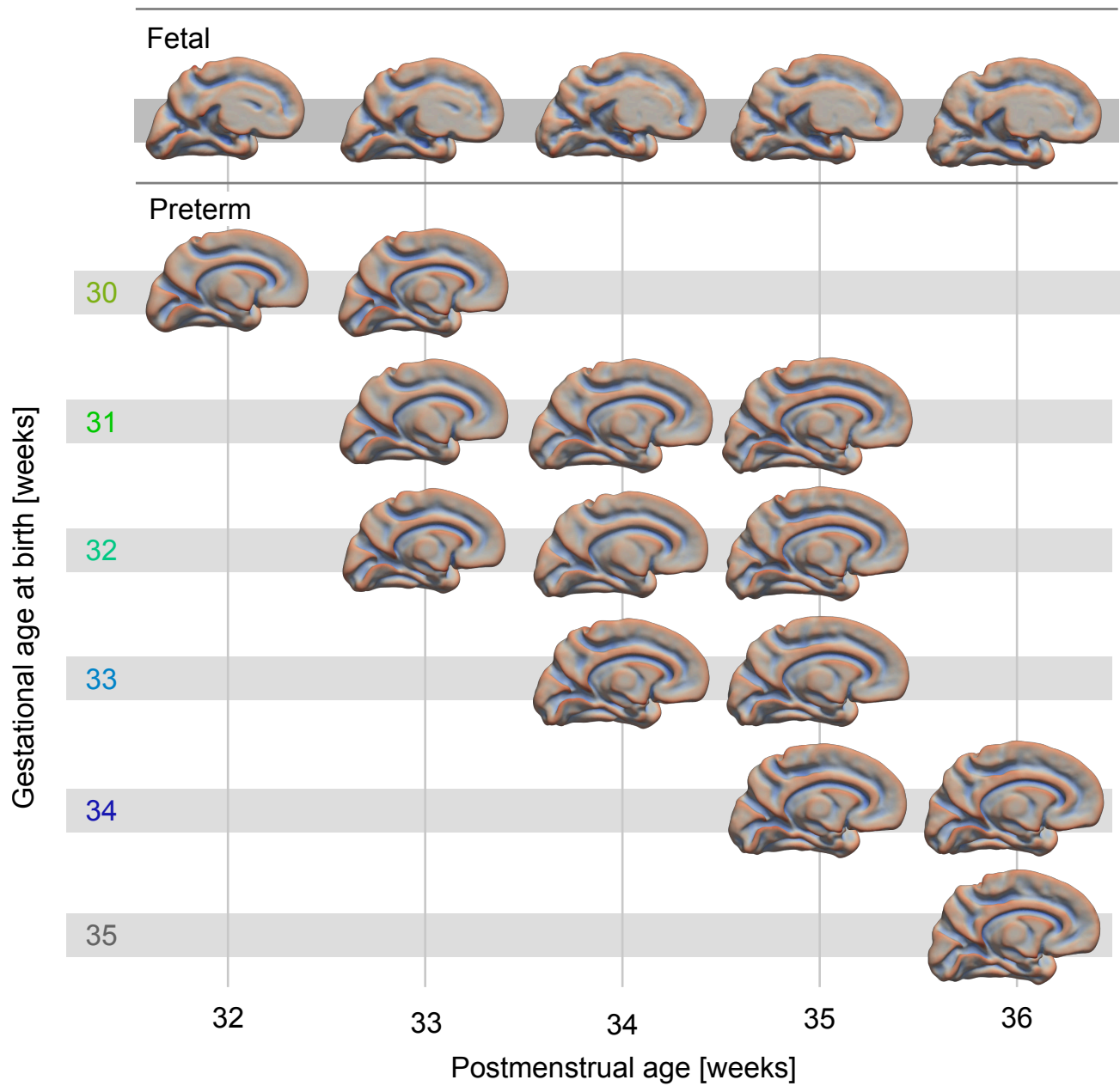


Figure 57: **Medial view, templates for extremely preterm babies (25-29 weeks GA at birth).**

