

Segmentation of Magnetic Resonance Images for Assessing Neonatal Brain Maturation



Siying Wang

Keble College

University of Oxford

Supervisors

Prof Julia A. Schnabel

Dr Maria Kuklisova-Murgasova

Prof J. Alison Noble

A thesis submitted for the degree of
Doctor of Philosophy in Engineering Science

Michaelmas 2016

Abstract

In this thesis, we aim to investigate the correlation between myelination and the gestational age for preterm infants, with the former being an important developmental process during human brain maturation. Quantification of myelin requires dedicated imaging, but the conventional magnetic resonance images routinely acquired during clinical imaging of neonates carry signatures that are thought to be associated with myelination. This thesis thus focuses on structural segmentation and spatio-temporal modelling of the so-called myelin-like signals on T_2 -weighted scans for early prognostic evaluation of the preterm brain.

The segmentation part poses the major challenges of this task: insufficient spatial prior information of myelination and the presence of substantial partial volume voxels in clinical data. Specific spatial priors for the developing brain are obtained from either probabilistic atlases or manually annotated training images, but none of them currently include myelin as an individual tissue type. This causes further difficulties in partial volume estimation which depends on the probabilistic atlases of the composing pure tissues.

Our key contribution is the development of an expectation-maximisation framework that incorporates an explicit partial volume class whose locations are configured in relation to the composing pure tissues in a predefined region of interest via second-order Markov random fields. This approach resolves the above challenges without requiring any probabilistic atlas of myelin. We also investigate atlas-based whole brain segmentation that generates the binary mask for the region of interest. We then construct a spatio-temporal growth model for myelin-like signals using logistic regression based on the automatic segmentations of 114 preterm infants aged between 29 and 44 gestational weeks. Lastly, we demonstrate the ability of age estimation using the normal growth model in a leave-one-out procedure.

Dedicated to my parents.

Acknowledgements

I would never have been able to complete my doctoral research without the help and support from my supervisors, family, friends and many other people. In particular, I would like to express my thanks to:

My supervisors Julia Schnabel, Maria Kuklisova-Murgasova and Alison Noble for giving me constant academic guidance and personal support, encouraging me when our manuscripts were rejected, tolerating my nature as a perfectionist, and most importantly, not giving me up when I struggled through the darkest days of depression.

Prof Joseph Hajnal at King's College London for authorising access to the dataset and proposing the term "myelin-like signals".

Prof Serena Counsell at King's College London for providing neurological insights and examining the manual annotations.

Dr Christian Ledig at Imperial College London who I have never had an opportunity to meet in person, but kindly helped me with the computational implementation of MRFs in the IRTK software, and always replied promptly with thorough comments on our co-authored papers.

My course mate Liam Cattell for sharing his thesis template that saved me from the painful writing, and taking care of my desk in the lab while I was on suspension.

Mischka Byworth at the Counselling Service and Dr Helen Steel at Jericho Health Centre, both of whom not only provided medical assistance, but also listened to my life stories with great patience.

The Clarendon Fund Scholarship, Department of Engineering Science and Keble College for the academic funding and travel grants over the four years of my DPhil.

Finally, Mum and Dad for loving me as always and teaching me what truly makes us a family.

Contents

List of Tables	vi
List of Figures	vii
List of Abbreviations	xii
1 Introduction	1
1.1 Myelination and human brain maturation	2
1.2 Imaging of the developing brain	4
1.2.1 Cranial ultrasonography	4
1.2.2 Principles of MRI	4
1.2.3 Conventional and advanced MRI techniques	5
1.2.4 Comparison of MRI techniques for assessing myelination	6
1.3 Challenges of neonatal brain MRI segmentation	7
1.4 Research contributions	10
1.5 Thesis overview	11
2 Review of brain MRI segmentation methods	13
2.1 Introduction	13
2.2 EM-based brain segmentation	14
2.2.1 Gaussian mixture model	14
2.2.2 Definition of the EM algorithm	15
2.2.3 Bias field correction	19
2.2.4 Probabilistic brain atlas	20

2.2.4.1	Unified segmentation in SPM	22
2.2.5	Markov random fields	23
2.2.6	Partial volume modelling	26
2.3	Registration-based brain segmentation	29
2.3.1	Image registration overview	29
2.3.2	Multi-atlas registration-based segmentation	31
2.4	Other brain segmentation methods	32
2.4.1	k -means clustering	32
2.4.2	k -nearest neighbour classification	33
2.4.3	Patch-based label propagation	34
2.4.4	Coupled level sets	36
2.5	Quantitative metrics for segmentation evaluation	37
2.5.1	Dice coefficient	37
2.5.2	Hausdorff distance	37
2.6	Developments for neonatal brain segmentation	38
3	Review of spatio-temporal image analysis methods	41
3.1	Introduction	41
3.2	Image registration	42
3.2.1	Similarity measures	42
3.2.2	Affine registration	43
3.2.3	Non-rigid registration using free-form deformations	44
3.2.4	Other deformation models	45
3.2.4.1	Elastic body model	46
3.2.4.2	Viscous fluid flow model	46
3.2.4.3	Diffusion model	47
3.2.4.4	Flows of diffeomorphisms	47
3.2.5	Groupwise registration	48
3.3	Spatial normalisation	49
3.4	Regression models	50

3.4.1	Kernel smoothing	51
3.4.2	Group of sigmoid functions	52
3.4.3	Polynomial functions	53
3.5	Applications of the atlas construction pipeline	54
3.5.1	Morphological age estimation	54
3.5.2	Deformation-based morphometry	55
3.6	Summary	56
4	Atlas-based segmentation of neonatal brain MRI	58
4.1	Introduction	58
4.2	Materials	59
4.2.1	NeoBrainS12 data	59
4.2.2	Segmentation protocol for manual annotations	59
4.2.3	Probabilistic neonatal brain atlas	60
4.3	Method	60
4.3.1	Image pre-processing	63
4.3.2	Unified segmentation in SPM	64
4.3.3	Segmentation refinements	66
4.3.3.1	Connected component labelling	66
4.3.3.2	Segmentation of myelinated white matter	66
4.4	Results	69
4.5	Discussion and conclusion	75
4.5.1	NeoBrainS12 challenge	75
4.5.2	Segmentation protocol for myelinated white matter	76
5	Partial volume modelling via second-order MRFs: Application to segmentation of myelin-like signals	79
5.1	Introduction	79
5.2	Materials	80
5.2.1	Subjects and MR acquisition	80

5.2.2	Segmentation protocol and manual annotations	81
5.3	Method	83
5.3.1	Image pre-processing	83
5.3.2	Purely intensity-based GMM without PV modelling	85
5.3.3	Explicit PV modelling	87
5.3.4	Spatial regularisation via second-order MRFs	88
5.3.5	Implementation	92
5.4	Results and discussion	94
5.4.1	Optimisation of the penalties in the connectivity tensor	95
5.4.2	Evaluation of the MLS segmentation methods	99
5.4.3	Dependence on the initial threshold	107
5.4.4	Comparison to the intra-rater reliability	112
5.4.5	Volume growths of myelin-like signals	112
5.5	Conclusion	116
6	Spatio-temporal analysis for assessing neonatal brain maturation	118
6.1	Introduction	118
6.2	Materials	119
6.3	Analysis pipeline	120
6.3.1	Construction of the spatio-temporal growth model	120
6.3.1.1	Transformation to the common reference space	120
6.3.1.2	Logistic regression	122
6.3.2	Estimation of gestational ages	123
6.4	Results and discussion	124
6.4.1	Jacobian determinants of the FFD transformations	125
6.4.2	Logistic simulation results compared to the known myelination patterns	126
6.4.2.1	Spatio-temporal trajectories of MLS in the thalami and brainstem	126
6.4.2.2	3D parameter maps	129

6.4.3	Comparing the growth models constructed using logistic and kernel-based methods	133
6.4.4	Evaluation of the logistic and kernel regression methods through leave-one-out age estimation	135
6.5	Conclusion	137
7	Conclusion	141
7.1	Summary of contributions	141
7.2	Limitations and future work	143
7.2.1	Short-term goals	143
7.2.2	Long-term outlooks	145
7.3	Conclusion	145
	List of Publications	147
	Bibliography	148

List of Tables

1.1	Outline of the thesis structure	11
3.1	Age-dependent sigmoid functions	53
4.1	Acquisition parameters for three sets of T_1w and T_2w brain MR images of preterm infants provided in the NeoBrainS12 challenge	59
5.1	Segmentation models for myelin-like signals	94
5.2	Optimal values of the penalties t_1 , t_2 and t_3 in different regions of interest determined using a coarse-to-fine grid search method	96
5.3	Average Dice coefficients of myelin-like signals segmented by optimal thresholding and the model-based methods for 16 test subjects aged be- tween 29 and 44 gestational weeks	99
5.4	Average Dice coefficients over eight test subjects obtained by threshold- ing and the model-based segmentation methods with respect to two sets of manual annotations compared to the intra-rater reliability	113
6.1	Myelination timetable of the posterior limbs of the internal capsule, ven- trolateral nuclei, subthalamic nuclei, medial lemnisci, medial longitudinal fasciculi and vestibular nuclei on T_2w images of preterm infants	129
6.2	Root mean squared errors of the gestational ages predicted for 114 preterm infants using the logistic and kernel-based growth models for myelin-like signals in the thalami and brainstem respectively	136
6.3	Root mean squared errors of the gestational ages predicted for 114 preterm infants using a different reference subject at 36 gestational weeks	137

List of Figures

1.1	Myelin-stained specimens obtained from three preterm infants at birth . .	3
1.2	Conventional T_1 w and T_2 w brain MR images for a preterm infant scanned at 42 gestational weeks	6
1.3	Simulated T_2 w images in the brainstem for a preterm infant at 38 gesta- tional weeks before and after downsampling	9
2.1	Gaussian mixture model obtained using the SPM software compared to the normalised histogram of the T_2 w image for an infant at 40 gestational weeks	14
2.2	T_2 w brain templates and the probability maps of grey matter, white matter and cerebrospinal fluid for adult and neonatal brain MRI	21
2.3	Misclassifications due to the partial volume effect in the segmentations of the cortical grey matter and white matter obtained using the SPM software on the T_2 w image for an infant at 40 gestational weeks	26
2.4	Intensity model for partial volume voxels and two composing pure tissues	28
3.1	Groupwise registration for spatial co-alignment of various images in an average reference space as compared to pairwise image registration to- wards a single reference subject	48
3.2	Common features of sigmoid functions	52
4.1	T_2 w brain template and the probability maps of all the tissue classes ex- cept myelinated white matter after amendments at 40 gestational weeks .	61

4.2	Schematic flow of the proposed method for automated segmentation of neonatal brain MRI	62
4.3	T ₂ w axial brain image for a subject at 40 gestational weeks before and after brain masking	63
4.4	Bias-corrected T ₂ w image and the initial SPM segmentations of all the tissue classes except myelinated white matter for a subject in the axial image set scanned at 40 gestational weeks	65
4.5	T ₁ w and T ₂ w axial brain images used to segment myelinated white matter in the brainstem for a subject at 40 gestational weeks	67
4.6	Clusters of myelinated white matter and the background tissue in the brainstem for a subject at 40 gestational weeks obtained by fitting a two-class multivariate Gaussian mixture model based on the T ₁ w and T ₂ w intensities	68
4.7	Automatic brain segmentations obtained using the proposed algorithm for a subject in the axial image set acquired at 40 gestational weeks	70
4.8	Automatic brain segmentations obtained using the proposed algorithm for a subject in the coronal image set acquired at 30 gestational weeks	71
4.9	Automatic brain segmentations obtained using the proposed algorithm for a subject in the coronal image set acquired at 40 gestational weeks	72
4.10	Dice coefficients of the automatic segmentations obtained using different methods in the NeoBrainS12 challenge for eight tissue classes averaged over three preterm infants in each image set	73
4.11	Hausdorff distances of the automatic segmentations obtained using different methods in the NeoBrainS12 challenge for eight tissue classes averaged over three preterm infants in each image set	74
4.12	Manual annotations of the myelinated tissue delineated according to different segmentation protocols for a training subject at 40 gestational weeks	78
5.1	Manual annotations of myelin-like signals in the thalami and brainstem of a subject at 42 gestational weeks	82

5.2	Schematic flow of the proposed segmentation approach for myelin-like signals on T ₂ w neonatal brain images	84
5.3	Two separate regions of interest for assessing myelin-like signals in the perinatal period	85
5.4	3D connectivity tensor for configuring spatial dependencies of myelin-like signals, background and partial volume voxels via second-order MRFs	89
5.5	Empirical effects on the posterior probability maps (a-c) when all the elements of matrix $\mathbf{T}_{PV}$ are set to zero, and (d-f) when all the elements of $\mathbf{T}_{PV}$ are set to zero except $\mathbf{T}_{PV}(PV, PV)$	92
5.6	Average Dice coefficients over 16 test subjects plotted against the test values of each penalty at the coarse and fine levels of grid search in the thalami	97
5.7	Average Dice coefficients over 16 test subjects plotted against the test values of each penalty at the coarse and fine levels of grid search in the brainstem	98
5.8	Dice coefficients of myelin-like signals segmented by optimal thresholding and the model-based methods in the thalami for 16 test subjects aged between 29 and 44 gestational weeks	101
5.9	Dice coefficients of myelin-like signals segmented by optimal thresholding and the model-based methods in the brainstem for 16 test subjects aged between 29 and 44 gestational weeks	102
5.10	Gaussian distributions of the posterior probability maps estimated using the segmentation models in Table 5.1 for myelin-like signals, background and partial volume voxels in the thalami of a subject at 37 gestational weeks	103
5.11	Gaussian distributions of the posterior probability maps estimated using the segmentation models in Table 5.1 for myelin-like signals, background and partial volume voxels in the brainstem of a subject at 37 gestational weeks	104

5.12	Posterior probability maps estimated using the segmentation methods with explicit PV modelling in the thalami of a subject at 37 gestational weeks .	105
5.13	Posterior probability maps estimated using the segmentation methods with explicit PV modelling in the brainstem of a subject at 37 gestational weeks	106
5.14	Average Dice coefficients of myelin-like signals in the thalami over 16 test subjects obtained by thresholding and the model-based methods using a range of initial thresholds	108
5.15	Average Dice coefficients of myelin-like signals in the brainstem over 16 test subjects obtained by thresholding and the model-based methods using a range of initial thresholds	109
5.16	Coronal images of myelin-like signals in the thalami segmented by thresholding and the model-based methods at the 1st, 5th, 10th, 15th and 20th percentile thresholds for a subject at 43 gestational weeks	110
5.17	Axial images of myelin-like signals in the brainstem segmented by thresholding and the model-based methods at the 21st, 25th, 30th, 35th and 40th percentile thresholds for a subject at 34 gestational weeks	111
5.18	Volume growth of the automatic segmentations for myelin-like signals in the thalami	114
5.19	Volume growth of the automatic segmentations for myelin-like signals in the brainstem	115
6.1	Schematic flow of the proposed spatio-temporal analysis	121
6.2	Voxelwise logistic function between the fraction of myelin-like signals in the reference space and gestational age	123
6.3	Minimum values in the individual Jacobian determinant maps of the free-form deformation transformations computed for co-aligning the regions of thalami and brainstem respectively for 114 preterm infants	125
6.4	Logistic simulation of the spatio-temporal trajectory for myelin-like signals in the thalami between 29 and 44 gestational weeks	127

6.5	Logistic simulation of the spatio-temporal trajectory for myelin-like signals in the brainstem between 29 and 44 gestational weeks	128
6.6	3D map for the average myelination age t_i^A in the thalami and brainstem .	130
6.7	3D map for the time interval Δt_i between $0.25\gamma_i$ and $0.75\gamma_i$ in the thalami and brainstem	131
6.8	3D map for the ultimate fraction of myelin-like signals γ_i in the thalami and brainstem	132
6.9	Spatio-temporal trajectories for myelin-like signals in the thalami between 38 and 41 gestational weeks simulated using logistic and kernel regression	134
6.10	Age distribution for 114 preterm infants at the time of scan	134
6.11	Gestational ages predicted for 114 preterm infants and age estimate errors obtained from the logistic and kernel-based growth models for myelin-like signals in the thalami	138
6.12	Gestational ages predicted for 114 preterm infants and age estimate errors obtained from the logistic and kernel-based growth models for myelin-like signals in the brainstem	139

List of Abbreviations

<i>k</i>NN	<i>k</i> -nearest neighbour
BGT	basal ganglia and thalami
BKG	background
BS	brainstem
CLS	coupled level sets
CoGM	cortical grey matter
CRB	cerebellum
CSF	cerebrospinal fluid
CT	computer tomography
DBM	deformation-based morphometry
DC	Dice coefficient
DGM	deep grey matter
DOF	degree of freedom
DSCP	decussation of the superior cerebellar peduncle
DTI	diffusion tensor imaging
ECSF	cerebrospinal fluid in the extracerebral space
EM	expectation-maximisation
FA	fractional anisotropy
FCMC	fuzzy <i>c</i> -means clustering
FFD	free-form deformation
GA	gestational age
GM	grey matter

GMAT	germinal matrix
GMM	Gaussian mixture model
GP	globus pallidus
GRF	Gibbs random field
HD	Hausdorff distance
IC	inferior colliculus
ICM	iterated conditional modes
KMC	<i>k</i> -means clustering
LDDMM	large deformation diffeomorphic metric mapping
LL	lateral lemniscus
MAP	maximum a posteriori
MI	mutual information
ML	medial lemniscus
MLF	medial longitudinal fasciculus
MLS	myelin-like signals
MRF	Markov random field
MRI	magnetic resonance imaging
MST	minimum spanning tree
MTI	magnetisation transfer imaging
MTR	magnetisation transfer ratio
MWF	myelin water fraction
MWI	myelin water imaging
MWM	myelinated white matter
NCC	normalised cross correlation
NMI	normalised mutual information
PDF	probability density function

PET	positron emission tomography
PLIC	posterior limb of the internal capsule
PPM	posterior probability map
PV	partial volume
RF	radio frequency
RMSE	root mean squared error
ROI	region of interest
SCP	superior cerebellar peduncle
SD	standard deviation
SPM	statistical parametric mapping
SSD	sum of squared differences
STN	subthalamic nucleus
TE	echo time
TR	repetition time
US	ultrasonography
UWM	unmyelinated white matter
VEN	ventricles
VLN	ventrolateral nucleus
VN	vestibular nucleus
WM	white matter

Chapter 1

Introduction

A global investigation of the world health organisation (WHO) in 2010 estimated that 14.9 million babies were born prematurely before 37 completed weeks of gestation, and the number is still increasing (Blencowe et al., 2012). The survival rates for preterm births have improved enormously in recent years because of the advances in neonatal intensive care (Saigal and Doyle, 2008). The growing population of preterm survivors is at high risk of neurodevelopmental disabilities such as learning difficulties and behavioural problems at school age (Saigal et al., 2003; Mikkola et al., 2005), despite the fact that most children adjust well during growth (Saigal and Doyle, 2008). The aim of this thesis is to develop a screening tool for detecting abnormal brain development during the perinatal period, that facilitates early medical interventions and reduces long-term morbidity.

Section 1.1 introduces an important maturational process called myelination. We explain various magnetic resonance imaging techniques in Section 1.2, and discuss the challenges of analysing neonatal brain image data in Section 1.3. The research contributions and thesis overview are presented in Sections 1.4 and 1.5. The age of an infant is defined throughout this thesis as the gestational age (GA) estimated from the last normal menstrual period (Savitz et al., 2002).

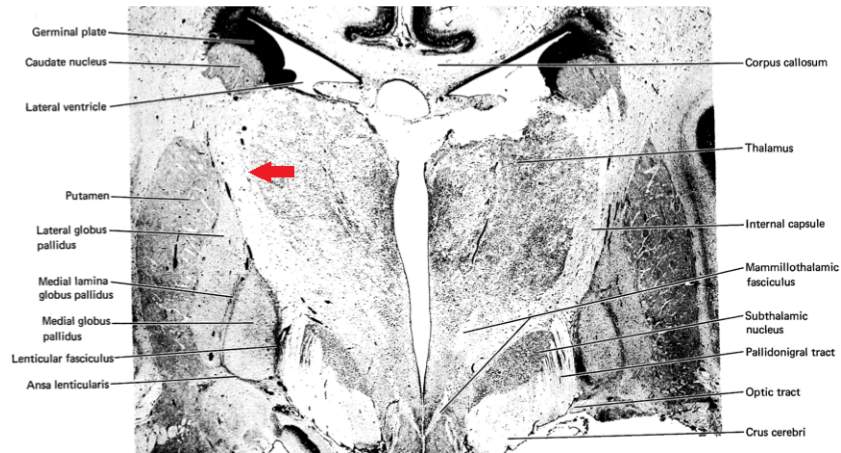
1.1 Myelination and human brain maturation

Myelin has a membranous structure made of lipid bilayers arranged alternately with protein, enveloping axons in the neural system to increase the transmission speed of neural impulses (Laule et al., 2007). The process of myelin growth and axon ensheathment is referred to as myelination.

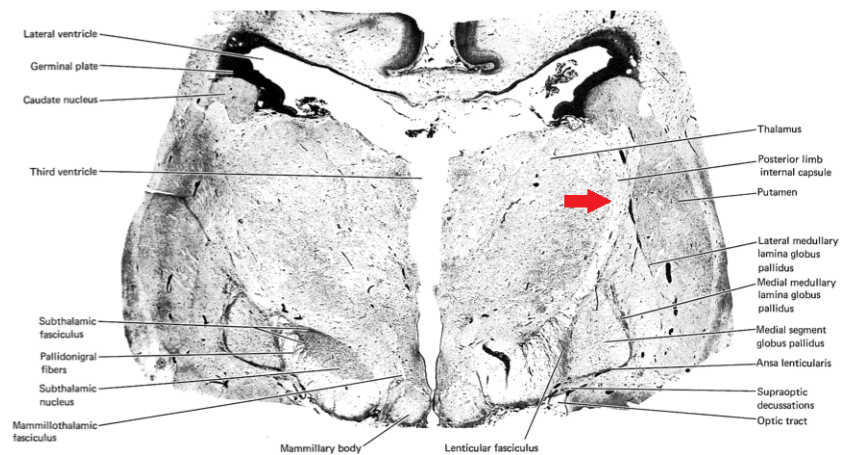
Human brain maturation involves a complex series of morphological, structural and functional changes that occur during the embryonic, fetal and neonatal stages, and continue by one or two years after birth. It has been demonstrated histologically that myelination begins during the second trimester of pregnancy, and proceeds in a systematic fashion from the central to peripheral, from the posterior to anterior, and from the inferior to superior regions of the brain (Yakolev and Lecours, 1967; Rorke and Riggs, 1969; Kinney et al., 1988). Myelin first appears in the posterior brainstem, followed by the anterior brainstem, internal capsules and cerebral hemispheres.

In clinical practice, birth weights are commonly used for measuring gestational maturity. Fig. 1.1 presents the myelin-stained specimens obtained by Rorke and Riggs (1969) from three preterm infants at birth. Infant B shows much less myelin than infant C, and the myelination status appears more analogous to that in infant A, which cannot be fully reflected from their birth weights. In fact, Rorke and Riggs (1969) reported that some well-developed infants based on the measurements of body weight show retarded brain maturation. Moreover, when extensive myelination occurs, the neurodevelopmental patterns in preterm infants are similar to those found in full-term babies, regardless of the birth weights.

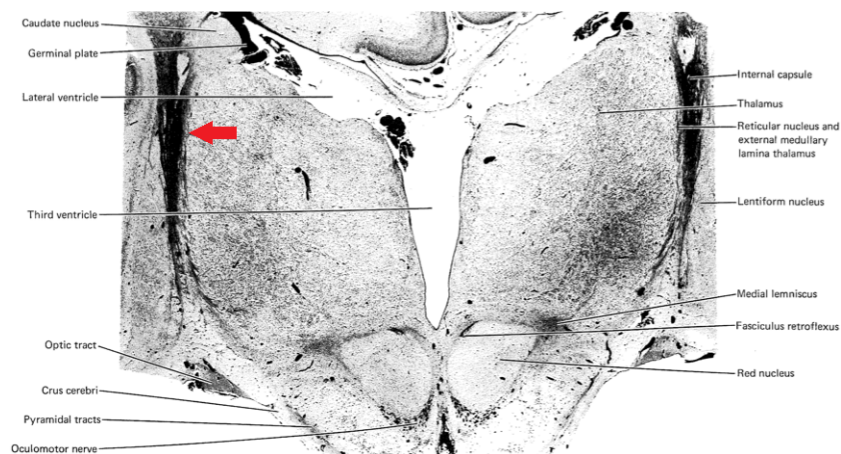
Considering the critical role of myelination in complex cognitive activities and the close relationship with neurodevelopment (Fornari et al., 2007), we propose to use myelin as a potential marker for assessing neonatal brain maturation. However, long-term *in vivo* investigations pose a significant burden for histological studies, and demand the use of non-invasive imaging techniques.



(a) Preterm infant A with birth weight of 1400 g



(b) Preterm infant B with birth weight of 2020 g



(c) Preterm infant C with birth weight of 2260 g

Figure 1.1: Myelin-stained specimens (coronal section through the prosencephalon) obtained by Rorke and Riggs (1969) from three preterm infants at birth. The red arrows point out the posterior limb of the internal capsule. Infant B shows much less myelin than infant C, and the myelination status appears more analogous to that in infant A, which cannot be fully reflected from their birth weights.

1.2 Imaging of the developing brain

Serial assessment of the developing brain requires non-invasive *in vivo* imaging techniques. Both magnetic resonance imaging (MRI) and cranial ultrasonography (US) are commonly used in neonatology due to the non-ionising property in contrast to computer tomography (CT) and positron emission tomography (PET) (O'Shea et al., 2005).

1.2.1 Cranial ultrasonography

US is a portable, relatively inexpensive imaging modality that can be used at the bedside with little disturbance to the infant. An US image is formed by transmitting ultrasound waves through the human body using a transducer array (Rumack et al., 2005). The reflections of the waves at various tissue boundaries produce echoes that can be received and processed to measure the distances between the transducers and the tissue boundaries. For a standard neonatal US protocol, the anterior fontanelle is used as the acoustic window, and scanning is performed with a transducer frequency of 7 MHz or higher (O'Shea et al., 2005).

US for preterm infants enables early detection of the typical brain lesions, such as haemorrhage and ventriculomegaly (De Vries et al., 1998). However, US also has several significant limitations. Because ultrasound waves are poorly transmitted through bones, the US technique is only valid for neuroimaging before the closure of the anterior fontanelle, which usually occurs between 7 and 19 months after birth (Aisenson, 1950). More importantly, in case of suspected developmental delays, myelination is not depicted by US, making MRI more suitable for tracking brain maturation during infancy and childhood (Van Wezel-Meijler et al., 2010).

1.2.2 Principles of MRI

MRI is based on the principles of nuclear magnetic resonance (Brown and Semelka, 2003; Mitchell and Cohen, 2004). Scanners are built with a strong static magnetic field, normally between 1.5 T and 3 T for clinical neuroimaging. When placed in the field, the spins

of hydrogen nuclei in water molecules become aligned, resulting in longitudinal magnetisation parallel to the static field. A radio frequency (RF) pulse at the resonance frequency of hydrogen nuclei is applied to flip the angle of the spins, which creates transverse magnetisation perpendicular to the static field, and further induces electrical currents in the receiver coils. Once the RF pulse is switched off, the relaxation process of the hydrogen nuclei is measured based on two time constants: T_1 for the exponential recovery of the longitudinal magnetisation and T_2 for the exponential decay of the transverse magnetisation. The MR image contrast is provided by the tissue-specific T_1 and T_2 relaxation times. Spatial information is encoded by applying additional magnetic field gradients in three-dimensional (3D) space.