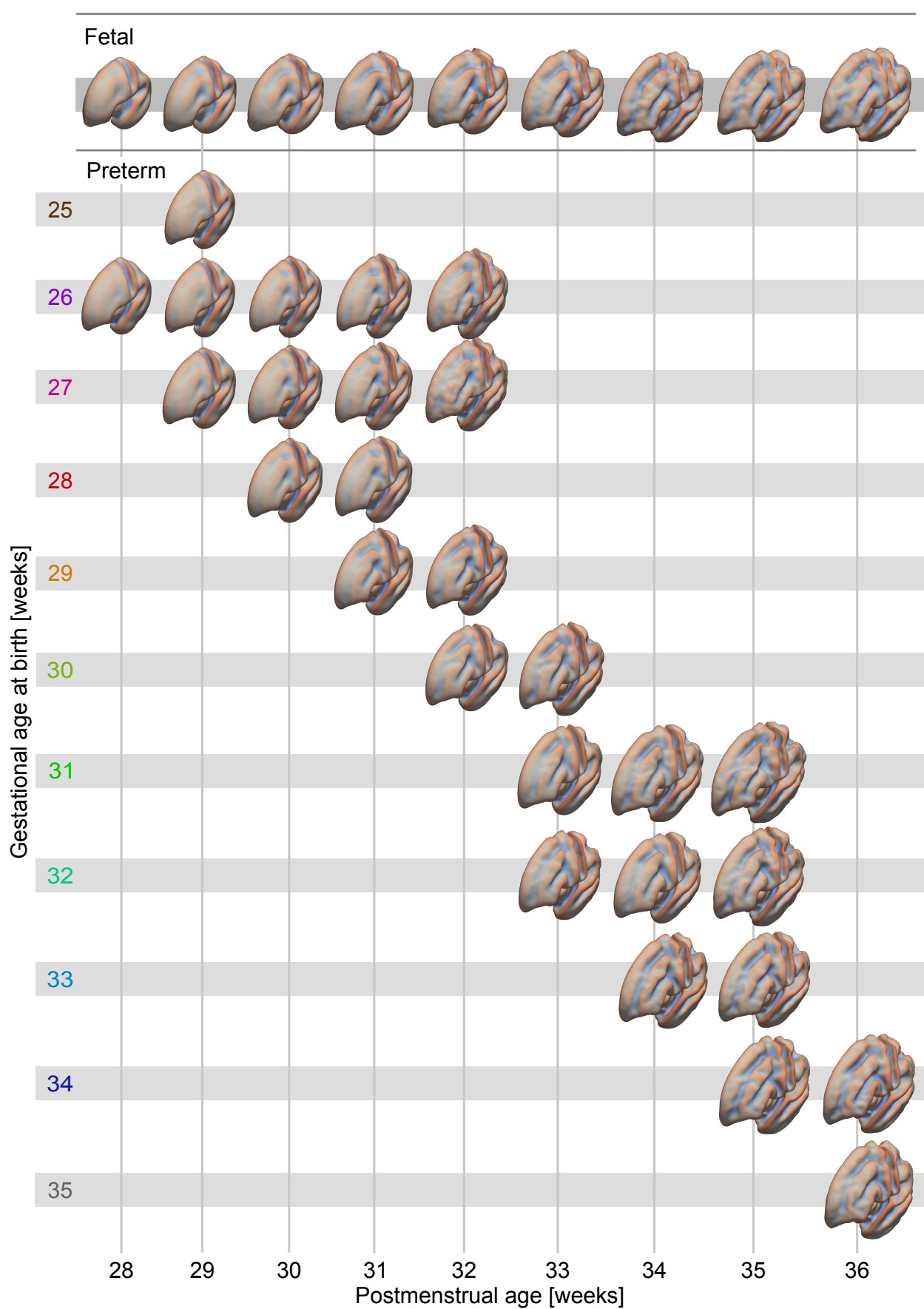


Figure 58: Superior/lateral view, templates for all preterm babies.