1.2.3 Conventional and advanced MRI techniques

Many MRI techniques have been developed to acquire images using specially designed pulse sequences. Conventional T_1 -weighted (T_1w) images are obtained using short repetition times (TRs) and short echo times (TEs). The contrast is predominantly caused by the differences in tissue T_1 values. T_2w images are acquired using long TRs and long TEs. As shown in Fig. 1.2, myelin may be recognised from the high intensities on T_1w neonatal brain MR images, which is reversed in the T_2w contrast (Barkovich, 1998). This is because T_1 and T_2 relaxation times are shortened due to the high lipid content in myelin and the ensheathment around axons (Rutherford, 2002).

While the conventional T_2w MRI uses a single TE, myelin water imaging (MWI) via T_2 relaxometry collects images at multiple TEs to produce a T_2 decay curve (Alonso-Ortiz et al., 2015). Apart from the cerebrospinal fluid (CSF), two peaks can be identified in the T_2 spectrum: a long component near 100 ms associated with intracellular and extracellular water and a short component near 20 ms arising from the so-called “myelin water” trapped within the myelin sheath (Laule et al., 2007). The area under the spectrum between 10 ms and 50 ms due to the myelin water is divided by the total area of the spectrum to calculate the myelin water fraction (MWF).

Magnetisation transfer imaging (MTI) also uses a two-pool model, and is based on the

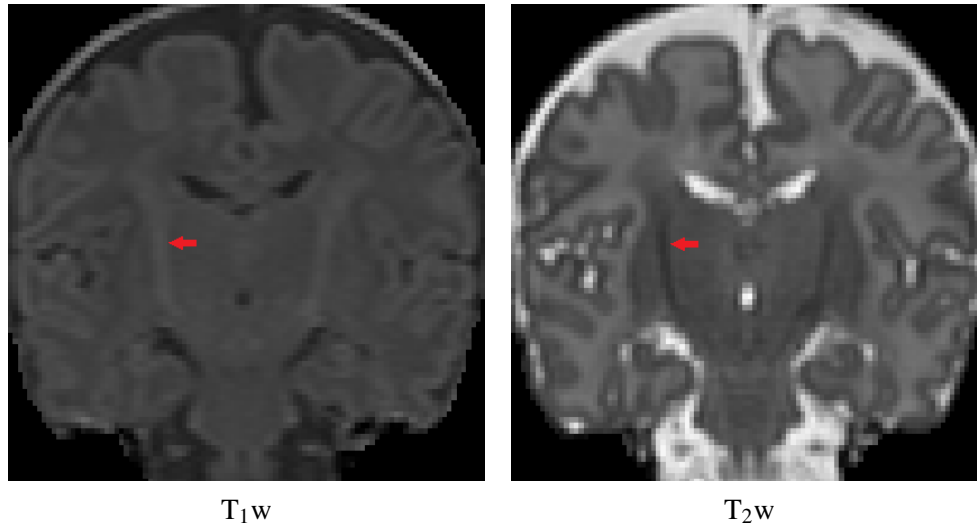


Figure 1.2: Conventional T_1w and T_2w brain MR images (coronal view) for a preterm infant scanned at 42 gestational weeks. The myelinated posterior limb of the internal capsule can be recognised from the high intensities relative to the thalami on the T_1w image, and from the low intensities on the T_2w image, as pointed out by the arrows.

exchange of magnetisation between a pool of free water and a pool of water molecules that are bound to macromolecules, such as the proteins and lipids of myelin (Henkelman et al., 2001). An off-resonance RF pulse is applied to saturate the macromolecular spins, and the magnetisation transfer ratio (MTR) is measured as the percentage loss of signal intensities due to the magnetisation transfer between the two pools (Van Buchem et al., 2001).

Diffusion tensor imaging (DTI) measures the magnitude and direction of water diffusion, and allows tissue microstructures, such as cell membranes, fibers and myelin, to be investigated non-invasively (Le Bihan et al., 2001). Basser et al. (1994) introduced a diffusion tensor to describe 3D molecular mobility in the presence of anisotropy. A minimum of six diffusion-weighted images are collected in non-collinear directions to compute the maps of fractional anisotropy (FA) and various diffusivities.

1.2.4 Comparison of MRI techniques for assessing myelination

T_1 relaxation as well as the MTR and FA obtained from MTI and DTI respectively have been found sensitive to the arrival of macromolecules and early fiber fasciculation in the absence of myelin at the pre-myelination stage (Wozniak and Lim, 2006; Nossin-

Manor et al., 2013; Dubois et al., 2014). T_2 relaxation and hence the MWM obtained via MWI are demonstrated to correlate best with the myelination process determined using histological methods (Rutherford, 2002; Laule et al., 2007). We therefore use T_2 w images to investigate myelination in the preterm brain in this thesis.

A drawback of choosing the conventional MRI is that the signal properties reflect several maturational processes that occur simultaneously, including chemical composition changes, increased densities of neurons and synapses as well as increased myelination (Korogi et al., 1996). On the other hand, T_1 w and T_2 w acquisitions remain the most routine MRI techniques compared to the advanced MWI, MTI and DTI which require elaborate scans that are constrained in clinical practice when imaging fragile preterm neonates (Rutherford et al., 2006; Srinivasan and Rutherford, 2008; Dubois et al., 2014).

Qualitative descriptions of myelination on T_2 w images (Barkovich, 1998; Counsell et al., 2002) are in great consistency with the histological observations (Yakolev and Lecours, 1967; Kinney et al., 1988). Thus the ability to numerically assess myelination using the conventional T_2 w scans would facilitate the automatic evaluation of brain maturation in preterm infants directly from clinical MRI data in large-scale studies.

1.3 Challenges of neonatal brain MRI segmentation

T_2 w neonatal brain MRI shows the tissue that is likely to contain myelin, which we will refer to as myelin-like signals (MLS) in this thesis. Automatic delineation of MLS is not straightforward. Below we explain the challenges of general neonatal brain segmentation as well as the additional complexities for segmenting MLS.

MR images of the neonatal brain have a much lower contrast-to-noise ratio than the adult brain images (Makropoulos et al., 2014). This is because the immature brain with higher water content causes a substantial increase in the T_1 and T_2 values, and lowers the tissue contrast (Huppi and Inder, 2001; Rutherford, 2002). The newborn brain grows rapidly, whose size and shape differ significantly from the adult brain. Another confounding issue is the partial volume (PV) effect when multiple tissues are present in a single voxel due to the small head size and low image resolution of neonatal brain MRI

(Kuklisova-Murgasova et al., 2007). For instance, the PV voxels containing both the cortical grey matter (CoGM) and CSF at the CoGM/CSF boundary are frequently mislabelled as white matter (WM) due to similar intensities (Xue et al., 2007; Kuklisova-Murgasova et al., 2011). As a result, most neonatal brain segmentation algorithms use a probabilistic atlas or a database of manual annotations to initialise intensity distributions, and more importantly, to obtain spatial prior information on the expected tissue locations. We will demonstrate this point in Chapter 2.

MLS segmentation is even more challenging for the following reasons:

- The existing probabilistic neonatal brain atlases (Kuklisova-Murgasova et al., 2011; Serag et al., 2012) do not include an individual probability map for MLS, and there is no available database of MLS manual annotations.
- The PV voxels containing both MLS and the background (BKG) tissue in a region of interest (ROI) need to be modelled explicitly when segmenting MLS (see Fig. 1.3). The current PV estimation methods depend on prior knowledge of the spatial locations of the composing pure tissues to guide the search for PV voxels at tissue boundaries (Shattuck et al., 2001; Van Leemput et al., 2003; Tohka et al., 2004; Manjon et al., 2010; Cardoso et al., 2011; Ledig et al., 2012). However, there is very limited spatial prior information available for MLS due to the absence of any probabilistic atlas or manual annotation database.

Fig. 1.3(a) shows an artificial T_2w image in the brainstem for a preterm infant at 38 weeks GA. We labelled MLS based on the manual annotation and generated the intensities of MLS and BKG from two separate Gaussians with random noise. It can be seen from Figs. 1.3(b) and 1.3(c) that a commonly used two-class Gaussian mixture model (GMM) correctly recovers the distributions with respect to the histograms of the simulated image. In Fig. 1.3(d), we downsampled the image to deliberately create PV voxels. Figs. 1.3(e) and 1.3(f) demonstrate that the PV voxels shift the histogram of manually annotated MLS towards BKG, and the simple GMM without PV estimation can no longer approximate the histograms accurately. Therefore, explicit PV modelling is vital for MLS segmentation.

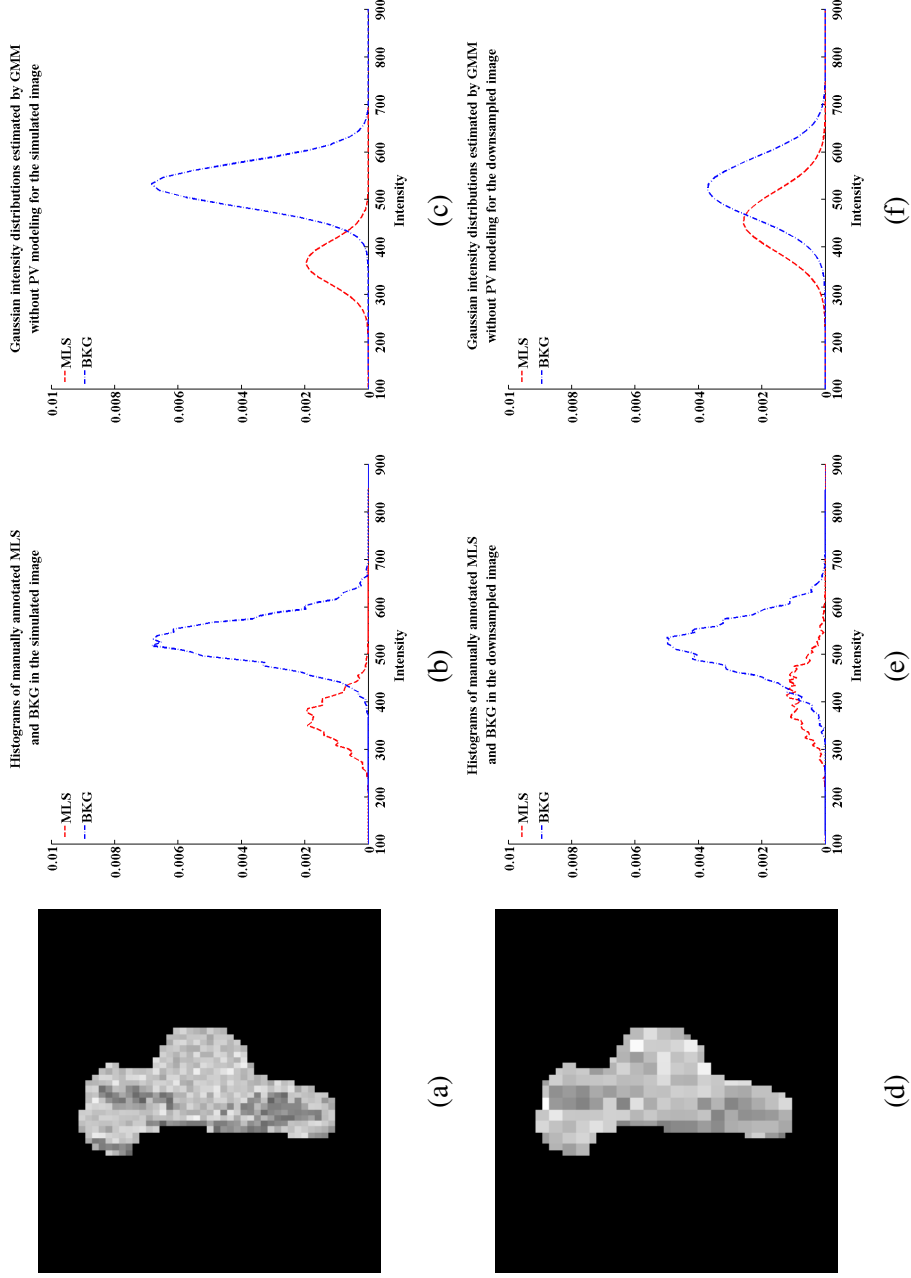


Figure 1.3: Simulated T_2w images in the brainstem (sagittal view) for a preterm infant at 38 gestational weeks (a) before and (d) after downsampling. The intensities of myelin-like signals (MLS) and the remaining background (BKG) are generated from two Gaussians, and the MLS voxels are labelled based on the manual annotation. A two-class Gaussian mixture model (GMM) recovers the distributions correctly before downsampling. Partial volume (PV) voxels created after downsampling shift the histogram of manually annotated MLS towards BKG, and the simple GMM without PV estimation can no longer approximate the histograms accurately.

1.4 Research contributions

The overall contribution of this thesis is the development of an image analysis pipeline based on the characteristic MLS on T₂w neonatal brain MR images, so that brain maturation can be assessed automatically and directly using clinical MRI data.

The myelination process has been investigated extensively using MWI, MTI and DTI in research work (Dean III et al., 2014; Deoni et al., 2015; Nossin-Manor et al., 2015). However, these advanced MRI techniques use dedicated pulse sequences that are constrained during clinical imaging of neonates (Rutherford et al., 2006; Dubois et al., 2014). To the best of our knowledge, no such investigations have been carried out using the conventional T₁w and T₂w acquisitions that are routinely performed in clinical practice.