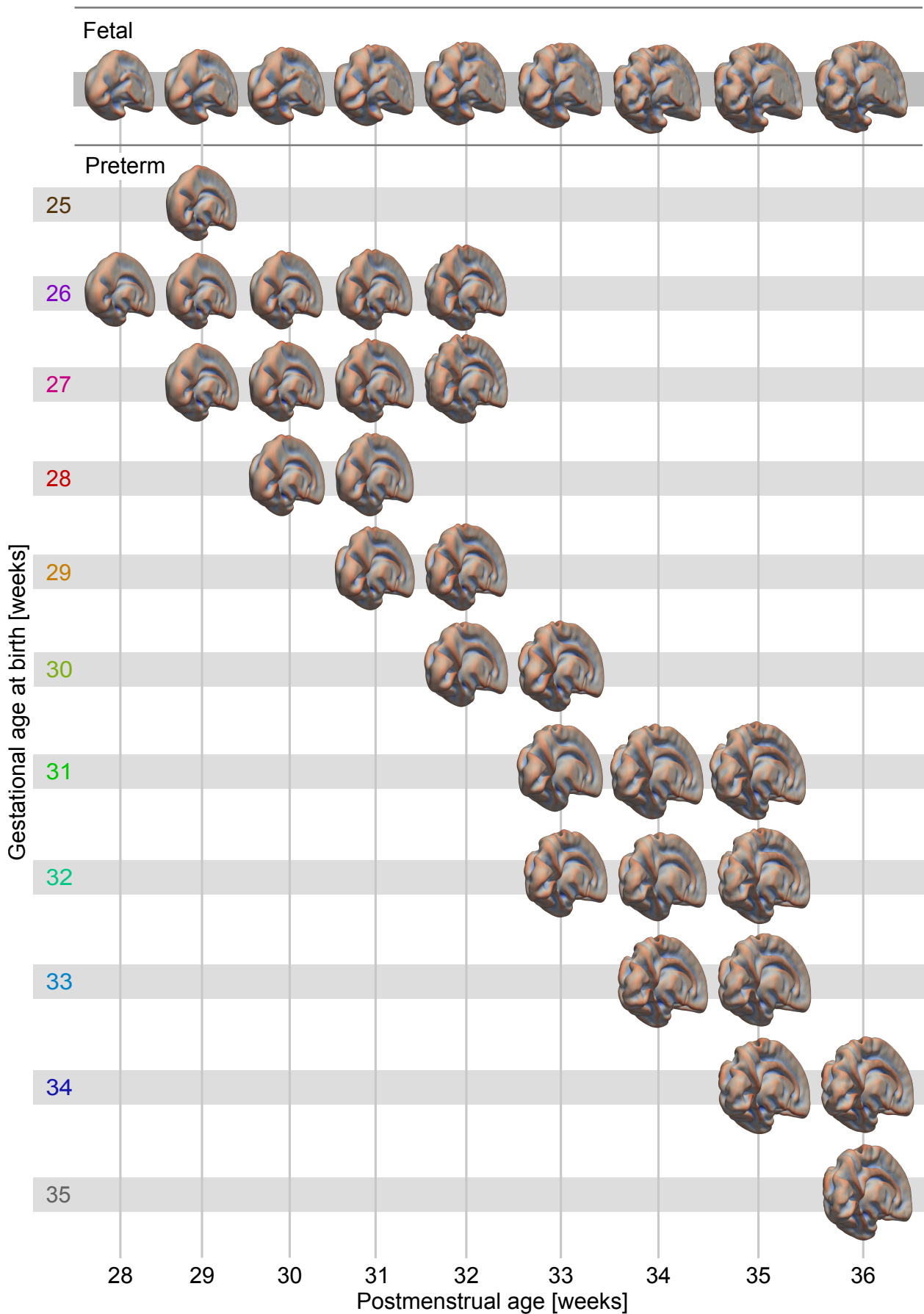


Figure 59: Inferior/medial view, templates for all preterm babies.

A.2 CORTICAL SURFACE DISPLACEMENT MAPS

In this appendix, average cortical surface templates constructed for preterm babies and fetuses are shown (see Chapter 5). Each preterm template has a colour mapping which visualises the displacements for matched locations on its corresponding fetal template, i.e. the template constructed for subjects with the same PMA. Specifically the colour mapping corresponds to the scalar projection of the computed displacements along the surface normal. Thus positive values depicted by hotter colours, show outward displacements relative to the fetal template, while negative values depicted by cooler colours show inwards displacements.

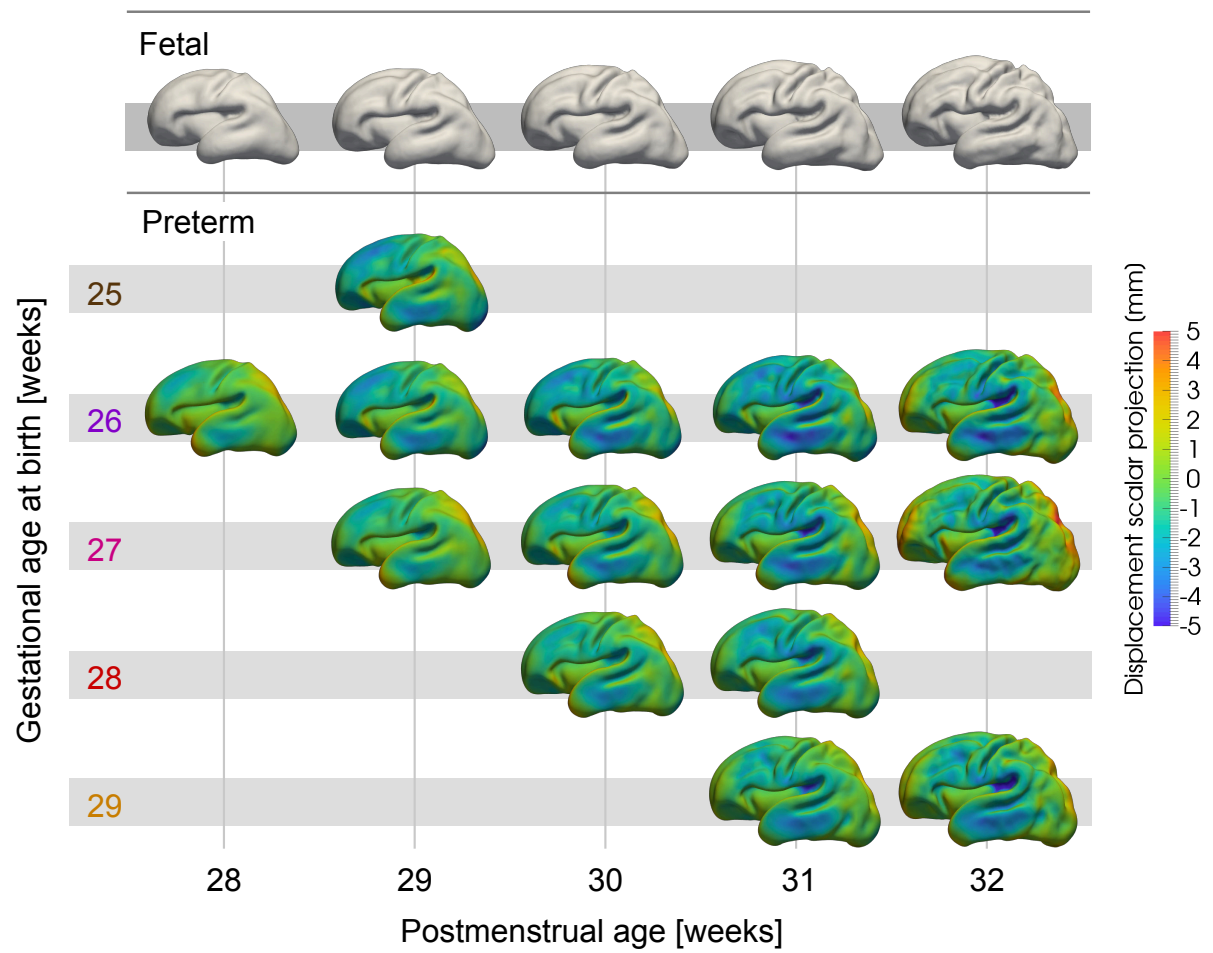


Figure 60: Lateral view, templates for extremely preterm babies (25-29 weeks GA at birth).