The main contributions include:

- 1) We propose an atlas-based neonatal brain MRI segmentation method that adapts the segmentation tool in the Statistical Parametric Mapping (SPM) software¹ (Ashburner and Friston, 2005) to a probabilistic neonatal brain atlas² (Kuklisova-Murgasova et al., 2011). This method achieves brain extraction, bias field correction as well as whole brain segmentation.

- 2) We develop a novel segmentation method for MLS in a predefined ROI on T₂w images. We first obtain the ROI using the previously proposed atlas-based method. The subsequent MLS segmentation in the ROI is atlas-free. We introduce an explicit PV class to the standard two-class GMM, and generate the spatial priors by configuring the PV locations in relation to the composing pure tissues via second-order Markov random fields (MRFs). Our approach of PV estimation achieves robust MLS segmentations, and the independence from an MLS atlas is important such that the segmentations are not biased towards the atlas spatial prior, which would prevent accurate age estimation as explained next.

- 3) We construct an MLS growth model using the automatic segmentations obtained from 114 preterm infants at different GAs to simulate the myelination progress in the preterm

¹<http://www.fil.ion.ucl.ac.uk/spm>

²<http://brain-development.org/brain-atlases>

brain during the perinatal period. We re-estimate the GA of an individual subject by calculating the deviation between its MLS segmentation and the average growth model. Patients with abnormal MLS development compared to the majority of the preterm population can be identified.

1.5 Thesis overview

The thesis structure is outlined in Table 1.1, followed by a brief overview of the individual chapters.

Chapter 1	Introduction and background information
Chapters 2 and 3	Literature review on brain MRI segmentation and spatio-temporal modelling
Chapters 4, 5 and 6	Research contributions, including descriptions of the datasets, methodologies, evaluation experiments and results
Chapter 7	Limitations, future work and conclusion

Table 1.1: Outline of the thesis structure.

Chapter 1 provides the background information for the later chapters. We first describe the myelination process and its critical role in brain maturation. We then analyse the advantages and challenges of assessing myelination using T₂w neonatal brain MRI, followed by a summary of our research contributions.

Chapter 2 presents a comprehensive review of brain MRI segmentation methods developed for both adults and neonates. We describe in detail a probabilistic expectation-maximisation (EM) framework for tissue classification, and briefly introduce some alternative approaches. The majority of the existing brain segmentation methods, especially those for the developing brain, require a probabilistic atlas or a large number of manual annotations to obtain spatial prior information. We will derive our own methods in Chapters 4 and 5 on the basis of the EM approach.

Chapter 3 describes an image analysis pipeline that has been established in the literature for constructing spatio-temporal brain atlases. The pipeline consists of image reg-

istration, spatial normalisation and application of voxelwise regression models. Relevant techniques will be explored in Chapter 6 to build an MLS growth model.

Chapter 4 proposes an atlas-based whole brain segmentation method dedicated to neonatal MRI data for participation in the NeoBrainS12 challenge (Isgum et al., 2015). We segment myelinated white matter (MWM) according to the challenge segmentation protocol using a multivariate GMM. However, MWM segmentation is significantly underperformed compared to the other tissue types across all the algorithms in the challenge, which leads to our investigation in Chapter 5.

Chapter 5 introduces the term MLS for which we develop a novel segmentation method on T_2w images. We first define an MLS segmentation protocol for delineating manual annotations. We then improve the attempt in Chapter 4 by adding an explicit PV class in the GMM, and incorporate spatial regularisation via second-order MRFs. We conduct a series of experiments to demonstrate the accuracy and robustness of the proposed method against different initialisations over a wide age range in the absence of any MLS atlas.

Chapter 6 constructs a spatio-temporal growth model based on the MLS segmentations obtained from Chapter 5. We further investigate the potential of age estimation using the average growth model, and notice that the logistic regression method better predicts the GAs than commonly used fixed-width kernel smoothing that introduces additional errors due to the non-uniform distribution of subjects at different ages.

Chapter 7 concludes this thesis by discussing the limitations of our work, and by outlining the directions for future research.

Chapter 2

Review of brain MRI segmentation methods

2.1 Introduction

In Chapter 1, we explained the advantages and challenges of neonatal brain segmentation using the conventional T_1w and T_2w acquisitions compared to other advanced MRI techniques. In this chapter, we review the segmentation algorithms developed for adult and neonatal brain MRI. These algorithms play a fundamental part in various neuroimaging studies, such as predicting neurodevelopmental outcomes (Young et al., 2015), diagnosis in Alzheimer’s disease (Kloppel et al., 2008) and construction of a spatio-temporal brain atlas (Kuklisova-Murgasova et al., 2011).

The chapter is organised as follows. Section 2.2 describes a probabilistic EM framework based on a Gaussian mixture model (GMM). Bias field correction, probabilistic brain atlas, MRF neighbourhood information and partial volume (PV) estimation are further incorporated in the EM framework to eliminate the influences of MR imaging artefacts and enhance segmentation robustness. Section 2.3 introduces a different registration-based approach that involves propagation and fusion of multiple labels via deformable image registration. In Section 2.4, we briefly explain some alternative brain segmentation techniques, such as k -means clustering, k -nearest neighbour classification, patch-based la-

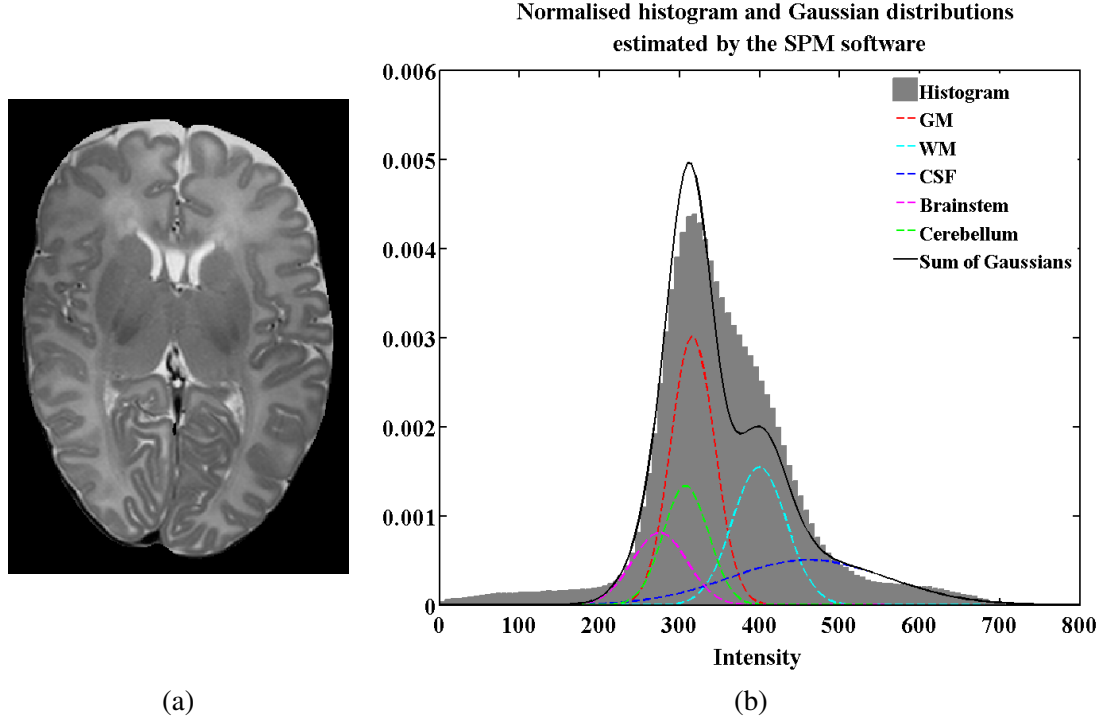


Figure 2.1: (a) T₂w axial brain image for an infant at 40 gestational weeks. The non-brain tissues have been removed. (b) Gaussian mixture model obtained using the Statistical Parametric Mapping (SPM) software compared to the normalised histogram of the T₂w image. Abbreviations: GM—grey matter in the cortex, basal ganglia and thalami, WM—white matter, CSF—cerebrospinal fluid.

bel propagation and coupled level sets. Quantitative metrics for evaluation are described in Section 2.5. Lastly, we discuss the application of these adult brain segmentation methods to neonatal brain MRI in Section 2.6.

2.2 EM-based brain segmentation

2.2.1 Gaussian mixture model

A great number of segmentation algorithms for both adult and neonatal brain MRI are derived on the basis of a GMM (Wells et al., 1996; Van Leemput et al., 1999a; Van Leemput et al., 1999b; Zhang et al., 2001; Ashburner and Friston, 2005; Cardoso et al., 2011; Makropoulos et al., 2014), as illustrated in Fig. 2.1.

An MR image of N voxels can be represented using a parameterised model. The observed intensities $\mathbf{y} = \{y_1, y_2, \dots, y_N\}$ are generated by the hidden labels $\mathbf{z} = \{z_1, z_2, \dots, z_N\}$

with a conditional probability density function (PDF) $p(\mathbf{y} | \mathbf{z}, \Phi_y)$ parameterised by Φ_y . Assuming there are K tissue classes, $\mathbf{z}_i = \mathbf{e}_k$ indicates to which class voxel i belongs and $\mathbf{e}_k$ is a unit vector of length K whose components are equal to zero except the k th component. Using Bayes' theorem, the joint probability of observing intensity y_i at voxel i of class k is:

$$p(y_i, \mathbf{z}_i = \mathbf{e}_k | \Phi_y) = p(y_i | \mathbf{z}_i = \mathbf{e}_k, \Phi_y) p(\mathbf{z}_i = \mathbf{e}_k | \Phi_y) \quad (2.1)$$

Thus, the total probability of observing intensity y_i at voxel i is:

$$p(y_i | \Phi_y) = \sum_{k=1}^K p(y_i | \mathbf{z}_i = \mathbf{e}_k, \Phi_y) p(\mathbf{z}_i = \mathbf{e}_k | \Phi_y) \quad (2.2)$$

where $p(y_i | \mathbf{z}_i = \mathbf{e}_k, \Phi_y)$ is the conditional PDF, and $p(\mathbf{z}_i = \mathbf{e}_k | \Phi_y)$ the prior probability that voxel i belongs to class k .

For the case of a GMM, the conditional PDF for each tissue class is a Gaussian function with mean μ_k and standard deviation (SD) σ_k , and the prior probability is modelled by the spatially constant mixing proportion c_k of class k :

$$p(y_i | \mathbf{z}_i = \mathbf{e}_k, \Phi_y) = G(y_i, \mu_k, \sigma_k) = \frac{1}{\sqrt{2\pi\sigma_k^2}} e^{-\frac{(y_i - \mu_k)^2}{2\sigma_k^2}}$$

$$p(\mathbf{z}_i = \mathbf{e}_k | \Phi_y) = c_k$$

In the next section, we will estimate the underlying hidden labels $\mathbf{z}$ from the observed intensities $\mathbf{y}$ using the EM algorithm proposed by Dempster et al. (1977).

2.2.2 Definition of the EM algorithm

In order to estimate the hidden labels from the observed intensities, Wells et al. (1996) first employed the EM algorithm (Dempster et al., 1977) for brain segmentation:

- E-step: Calculate the posterior probability $p_{ik}^{(m+1)}$ that voxel i belongs to class k based on its intensity y_i and the current parameter estimates $\Phi_y^{(m)}$. According to

Bayes' rule,

$$p_{ik}^{(m+1)} = p(\mathbf{z}_i = \mathbf{e}_k \mid y_i, \Phi_y^{(m)}) = \frac{p(y_i, \mathbf{z}_i = \mathbf{e}_k \mid \Phi_y^{(m)})}{p(y_i \mid \Phi_y^{(m)})} \quad (2.3)$$

The symbol m is the current EM iteration number. By substituting Eqs. 2.1 and 2.2 into Eq. 2.3, we obtain:

$$\begin{aligned} p_{ik}^{(m+1)} &= \frac{p(y_i \mid \mathbf{z}_i = \mathbf{e}_k, \Phi_y^{(m)}) p(\mathbf{z}_i = \mathbf{e}_k \mid \Phi_y^{(m)})}{\sum_{k'=1}^K p(y_i \mid \mathbf{z}_i = \mathbf{e}_{k'}, \Phi_y^{(m)}) p(\mathbf{z}_i = \mathbf{e}_{k'} \mid \Phi_y^{(m)})} \\ &= \frac{G(y_i, \mu_k^{(m)}, \sigma_k^{(m)}) c_k^{(m)}}{\sum_{k'=1}^K G(y_i, \mu_{k'}^{(m)}, \sigma_{k'}^{(m)}) c_{k'}^{(m)}} \end{aligned} \quad (2.4)$$

- M-step: Update the mean μ_k , SD σ_k and mixing proportion c_k of class k with the following equations:

$$\mu_k^{(m+1)} = \frac{\sum_{i=1}^N p_{ik}^{(m+1)} y_i}{\sum_{i=1}^N p_{ik}^{(m+1)}} \quad (2.5)$$

$$\left(\sigma_k^{(m+1)}\right)^2 = \frac{\sum_{i=1}^N p_{ik}^{(m+1)} \left(y_i - \mu_k^{(m+1)}\right)^2}{\sum_{i=1}^N p_{ik}^{(m+1)}} \quad (2.6)$$

$$c_k^{(m+1)} = \frac{1}{N} \sum_{i=1}^N p_{ik}^{(m+1)} \quad (2.7)$$

where N is the number of voxels in the image.