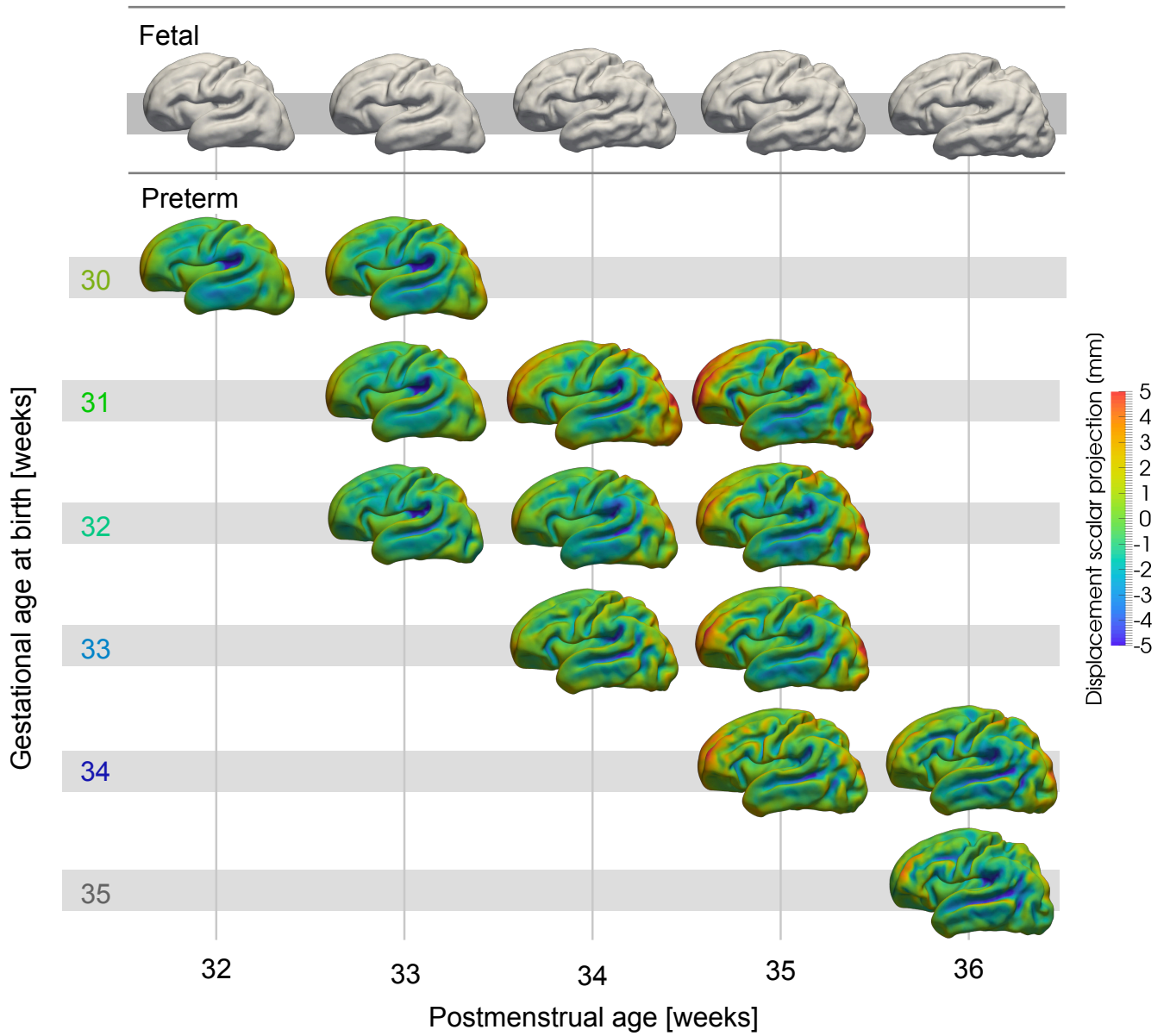


Figure 61: Lateral view, templates for very and moderately preterm babies (30-35 weeks GA at birth).

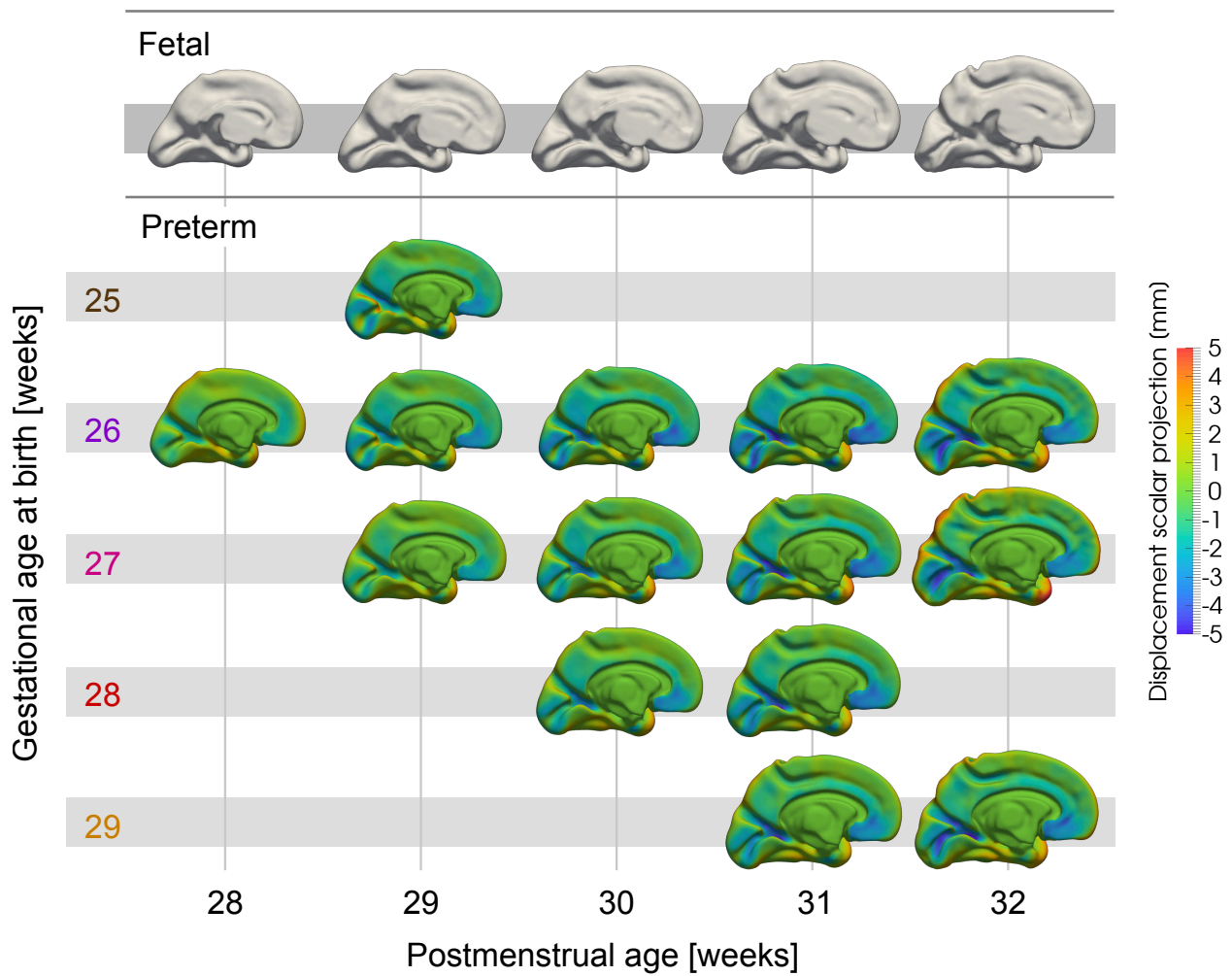


Figure 62: Medial view, templates for extremely preterm babies (25-29 weeks GA at birth).