Eqs. 2.5, 2.6 and 2.7 can be derived using the general definition of the EM algorithm (Dempster et al., 1977):

- E-step: Calculate the function $Q(\Phi_y \mid \Phi_y^{(m)})$ which is the expected value of the log likelihood $L_c(\Phi_y)$ for the complete data $\mathbf{q} = \{\mathbf{y}, \mathbf{z}\}$ based on the observed intensities $\mathbf{y}$ and the current parameter estimates $\Phi_y^{(m)}$:

$$Q(\Phi_y \mid \Phi_y^{(m)}) = E \left[L_c(\Phi_y) \mid \mathbf{y}, \Phi_y^{(m)} \right] \quad (2.8)$$

where

$$L_c(\Phi_y) = \ln p(\mathbf{y}, \mathbf{z} \mid \Phi_y)$$

- M-step: Obtain new estimates $\Phi_y^{(m+1)}$ that maximise $Q(\Phi_y | \Phi_y^{(m)})$:

$$\Phi_y^{(m+1)} = \arg \max_{\Phi_y} Q(\Phi_y | \Phi_y^{(m)})$$

We assume that all the image voxels are independent in the GMM, an assumption that we will revisit in Section 2.2.5 when we incorporate neighbourhood information, and write $L_c(\Phi_y)$ as:

$$L_c(\Phi_y) = \ln p(\mathbf{y}, \mathbf{z} | \Phi_y) = \sum_{i=1}^N \ln p(y_i, \mathbf{z}_i | \Phi_y) \quad (2.9)$$

Since $\mathbf{z}_i$ is a $K \times 1$ unit vector whose components are all equal to zero except the k th component, Eq. 2.9 can be transformed into matrix notation:

$$L_c(\Phi_y) = \sum_{i=1}^N \ln p(y_i, \mathbf{z}_i | \Phi_y) = \sum_{i=1}^N \mathbf{z}_i^T \mathbf{V}(y_i | \Phi_y) \quad (2.10)$$

where

$$\mathbf{V}(y_i | \Phi_y) = \begin{bmatrix} \ln p(y_i, \mathbf{z}_i = \mathbf{e}_1 | \Phi_y) \\ \ln p(y_i, \mathbf{z}_i = \mathbf{e}_2 | \Phi_y) \\ \vdots \\ \ln p(y_i, \mathbf{z}_i = \mathbf{e}_K | \Phi_y) \end{bmatrix} \quad (2.11)$$

Before we proceed, let us first define the conditional expected value:

$$E[f(A) | B] = \sum_a f(a) p(A = a | B)$$

with the following properties:

$$E[A + B | C] = E[A | C] + E[B | C]$$

$$E[dA | C] = dE[A | C]$$

Here A , B and C represent random variables, and d is a constant. Therefore, we rearrange $Q(\Phi_y | \Phi_y^{(m)})$ by substituting Eq. 2.10 into Eq. 2.8:

$$\begin{aligned} Q(\Phi_y | \Phi_y^{(m)}) &= E \left[\sum_{i=1}^N \mathbf{z}_i^T \mathbf{V}(y_i | \Phi_y) | \mathbf{y}, \Phi_y^{(m)} \right] \\ &= \sum_{i=1}^N E \left[\mathbf{z}_i | \mathbf{y}, \Phi_y^{(m)} \right]^T \mathbf{V}(y_i | \Phi_y) \end{aligned} \quad (2.12)$$

According to the definition of the conditional expected value, $E \left[\mathbf{z}_i | \mathbf{y}, \Phi_y^{(m)} \right]$ can be explicitly expressed as:

$$E \left[\mathbf{z}_i | \mathbf{y}, \Phi_y^{(m)} \right] = \sum_{l=1}^K \mathbf{e}_l p(\mathbf{z}_i = \mathbf{e}_l | \mathbf{y}, \Phi_y^{(m)}) \quad (2.13)$$

where $p(\mathbf{z}_i = \mathbf{e}_l | \mathbf{y}, \Phi_y^{(m)})$, when denoted as $p_{il}^{(m+1)}$, is the posterior probability calculated in the E-step of the simplified EM algorithm. Finally, we substitute Eqs. 2.13 and 2.11 into Eq. 2.12, and derive the following objective function for brain segmentation:

$$\begin{aligned} Q(\Phi_y | \Phi_y^{(m)}) &= \sum_{i=1}^N \sum_{l=1}^K p_{il}^{(m+1)} \ln p(y_i, \mathbf{z}_i = \mathbf{e}_l | \Phi_y) \\ &= \sum_{i=1}^N \sum_{l=1}^K p_{il}^{(m+1)} \ln p(y_i | \mathbf{z}_i = \mathbf{e}_l, \Phi_y) + \sum_{i=1}^N \sum_{l=1}^K p_{il}^{(m+1)} \ln p(\mathbf{z}_i = \mathbf{e}_l | \Phi_y) \end{aligned}$$

For the GMM, we have:

$$Q(\Phi_y | \Phi_y^{(m)}) = \sum_{i=1}^N \sum_{l=1}^K p_{il}^{(m+1)} \ln G(y_i, \mu_l, \sigma_l) + \sum_{i=1}^N \sum_{l=1}^K p_{il}^{(m+1)} \ln c_l \quad (2.14)$$

where

$$\ln G(y_i, \mu_l, \sigma_l) = -\frac{1}{2} \ln(2\pi\sigma_l^2) - \frac{(y_i - \mu_l)^2}{2\sigma_l^2}$$

Maximisation of the objective function $Q(\Phi_y | \Phi_y^{(m)})$ in Eq. 2.14 with respect to mean μ_k , SD σ_k^2 and mixing proportion c_k of class k yields Eqs. 2.5, 2.6 and 2.7 for updating the GMM parameters in the M-step of the simplified EM algorithm.

2.2.3 Bias field correction

MR images are often degraded by different artefacts. The artefact that is of major concern to automatic MR image processing is a low frequency spatially varying intensity inhomogeneity, known as the bias field that arise due to non-uniform field gradients (Liew and Yan, 2006).

The idea of bias field estimation in the EM framework was first proposed by Wells et al. (1996). The bias field is multiplicative for original MR intensities, and hence additive for log transformed intensities. Ashburner and Friston (2005) modelled the original multiplicative effect whereas Wells et al. (1996) and Van Leemput et al. (1999a) worked in the log space for bias field correction. We will discuss the work of Ashburner and Friston (2005) in the next section, which incorporates tissue classification, bias field correction and non-rigid atlas registration within a single probabilistic framework.

The additive bias field can be approximated as a linear combination of smooth basis functions $\Psi_l(\mathbf{x}_i)$ (Van Leemput et al., 1999a), such as polynomials and splines:

$$y_i = \tilde{y}_i + \sum_{l=1}^L a_l \Psi_l(\mathbf{x}_i)$$

Here y_i and $\tilde{y}_i$ are the observed and ideal intensities respectively at voxel i after log transformation. $\mathbf{A} = \{a_1, a_2, \dots, a_L\}$ denotes the bias field parameters, L the number of basis functions, and $\mathbf{x}_i$ the spatial location of voxel i . Below we summarise the EM algorithm developed by Van Leemput et al. (1999a) with bias field correction:

- E-step: The spatially constant mixing proportion c_k (Eq. 2.4) for the GMM is replaced with spatial prior p_{ik}^{atlas} obtained from a probabilistic brain atlas, which we will discuss in the next section:

$$p_{ik}^{(m+1)} = \frac{G(y_i - \sum_{l=1}^L a_l^{(m)} \Psi_l(\mathbf{x}_i), \mu_k^{(m)}, \sigma_k^{(m)}) p_{ik}^{atlas}}{\sum_{k'=1}^K G(y_i - \sum_{l=1}^L a_l^{(m)} \Psi_l(\mathbf{x}_i), \mu_{k'}^{(m)}, \sigma_{k'}^{(m)}) p_{ik'}^{atlas}} \quad (2.15)$$

- M-step: The ideal intensity $\tilde{y}_i$ at voxel i that is not corrupted by the bias field is reconstructed from the estimated posterior probabilities and Gaussian parameters.

$\tilde{y}_i$ is then subtracted from the observed intensity y_i at each location to form the residual image $\mathbf{R}$ from which the bias field is computed:

$$\begin{aligned}\mu_k^{(m+1)} &= \frac{\sum_{i=1}^N p_{ik}^{(m+1)} \left(y_i - \sum_{l=1}^L a_l^{(m)} \Psi_l(\mathbf{x}_i) \right)}{\sum_{i=1}^N p_{ik}^{(m+1)}} \\ \left(\sigma_k^{(m+1)} \right)^2 &= \frac{\sum_{i=1}^N p_{ik}^{(m+1)} \left(y_i - \sum_{l=1}^L a_l^{(m)} \Psi_l(\mathbf{x}_i) - \mu_k^{(m+1)} \right)^2}{\sum_{i=1}^N p_{ik}^{(m+1)}} \\ \left(\mathbf{A}^{(m+1)} \right)^T &= \left(\mathbf{F}^T \mathbf{W}^{(m+1)} \mathbf{F} \right)^{-1} \mathbf{F}^T \mathbf{W}^{(m+1)} \mathbf{R}^{(m+1)}\end{aligned}$$

where N is the number of voxels, and:

$$\begin{aligned}\mathbf{F} &= \begin{bmatrix} \Psi_1(\mathbf{x}_1) & \Psi_2(\mathbf{x}_1) & \dots & \Psi_L(\mathbf{x}_1) \\ \Psi_1(\mathbf{x}_2) & \Psi_2(\mathbf{x}_2) & \dots & \Psi_L(\mathbf{x}_2) \\ \vdots & & \ddots & \vdots \\ \Psi_1(\mathbf{x}_N) & \Psi_2(\mathbf{x}_N) & \dots & \Psi_L(\mathbf{x}_N) \end{bmatrix} \\ \mathbf{W}^{(m+1)} &= \text{diag} \left(\omega_i^{(m+1)} \right) \quad \text{with} \quad \omega_i^{(m+1)} = \sum_{k=1}^K \frac{p_{ik}^{(m+1)}}{\left(\sigma_k^{(m+1)} \right)^2} \\ \mathbf{R} &= \begin{bmatrix} y_1 - \tilde{y}_1^{(m+1)} \\ y_2 - \tilde{y}_2^{(m+1)} \\ \vdots \\ y_N - \tilde{y}_N^{(m+1)} \end{bmatrix} \quad \text{with} \quad \tilde{y}_i^{(m+1)} = \frac{\sum_{k=1}^K \omega_{ik}^{(m+1)} \mu_k^{(m+1)}}{\sum_{k'=1}^K \omega_{ik'}^{(m+1)}}\end{aligned}$$

2.2.4 Probabilistic brain atlas

The previous GMM completely depends on image intensities without any spatial regularisation. A probabilistic brain atlas that contains an intensity template and individual tissue probability maps has been employed to provide spatial prior information of human brain anatomy, and hence to improve the accuracy and robustness of automatic brain segmentation algorithms (Van Leemput et al., 1999a; Pohl et al., 2002; Ashburner and Friston, 2005). Fig. 2.2 shows the T₂w templates and the probability maps of grey matter (GM), white matter (WM) and cerebrospinal fluid (CSF) for the adult and neonatal brains. The

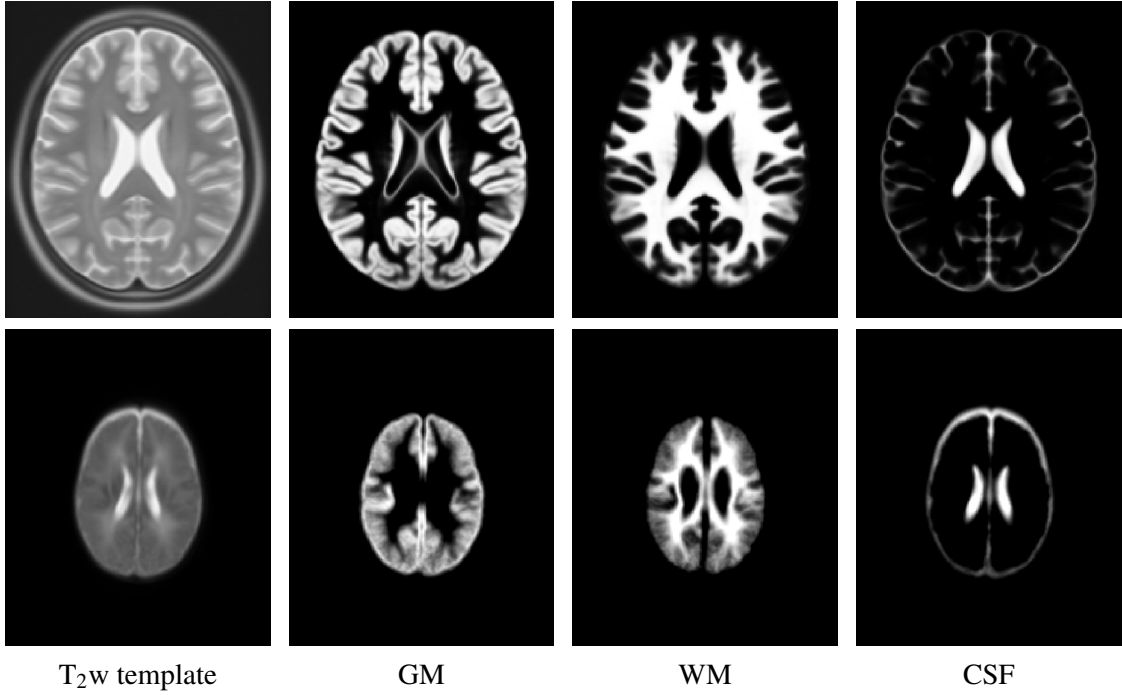


Figure 2.2: T₂w brain templates and the probability maps of grey matter (GM), white matter (WM) and cerebrospinal fluid (CSF) for adult (Fonov et al., 2011) and neonatal (Kuklisova-Murgasova et al., 2011) brain MRI. The relative sizes of the adult and neonatal brain atlases have been preserved when displaying the images. The neonatal probability map of GM only refers to the cortex without including the basal ganglia and thalami.

adult brain atlas, also known as the ICBM152 atlas¹ (Fonov et al., 2011), was created from 152 subjects aged between 18.5 and 43.5 years in the International Consortium for Brain Mapping database. The neonatal brain atlas² (Kuklisova-Murgasova et al., 2011) describes the anatomical variability of the developing brain based on 142 preterm infants between 29 and 44 weeks gestational age (GA).