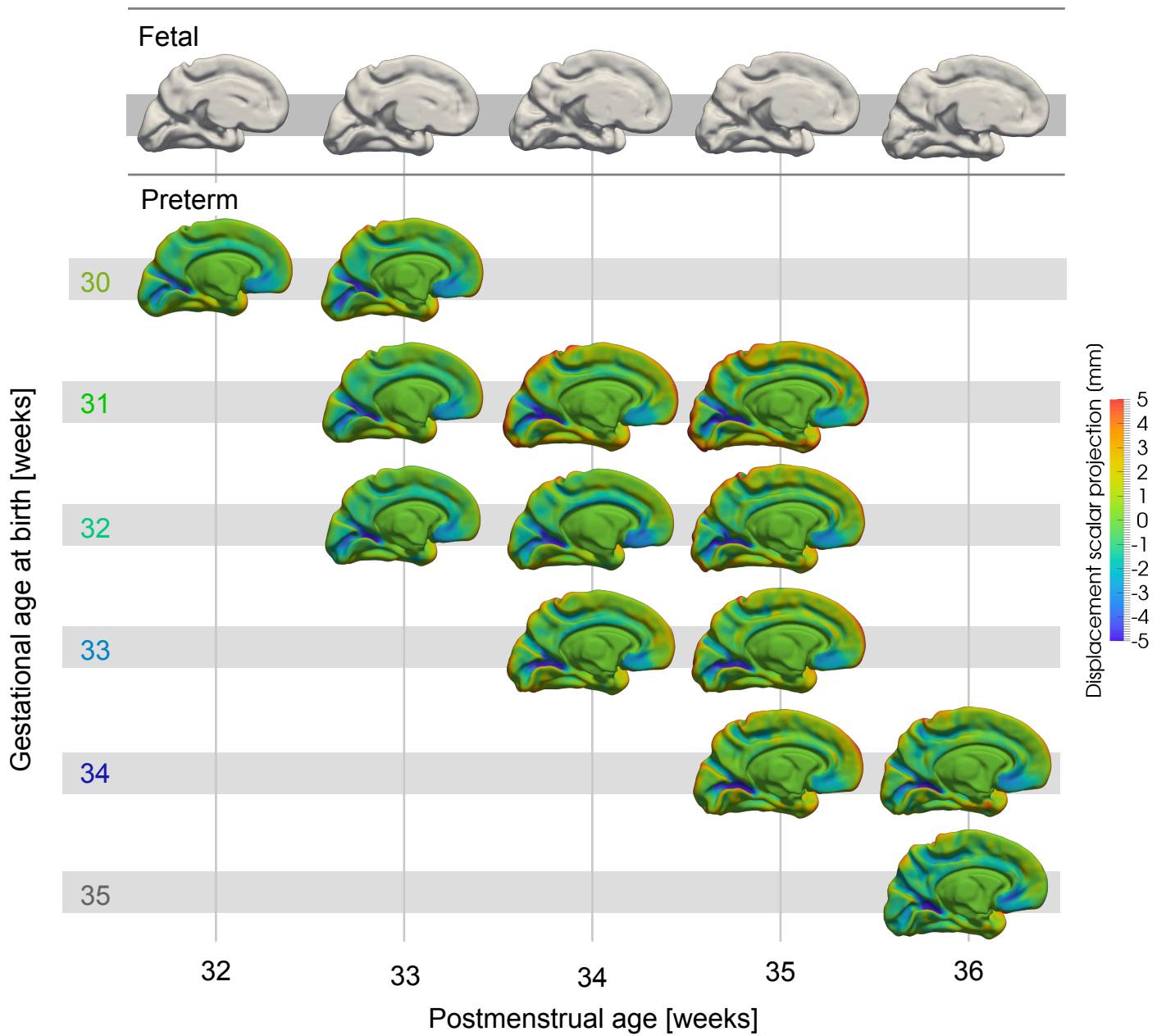


Figure 63: **Medial view, templates for extremely preterm babies (25-29 weeks GA at birth).**

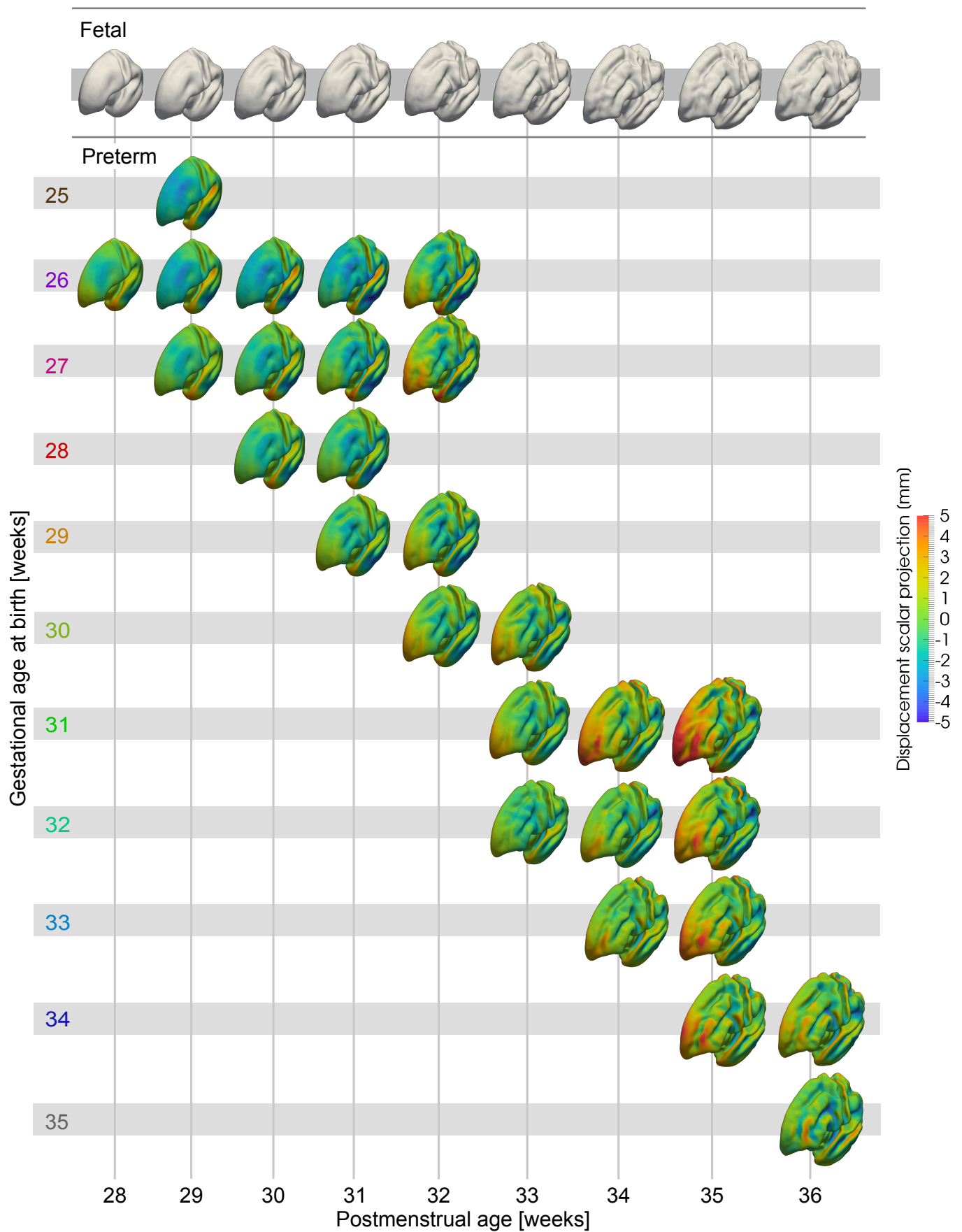


Figure 64: Superior/lateral view, templates for all preterm babies.