In order to incorporate the probabilistic atlas in the EM framework, Van Leemput et al. (1999a) aligned the atlas with the image to be segmented using affine registration (Maes et al., 1997). The prior probability originally modelled as the mixing proportion c_k is now replaced with the value p_{ik}^{atlas} obtained from the affinely registered probability map of class k at voxel i , as shown in Eq. 2.15 in Section 2.2.3. Pohl et al. (2002) suggested aligning the atlas using non-rigid registration (Guimond et al., 2001) in case the brain image of an individual subject is considerably different from the average template.

Both Van Leemput et al. (1999a) and Pohl et al. (2002) implemented atlas registration

¹<http://www.bic.mni.mcgill.ca/ServicesAtlases/ICBM152NLin2009>

²<http://brain-development.org/brain-atlases>

as a separate pre-processing step. A different approach for joint segmentation and registration (Ashburner and Friston, 2005; Pohl et al., 2006a) proved successful in practice, particularly the unified segmentation model developed by Ashburner and Friston (2005) for the Statistical Parametric Mapping (SPM) software.

2.2.4.1 Unified segmentation in SPM

The SPM segmentation model simultaneously achieves tissue classification, bias field correction and non-rigid atlas registration within a single EM framework. The major differences compared to the work of Van Leemput et al. (1999a) are:

- Ashburner and Friston (2005) modelled the multiplicative bias field for the original intensities as the exponential of a linear combination of low frequency basis functions. The Gaussian PDF of class k has modified mean $\mu_k/\rho_i(\beta)$ and SD $\sigma_k/\rho_i(\beta)$, where ρ_i denotes the multiplicative bias field at voxel i , and β a vector of the linear coefficients.
- Instead of using the affinely registered probability map, SPM combines the stationary mixing proportion c_k with the tissue probability map which is allowed to deform according to parameters α . The displacement field is parameterised using a linear combination of cosine transform bases.

The posterior probability of voxel i being drawn from class k then becomes:

$$p_{ik}^{(m+1)} = \frac{G\left(y_i, \mu_k^{(m)}/\rho_i(\beta^{(m)}), \sigma_k^{(m)}/\rho_i(\beta^{(m)})\right) p(z_i = \mathbf{e}_k \mid \Phi_y^{(m)}, \alpha^{(m)})}{\sum_{k'=1}^K G\left(y_i, \mu_{k'}^{(m)}/\rho_i(\beta^{(m)}), \sigma_{k'}^{(m)}/\rho_i(\beta^{(m)})\right) p(z_i = \mathbf{e}_{k'} \mid \Phi_y^{(m)}, \alpha^{(m)})}$$

with the following spatial prior:

$$p(z_i = \mathbf{e}_k \mid \Phi_y^{(m)}, \alpha^{(m)}) = \frac{c_k^{(m)} p_{ik}^{atlas}(\alpha^{(m)})}{\sum_{k''=1}^K c_{k''}^{(m)} p_{ik''}^{atlas}(\alpha^{(m)})}$$

Ashburner and Friston (2005) defined the objective function F as the negative log likelihood based on Eq. 2.14 described in Section 2.2.2:

$$F = - \sum_{i=1}^N \sum_{l=1}^K p_{il}^{(m+1)} \ln G\left(y_i, \mu_l^{(m)} / \rho_i(\beta^{(m)}), \sigma_l^{(m)} / \rho_i(\beta^{(m)})\right) \\ - \sum_{i=1}^N \sum_{l=1}^K p_{il}^{(m+1)} \ln \frac{c_l^{(m)} p_{il}^{atlas}(\alpha^{(m)})}{\sum_{l'=1}^K c_{l'}^{(m)} p_{il'}^{atlas}(\alpha^{(m)})} - \ln P(\beta) - \ln P(\alpha)$$

Here $P(\beta)$ and $P(\alpha)$ are regularisation terms that represent the prior knowledge about the bias and displacement fields respectively. F is then optimised using an iterated conditional modes (ICM) approach (Besag, 1986). At each iteration, three parameter groups, including GMM parameters μ_k , σ_k and c_k , bias field parameters β and deformation parameters α , are optimised one at a time while fixing the others at their current best values. The GMM parameters are updated using the EM algorithm:

$$\mu_k^{(m+1)} = \frac{\sum_{i=1}^N p_{ik}^{(m+1)} \rho_i(\beta^{(m)}) y_i}{\sum_{i=1}^N p_{ik}^{(m+1)}} \\ \left(\sigma_k^{(m+1)}\right)^2 = \frac{\sum_{i=1}^N p_{ik}^{(m+1)} \left(\rho_i(\beta^{(m)}) y_i - \mu_k^{(m+1)}\right)^2}{\sum_{i=1}^N p_{ik}^{(m+1)}} \\ c_k^{(m+1)} = \frac{\sum_{i=1}^N p_{ik}^{(m+1)}}{\sum_{i=1}^N \frac{p_{ik}^{atlas}(\alpha^{(m)})}{\sum_{k'=1}^K c_{k'}^{(m)} p_{ik'}^{atlas}(\alpha^{(m)})}}$$

The parameters for the bias field and deformed atlas are both estimated using the Levenberg-Marquardt algorithm (Levenberg, 1944) which requires the first and second derivatives of F with respect to each parameter group. The detailed optimisation process can be found in Ashburner and Friston (2005).

2.2.5 Markov random fields

Van Leemput et al. (1999b) and Zhang et al. (2001) further incorporated neighbourhood information in the EM framework using MRFs (Li, 1995) to reduce noise in the resultant brain segmentations.

Let $\mathbf{Z} = \{\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_N\}$ be a family of random variables for an image of N voxels.

Li (1995) defined that $\mathbf{Z}$ is an MRF if:

$$p(\mathbf{z} | \Phi_z) > 0 \quad \forall \mathbf{z} \in \mathbb{Z} \quad \text{and} \quad p(\mathbf{z}_i | \mathbf{z}_S, \Phi_z) = p(\mathbf{z}_i | \mathbf{z}_{\mathcal{N}_i}, \Phi_z)$$

The set S includes all the voxels in the image except i , and $\mathbf{z}_{\mathcal{N}_i} = \{\mathbf{z}_{i'} | i' \in \mathcal{N}_i\}$. This means that the probability that voxel i belongs to a certain class depends on the set of labels in the local neighbourhood $\mathcal{N}_i$ of voxel i as well as the MRF parameters Φ_z . The Hammersley-Clifford theorem (Besag, 1974) states that such an MRF is equivalent to a Gibbs random field (GRF) which obeys the Gibbs distribution:

$$p(\mathbf{z} | \Phi_z) = \frac{1}{f(\Phi_z)} e^{-U_{\text{MRF}}(\mathbf{z} | \Phi_z)} \quad (2.16)$$

The normalisation constant $f(\Phi_z) = \sum_{\mathbf{z}} e^{-U_{\text{MRF}}(\mathbf{z} | \Phi_z)}$ is the partition function summed over all possible configurations of the hidden labels, and $U_{\text{MRF}}(\mathbf{z} | \Phi_z)$ is an energy function.

The EM algorithm applied to the independent GMM needs to be modified due to the interactions among the labels (Van Leemput et al., 1999b). Denoting the Gaussian means and SDs as Φ_y and $\Phi = \{\Phi_y, \Phi_z\}$, the posterior probability that voxel i belongs to class k based on its intensity y_i and the current parameter estimates $\Phi^{(m)}$ can be written as:

$$\begin{aligned} p_{ik}^{(m+1)} &= \frac{p(y_i, \mathbf{z}_i = \mathbf{e}_k | \Phi^{(m)})}{p(y_i | \Phi^{(m)})} \\ &= \frac{p(y_i | \mathbf{z}_i = \mathbf{e}_k, \Phi_y^{(m)}) p(\mathbf{z}_i = \mathbf{e}_k | \mathbf{z}_{\mathcal{N}_i}, \Phi_z^{(m)})}{\sum_{k'=1}^K p(y_i | \mathbf{z}_i = \mathbf{e}_{k'}, \Phi_y^{(m)}) p(\mathbf{z}_i = \mathbf{e}_{k'} | \mathbf{z}_{\mathcal{N}_i}, \Phi_z^{(m)})} \end{aligned}$$

As explained before, K is the number of classes, m the current iteration number, and $\mathbf{e}_k$ a column vector of length K with the k th component equal to one. The conditional PDF of class k $p(y_i | \mathbf{z}_i = \mathbf{e}_k, \Phi_y^{(m)})$ is still a Gaussian function $G(y_i, \mu_k^{(m)}, \sigma_k^{(m)})$, whereas the spatially constant mixing proportion c_k in the GMM has been replaced by the prior obtained from the MRF. Van Leemput et al. (1999b) adopted the mean field approach (Zhang, 1992) which assumed that the influence of the neighbouring labels $\mathbf{z}_{\mathcal{N}_i}$ can be

approximated by the influence of the probabilities $\mathbf{p}_{\mathcal{N}_i}^{(m)}$. We now have:

$$p_{ik}^{(m+1)} \simeq \frac{p(y_i | \mathbf{z}_i = \mathbf{e}_k, \Phi_y^{(m)}) p(\mathbf{z}_i = \mathbf{e}_k | \mathbf{p}_{\mathcal{N}_i}^{(m)}, \Phi_z^{(m)})}{\sum_{k'=1}^K p(y_i | \mathbf{z}_i = \mathbf{e}_{k'}, \Phi_y^{(m)}) p(\mathbf{z}_i = \mathbf{e}_{k'} | \mathbf{p}_{\mathcal{N}_i}^{(m)}, \Phi_z^{(m)})} \quad (2.17)$$

and the MRF prior takes the following Gibbs distribution form:

$$p(\mathbf{z}_i = \mathbf{e}_k | \mathbf{p}_{\mathcal{N}_i}^{(m)}, \Phi_z^{(m)}) = \frac{e^{-U_{\text{MRF}}(\mathbf{e}_k | \mathbf{p}_{\mathcal{N}_i}^{(m)}, \Phi_z^{(m)})}}{\sum_{k''=1}^K e^{-U_{\text{MRF}}(\mathbf{e}_{k''} | \mathbf{p}_{\mathcal{N}_i}^{(m)}, \Phi_z^{(m)})}} \quad (2.18)$$

Van Leemput et al. (1999b) computed the energy function for first-order MRFs as follows:

$$U_{\text{MRF}}(\mathbf{e}_k | \mathbf{p}_{\mathcal{N}_i}^{(m)}, \Phi_z^{(m)}) = \sum_{k_1=1}^K \sum_{j \in \mathcal{N}_i} h_{kk_1}^{(m)} p_{jk_1}^{(m)}$$

The element of a $K \times K$ connectivity matrix h_{kk_1} indicates the penalty when class k_1 is present in the neighbourhood of class k . Such a first-order MRF model favours homogeneous regions, and helps to smooth out isolated small misclassifications resulted from image noise.

We update the log likelihood for the complete data $\mathbf{q} = \{\mathbf{y}, \mathbf{z}\}$ using Eq. 2.16:

$$L_c(\Phi) = \sum_{i=1}^N \ln p(y_i | \mathbf{z}_i, \Phi_y) - U_{\text{MRF}}(\mathbf{z} | \Phi_z) - \ln f(\Phi_z)$$

The previous Q -function (Eq. 2.14) now becomes:

$$\begin{aligned} Q(\Phi | \Phi^{(m)}) &= \sum_{l=1}^K \sum_{i=1}^N p_{il}^{(m+1)} \ln G(y_i, \mu_l, \sigma_l) \\ &\quad - E \left[U_{\text{MRF}}(\mathbf{z} | \Phi_z) + \ln f(\Phi_z) | \mathbf{y}, \Phi^{(m)} \right] \end{aligned} \quad (2.19)$$

Therefore, the Gaussian means and SDs can be estimated in the same way as the independent GMM. Van Leemput et al. (1999b) optimised the MRF penalties h_{kk_1} using a least squares fit procedure.

Zhang et al. (2001) incorporated a hidden MRF model in the EM framework along with the bias field correction algorithm developed by Guillemaud and Brady (1997). The

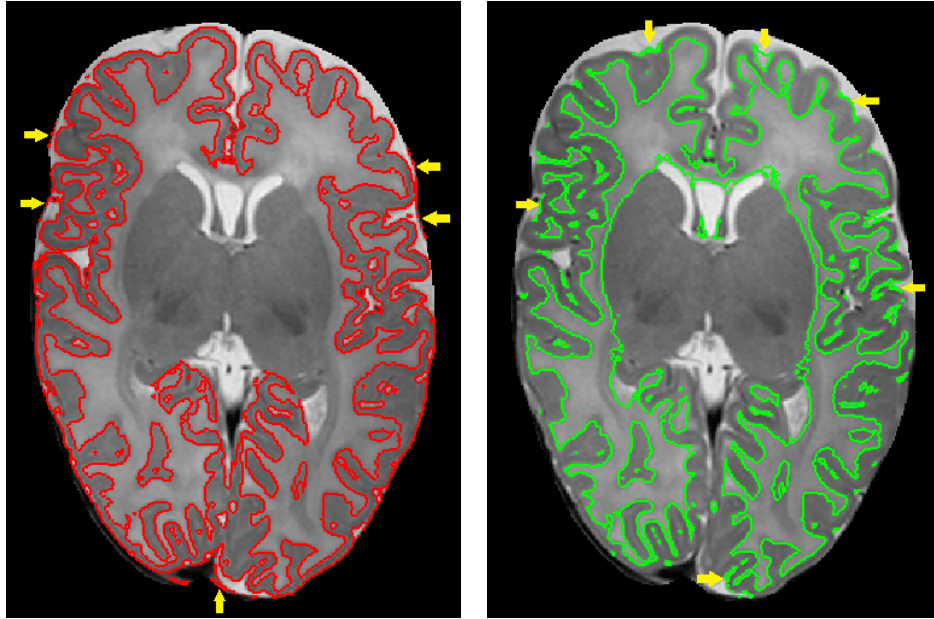


Figure 2.3: Misclassifications due to the partial volume effect in the segmentations of the cortical grey matter (CoGM) and white matter (WM) obtained using the SPM software on the T_2w image for an infant at 40 gestational weeks. The CoGM and WM are labelled in red and green respectively. Typical misclassifications have been pointed out by the yellow arrows. Voxels are frequently misclassified as CoGM at the boundary between the background and cerebrospinal fluid (CSF), and as WM at the CoGM/CSF boundary.

objective function was formulated based on the *maximum a posteriori* (MAP) principle, unlike the maximum likelihood estimation in Van Leemput et al. (1999b), and optimisation was achieved using the ICM algorithm (Besag, 1986). This method of Zhang et al. (2001) formed the basis of the segmentation tool FAST in the FSL software³ (Jenkinson et al., 2012) developed at the FMRIB Centre, Oxford.