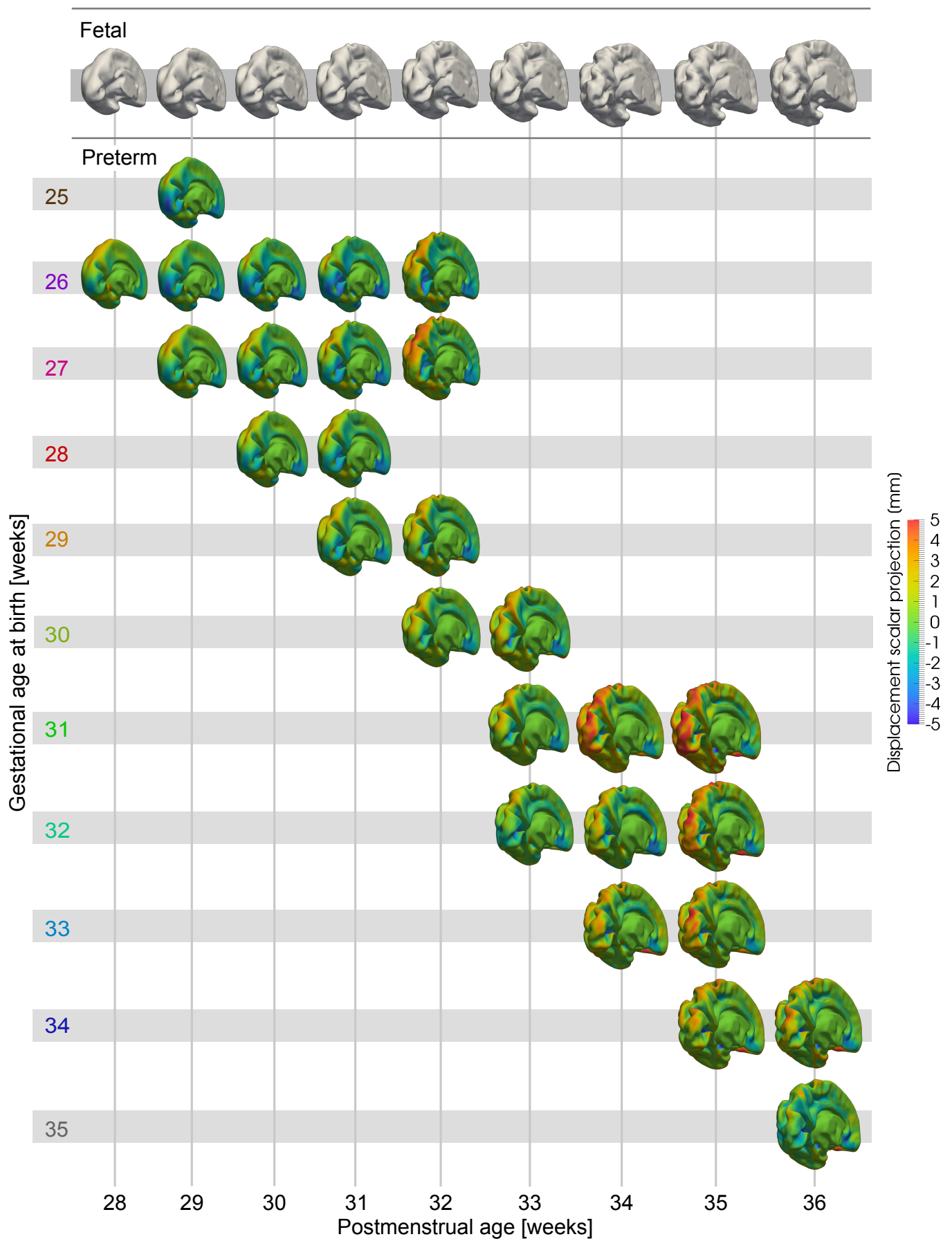


Figure 65: Inferior/medial view, templates for all preterm babies.