2.2.6 Partial volume modelling

Another MR imaging artefact that affects automatic brain segmentation algorithms is the PV effect (Liew and Yan, 2006). Multiple tissues are present in a single voxel, and fine anatomical structures are lost due to the limited image resolution or the small head size in the case of neonatal brain MRI. We demonstrated in Chapter 1 that segmentation of myelination becomes extremely challenging due to the presence of PV voxels that cause a heavy tail between the intensity distributions of the composing pure tissues in the im-

³<https://fsl.fmrib.ox.ac.uk/fsl/fslwiki>

age histogram. General brain segmentation also suffers from the PV effect. Fig. 2.3 shows that voxels are frequently mislabelled as the cortical grey matter (CoGM) at the CSF/background (BKG) boundary, and as WM at the CoGM/CSF boundary.

To account for the fact that the intensity distribution of each tissue class may not be strictly Gaussian in the presence of PV voxels, Ashburner and Friston (2005) used more than one Gaussian per class in SPM, which is, however, insufficient in reducing PV misclassifications (see Fig. 2.3). This is because the Gaussians are mixed according to the global distributions, whereas the PV artefact occurs in the local neighbourhood which is not considered in the SPM segmentation model. Ledig et al. (2012) proposed an atlas-based segmentation method that suppresses classification of small WM patches at the boundary between CoGM and CSF via second-order MRFs. The first-order MRF framework (Van Leemput et al., 1999b) described in the previous section is expanded to:

$$p_{ik}^{(m+1)} = \frac{p_{ik}^{atlas} e^{-U_{1st}(\mathbf{e}_k | \mathbf{p}_{\mathcal{N}_i}^{(m)}, \Phi_z)} e^{-U_{2nd}(\mathbf{e}_k | \mathbf{p}_{\mathcal{N}_i}^{(m)}, \Phi_z)}}{\sum_{k'=1}^K p_{ik'}^{atlas} e^{-U_{1st}(\mathbf{e}_{k'} | \mathbf{p}_{\mathcal{N}_i}^{(m)}, \Phi_z)} e^{-U_{2nd}(\mathbf{e}_{k'} | \mathbf{p}_{\mathcal{N}_i}^{(m)}, \Phi_z)}} \quad (2.20)$$

Here p_{ik}^{atlas} is the value in the probabilistic atlas aligned with the subject's image at voxel i of class k . Ledig et al. (2012) defined the energy function for second-order MRFs as:

$$U_{2nd}(\mathbf{e}_k | \mathbf{p}_{\mathcal{N}_i}^{(m)}, \Phi_z) = \sum_{k_1=1}^K v_{k_1}^{(m)} \sum_{k_2=1}^K \mathbf{T}_k(k_1, k_2) v_{k_2}^{(m)} \quad (2.21)$$

where v_k calculates the sum of the k th class probabilities in the 6-neighbourhood $\mathcal{N}_i$, weighted by the anisotropic voxel spacing. The element of a 3D connectivity tensor, denoted as $\mathbf{T}_k(k_1, k_2)$, indicates the penalty when classes k_1 and k_2 are both present in the neighbourhood of class k . A large positive value can thus be assigned to $\mathbf{T}_{WM}(\text{CoGM}, \text{CSF})$ to penalise such an implausible tissue configuration caused by the PV effect.

Unlike the above methods in which one single tissue class is assigned to each voxel, Santiago and Gage (1993), Shattuck et al. (2001), Van Leemput et al. (2003), Tohka et al. (2004), Manjon et al. (2010), Cardoso et al. (2011) modelled the PV effect explicitly with a vector-valued random variable $\mathbf{t}_i$ that represents the local mixing fractions of K pure

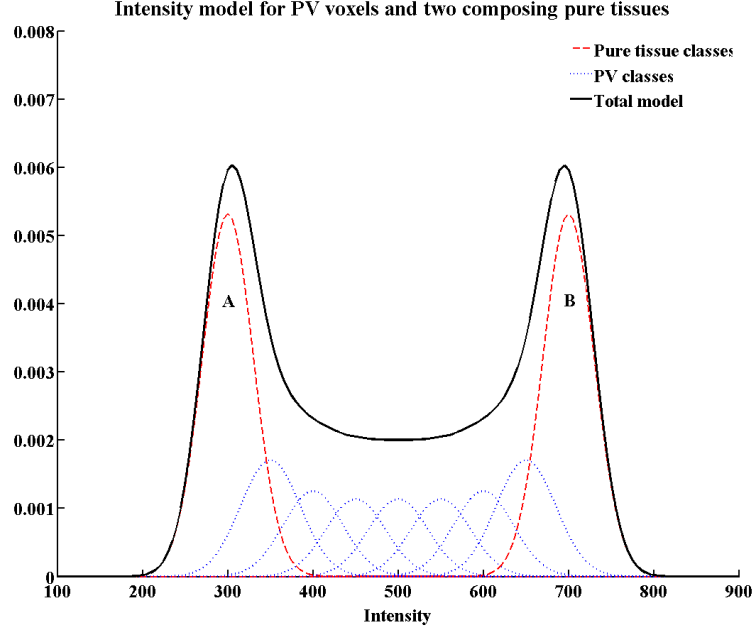


Figure 2.4: Intensity model for partial volume (PV) voxels and two composing pure tissues. A number of Gaussians are used to approximate the broad PV intensity distribution. The PV classes from left to right correspond to the following mixing fractions of pure tissue classes A and B: 87.5/12.5, 75/25, 62.5/37.5, 50/50, 37.5/62.5, 25/75, 12.5/87.5.

tissue classes in each voxel i . To simplify the problem, we only consider the situation where $K = 2$. The value of t_i was considered continuous except Van Leemput et al. (2003) who divided each voxel into subvoxels, so that only a discrete set of values are possible for t_i (see Fig. 2.4).

Santago and Gage (1993) assumed spatial independence of t_i between voxels. This approach addressed the problem of estimating the total volume of each pure tissue in the entire image, but not the amount in each individual voxel. Shattuck et al. (2001), Van Leemput et al. (2003), Tohka et al. (2004) and Manjon et al. (2010) applied a first-order MRF model to reduce the impact of noise during PV estimation. For example, Van Leemput et al. (2003) calculated the hidden label l_j of subvoxel j with the following energy function:

$$U(l_j = e_k | l_{N_i}, \Phi_L) = \alpha_k + \frac{1}{2}\beta \sum_{j' \in N_i} \delta(l_j, l_{j'})$$

The function $\delta(l_j, l_{j'})$ equals one if the labels of subvoxel j and its neighbour j' are different. The parameter β favours homogeneous regions of pure tissues, and α_k regulates the global mixing proportion of class k . The disadvantage of the first-order MRF model

is that it tends to over-smooth sharp transitions in the value of t_i between neighbouring voxels, which is undesirable for convoluted brain structures such as the cortex. Alternatively, Cardoso et al. (2011) used the subject-specific posterior probability maps of the composing pure tissues where the voxelwise normalised geometric mean was generated as *ad hoc* prior for PV voxels:

$$p_{ipv} = \frac{1}{\Pi_i} \frac{\sqrt{p_{ik_1} p_{ik_2}}}{0.5}$$

with Π_i a normalisation constant over all pure and mixed classes.

All these explicit PV estimation methods require the spatial locations of the composing pure tissues to be known first (Manjon et al., 2010). This is because the extra Gaussians introduced for explicit PV modelling hamper the segmentation robustness due to the large overlap in the intensity domain, and spatial prior knowledge is thus crucial for robust PV estimation (Van Leemput et al., 2003). Ledig et al. (2012) did not model the PV voxels explicitly, but the probabilistic atlases of the composing pure tissues were incorporated in the EM framework, that help to identify the implausible tissue configurations caused by PV.

2.3 Registration-based brain segmentation

Image registration is the process of determining a mapping from the coordinates in one space to another in order to achieve an optimal anatomical or functional correspondence (Fitzpatrick et al., 2000), and thus can be used to transform manual annotations for the purpose of brain segmentation (Heckemann et al., 2006). We will give a brief overview of the registration methodology in this section, and explain further details in Chapter 3 as part of the spatio-temporal analysis pipeline.

2.3.1 Image registration overview

A registration method generally consists of three components:

- A spatial transformation. The basic form of registration involves two images: a source or moving image I_S and a target or fixed image I_T . The two images are related by a transformation $\mathcal{T}$ that maps a point (x, y, z) in the target image to its corresponding point (x', y', z') in the source image:

$$\mathcal{T} : (x, y, z) \mapsto (x', y', z')$$

The transformation is defined unintuitively from the target image to the source image, known as backward mapping, which is easier than the opposite forward mapping from an implementation point of view (Sotiras et al., 2013). When voxel locations in the target space are mapped to non-voxel locations in the source space, the intensities of the source image at these non-voxel locations can be easily calculated by interpolation.

Rigid transformations that preserve straight lines and all non-zero angles have six degrees of freedom (DOFs), including three translational and three rotational parameters. Affine transformations preserve the straightness of lines, and hence parallelism, but the angles between lines are allowed to change. There are now twelve DOFs associated with translation, rotation, scaling and skewing. Non-rigid transformations can have thousands of DOFs depending on the choice of deformation models (Sotiras et al., 2013). Rigid transformations can be used to align images of the same patient. However, non-rigid transformations are necessary if significant brain development or different patients are involved, in which case a global affine transformation is usually performed first. We will discuss different types of deformation models in the next chapter for creating spatio-temporal brain templates.

- An objective function. The aim of registration is to obtain an estimate for the transformation that optimises the objective function of the general form:

$$C_1(I_T, I_S \circ \mathcal{T}) + C_2(\mathcal{T})$$

Here C_1 measures the similarity between the target image I_T and the aligned source image denoted by $I_S \circ \mathcal{T}$, and C_2 is a regularisation term for the transformation.

- An optimisation method. Optimisation methods only find a local minimum in the parameter space, particularly for non-rigid transformations with no closed-form solutions. Klein et al. (2007) compared the commonly used optimisation algorithms in image registration, such as gradient descent, conjugate gradient, Powell’s conjugate directions, Quasi-Newton and Levenberg-Marquardt. The technical details are out of scope of this thesis, but descriptions of the standard optimisers can be found in *Numerical Recipes in C*⁴ (Press et al., 1992).

2.3.2 Multi-atlas registration-based segmentation

An atlas in the context of registration-based segmentation refers to the pair of a brain MR image and its corresponding label volume that contains the manual annotations of different brain structures. In addition, multiple atlases are used to improve the segmentation accuracy. The overall pipeline (Heckemann et al., 2006) can be described as follows:

- Step 1. Non-rigid registration. The geometric transformation between an atlas image and the subject’s image is estimated using a deformation model.
- Step 2. Label propagation. The corresponding label volume is then transformed accordingly from the atlas space to the subject’s space.
- Step 3. Decision fusion. A number of propagated label volumes obtained from different starting atlases are combined through voxelwise application of the maximum-vote rule (Kittler et al., 1998).

Various deformation models (see Section 3.2 for details) have been tested for propagating atlases. For example, Heckemann et al. (2006) cross-validated the free-form deformation (FFD) model (Rueckert et al., 1999) on 30 healthy adult brain MR atlases, and Avants et al. (2008) compared a symmetric diffeomorphic registration method with an elastic body

⁴<http://numerical.recipes/>

model (Broit, 1981) and a Demons approach (Thirion, 1998) for segmenting elderly and neurodegenerative brain MRI.

An important issue in multi-atlas brain segmentation is atlas selection. Heckemann et al. (2006) discovered that the segmentation accuracy increases asymptotically with the number of atlases used while taking into account the computational cost. Aljabar et al. (2009) thus proposed to select a specific subset of atlases for each subject based on both an image similarity measure and age. In practice, the similarity metric is evaluated in the space of an arbitrary reference image with which all the atlases and the subject's image are aligned. Better segmentations were achieved using the top-ranked atlases than those given by random atlas subsets.

2.4 Other brain segmentation methods

In this section, we briefly describe some alternative algorithms that have been used in adult and neonatal brain MRI segmentation.

2.4.1 *k*-means clustering

k-means clustering (KMC) (MacQueen, 1967) partitions N data points into K disjoint clusters so that the following objective function is minimised:

$$F = \sum_{k=1}^K \sum_{i=1}^{N_k} \|\mathbf{x}_i - \mathbf{c}_k\|^2$$

where N_k represents the number of data points in cluster k , $\mathbf{x}_i$ the vector-valued i th data point, $\mathbf{c}_k$ the centroid of cluster k , and $\|\cdot\|$ the Euclidean distance between two vectors. The algorithm is initialised by assigning every data point randomly to one of the K clusters. The next two steps are then implemented iteratively until there are no further changes in the assignments: 1) calculating the centroid of each cluster; 2) reassigning every data point to the cluster with the closest centroid.

Unlike the one-to-one assignments in KMC, Dunn (1974) developed fuzzy *c*-means clustering (FCMC) by incorporating a membership value w_{ik} in the objective function,

that indicates the degree to which data point $\mathbf{x}_i$ belongs to cluster k :

$$F = \sum_{i=1}^N \sum_{k=1}^K w_{ik}^q \|\mathbf{x}_i - \mathbf{c}_k\|^2$$

where

$$w_{ik} = \frac{\|\mathbf{x}_i - \mathbf{c}_k\|^{-\frac{2}{q-1}}}{\sum_{k'=1}^K \|\mathbf{x}_i - \mathbf{c}_{k'}\|^{-\frac{2}{q-1}}}$$

The parameter q determines the level of cluster fuzziness, and in the limiting case when $q = 1$, the FCMC algorithm reduces to KMC.

Before probabilistic brain atlases for neonates became available, Xue et al. (2007) applied KMC to obtain the initial segmentations of GM, WM and CSF on neonatal brain MR images. The segmentations are then blurred to provide subject-specific spatial priors for the EM-MRF algorithm. Pham and Prince (1999) incorporated intensity inhomogeneity correction in the FCMC algorithm. Compared to the previous EM- and registration-based approaches, a significant problem of the KMC and FCMC algorithms is that both of them are purely intensity-based without any spatial regularisation.

2.4.2 k -nearest neighbour classification

The k -nearest neighbour (k NN) algorithm (Cover and Hart, 1967) is particularly suitable for non-parametric supervised classification when a large training dataset is available. Suppose that the training dataset P consists of prototype vectors defined in a D -dimensional feature space, and each prototype vector $\mathbf{p} \in P$ has associated with it a label l_p . A pattern vector $\mathbf{v}_i$ at voxel i of the image to be segmented can then be classified by assigning the label which is most frequent among the k prototype vectors nearest to $\mathbf{v}_i$ in the feature space:

$$P_i = \{\mathbf{p} \in P : \|\mathbf{v}_i - \mathbf{p}\| \leq d_k\} \quad \text{and} \quad l_i = \arg \max_l (l_p = l \mid \mathbf{p} \in P_i)$$

We use $\|\mathbf{v} - \mathbf{p}\|$ to denote the commonly used Euclidean distance measure:

$$\|\mathbf{v} - \mathbf{p}\|^2 = \sum_{f=1}^D (v_f - p_f)^2$$

Warfield et al. (2000) adapted the k NN algorithm for brain MRI segmentation, which interleaves tissue classification with elastic matching of an anatomical template. Apart from image intensities, spatial localisation of different brain structures obtained from the non-rigidly aligned template is used as an additional feature to improve the classification process. This method, however, requires a large number of training samples to be selected manually for each tissue class. The segmentation results thus can be affected by the particular choice of the training set.

Following the work of Warfield et al. (2000), Cocosco et al. (2003) proposed a “pruning” strategy for automatic sample selection. Initial labelling is first obtained via affine transformation of a probabilistic brain atlas into the subject’s space. The subsequent pruning procedure makes use of minimum spanning tree (MST) in the feature space. Long edges in the MST are removed iteratively until there are four distinct clusters for the GM, WM, CSF and BKG. The samples with consistent initial labelling and pruning results are chosen for the final k NN classification. One limitation is that this pruning method puts equal trust in the intensities of pure tissue voxels and the intensities of PV voxels positioned between clusters in the feature space. It would be desirable to incorporate local intensity gradient information or spatial prior knowledge of tissue boundaries into the algorithm.

2.4.3 Patch-based label propagation

The registration-based segmentation approach described in Section 2.3.2 makes the implicit assumption that there exists a one-to-one mapping between the subject’s image and the atlas image. Inspired by the work of Buades et al. (2005) in non-local image denoising, Rousseau et al. (2011) recently developed a patch-based framework for adult brain MRI segmentation that does not require any non-rigid registration.

Let us consider the simplest pair-wise case that involves an input image I_S for an individual subject and a single atlas consisting of an image I_A and the corresponding label volume L_A . The atlas used by Rousseau et al. (2011) is in the same image domain Ω as the input image I_S . The key of the patch-based method is a weighted graph $w(\mathbf{x}, \mathbf{y})$, $(\mathbf{x}, \mathbf{y}) \in \Omega^2$, that links together voxels $\mathbf{x}$ of the input image I_S and voxels $\mathbf{y}$ of the atlas image I_A :

$$w(\mathbf{x}, \mathbf{y}) = \exp \left(- \frac{\sum_{\mathbf{x}' \in P_{I_S}(\mathbf{x}), \mathbf{y}' \in P_{I_A}(\mathbf{y})} (I_S(\mathbf{x}') - I_A(\mathbf{y}'))^2}{2N\beta\sigma^2} \right)$$

The input image intensity at voxel $\mathbf{x}$ is denoted as $I_S(\mathbf{x})$, and $P_{I_S}(\mathbf{x})$ is a 3D patch of the image I_S centred at voxel $\mathbf{x}$. Similar notations $I_A(\mathbf{y})$ and $P_{I_A}(\mathbf{y})$ apply to the atlas image. N equals the number of voxels in the 3D patch, σ the SD of image noise, and β a smoothing parameter.

Patch-based label propagation relies on the assumption that, if patches of the input image I_S are locally similar to patches of the atlas image I_A , they should have a similar label. The input image can then be labelled by propagating the atlas label L_A using the weighted graph w that reflects the local similarities between I_S and I_A . Rousseau et al. (2011) computed w only for a limited number of voxels in the atlas image, i.e. $w(\mathbf{x}, \mathbf{y})$, $\forall \mathbf{x} \in \Omega, \mathbf{y} \in \mathcal{N}(\mathbf{x})$, where $\mathcal{N}(\mathbf{x})$ is the neighbourhood of location $\mathbf{x}$ in the atlas image I_A . Thus, the label $L_S(\mathbf{x})$ at voxel $\mathbf{x}$ of the input image is determined as follows:

$$\forall \mathbf{x} \in \Omega, L_S(\mathbf{x}) = \frac{\sum_{\mathbf{y} \in \mathcal{N}(\mathbf{x})} w(\mathbf{x}, \mathbf{y}) L_A(\mathbf{y})}{\sum_{\mathbf{y} \in \mathcal{N}(\mathbf{x})} w(\mathbf{x}, \mathbf{y})}$$

The patch-based approach relaxes the one-to-one constraint in registration-based segmentation. As a one-to-many block matching technique, it takes into account several good candidates to estimate label patches. However, a significant drawback is that patch-based label propagation is purely intensity-based, in contrast to the EM-based methods that can incorporate probabilistic spatial atlases, and the registration-based methods which can propagate existing labels via transformations. Poorly contrasted tissue boundaries or intensity variations would disturb the search process of similar patches, and cause less relevant patch examples to be included for assigning the final label. Therefore, it might be

useful to combine the patch-based strategy with the EM or non-rigid registration framework.

2.4.4 Coupled level sets

The coupled level set (CLS) method (Wang et al., 2011) combines local intensity information, atlas spatial prior and cortical thickness constraint in a single level set framework.

Let I be a brain MR image, and Ω be the image domain that is divided into four disjoint subsets Ω_k , $k \in \{\text{GM}, \text{WM}, \text{CSF}, \text{BKG}\}$. Instead of estimating the global Gaussian distribution of each tissue class, Wang et al. (2011) modelled the intensities in the local neighbourhood $\mathcal{N}(\mathbf{x})$, $\forall \mathbf{x} \in \Omega$, using a Gaussian PDF with spatially varying mean $\mu_k(\mathbf{x})$ and SD $\sigma_k(\mathbf{x})$:

$$p_{k,\mathbf{x}}(I(\mathbf{y})) = \frac{1}{\sqrt{2\pi}\sigma_k(\mathbf{x})} \exp\left(-\frac{(\mu_k(\mathbf{x}) - I(\mathbf{y}))^2}{2\sigma_k^2(\mathbf{x})}\right), \text{ for } \mathbf{y} \in \mathcal{N}(\mathbf{x})$$

where $I(\mathbf{y})$ is the intensity of voxel $\mathbf{y}$ in the neighbourhood $\mathcal{N}(\mathbf{x})$ of the current voxel $\mathbf{x}$. The energy function to be minimised is formulated as follows, which combines the above local Gaussian distribution fitting with spatial prior $p_k^{atlas}(\mathbf{y})$ obtained from a brain atlas (Shi et al., 2011) as well as level set functions $\Phi = (\phi_1, \phi_2, \phi_3)$:

$$F = \int \left(\sum_{k=1}^4 \int -\omega_\sigma(\mathbf{x} - \mathbf{y}) \ln \left(p_k^{atlas}(\mathbf{y}) p_{k,\mathbf{x}}(I(\mathbf{y})) \right) M_k(\Phi(\mathbf{y})) d\mathbf{y} \right) d\mathbf{x} + \nu \sum_{k=1}^3 \mathcal{L}(\phi_k)$$

A Gaussian kernel ω_σ with scale σ has been introduced to control the size of the neighbourhood. The zero-level surfaces of ϕ_1 , ϕ_2 and ϕ_3 correspond to the GM/WM, GM/CSF and CSF/BKG interfaces respectively. The parameter ν tunes the length term $\mathcal{L}(\phi_k)$ to maintain a smooth cortical surface.

On one hand, the CLS framework allows the incorporation of the cortical thickness constraint which takes the advantage that the human cerebral cortex is a thin, folded sheet of GM with an almost constant thickness. On the other hand, less clearly defined boundaries can be problematic for this method due to low tissue contrast and the presence of PV voxels. Wang et al. (2011) found it difficult to segment the deep grey matter (DGM)

in neonatal brain MR images because of the low contrast between GM and WM in the central brain region.

2.5 Quantitative metrics for segmentation evaluation

In Chapters 4 and 5, we choose the the Dice coefficient (DC) (Dice, 1945) and Hausdorff distance (HD) (Huttenlocher et al., 1993) to evaluate automatic brain segmentations with respect to manually annotated ground truth.

2.5.1 Dice coefficient

A DC is defined for two overlapping volumes as the ratio of the number of voxels in the intersection to the average number of voxels in the two volumes (Dice, 1945):

$$DC(A, B) = \frac{2N(A \cap B)}{N(A) + N(B)} \quad (2.22)$$

where $N(A)$ denotes the number of voxels in volume A .

2.5.2 Hausdorff distance

An HD can be used to measure the degree of similarity between two superimposed segmentations, and is defined as (Huttenlocher et al., 1993):

$$HD(A, B) = \max(h(A, B), h(B, A)) \quad (2.23)$$

where

$$h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\|$$

The symbol $\|\cdot\|$ represents the Euclidean distance between two points. The directed HD from A to B , denoted as $h(A, B)$, identifies point a in volume A that lies farthest from any point in volume B , and calculates the Euclidean distance from point a to the nearest point b in volume B .

2.6 Developments for neonatal brain segmentation

Specific spatial prior information for the developing brain is helpful when adapting the above adult brain segmentation methods to neonatal data. As explained in Chapter 1, this is due to the low tissue contrast and the significant size and shape differences compared to the adult brain. The PV effect, which is a general problem for *in vivo* MRI, needs to be considered more carefully in neonatal brain segmentation because of the limited image resolution and the small size of the newborn brain. Acquisition of high resolution images requires subjects to be still during long scanning periods, which is difficult to arrange when imaging babies. The automatic neonatal brain segmentation algorithms developed in recent years are summarised as follows:

- Prastawa et al. (2005) applied atlas-based EM classification to T_1w and T_2w newborn brain MR images. The GM and CSF segmentations were refined using the MCD estimator (Rousseeuw and Van Driessen, 1999), and the myelinated and unmyelinated WM were separated using the MST technique (Cocosco et al., 2003). A drawback is that the atlas was simulated by averaging the semi-automatic segmentations from three subjects, and therefore cannot represent the brain anatomy of the neonatal population.
- Xue et al. (2007) first performed intensity-based KMC on T_2w images, that divided the developing brain into multiple regions without requiring any probabilistic or manual atlases. The EM-MRF segmentation algorithm (Van Leemput et al., 1999b) was then implemented locally in each brain region. The first-order MRFs help to reduce small clusters of PV misclassifications.
- Anbeek et al. (2008) trained the k NN classifier in a leave-one-out fashion to segment T_2w and inversion recovery brain scans of 13 newborn infants. The learning set is comprised of samples selected from the manual annotations of 12 patients based on the features of voxel intensities and spatial locations.
- Kuklisova-Murgasova et al. (2011) generated the neonatal tissue probability maps shown in Fig. 2.2 by averaging the automatic segmentations computed on T_2w

images. Two segmentation approaches were used: the intensity-based method of Xue et al. (2007) for delineating the CoGM, WM and CSF, and registration-based multi-atlas segmentation (Rueckert et al., 1999; Heckemann et al., 2006) for the DGM, brainstem and cerebellum.

- Cardoso et al. (2011) proposed three post-processing refinements for atlas-based EM classification: 1) a prior relaxation scheme to reduce the influence of atlas alignment; 2) explicit PV classes modelled as mixed distributions of the composing pure tissues; 3) a locally varying MRF model for delineation of deep sulci and gyri.
- Ledig et al. (2012) focused on eliminating PV misclassifications by incorporating second-order neighbourhood information in the atlas-based EM framework, as described in Section 2.2.6.
- Wang et al. (2014) derived a patch-driven CLS framework for cortical segmentation on T_2w scans of the developing brain, that combines the local intensity information, atlas spatial prior and cortical thickness constraint defined using level set functions.
- Liu et al. (2016) segmented T_1w brain images of preterm infants spanning 20 gestational weeks using patch-based augmentation of EM with weighted fusion. The local patch search is constrained using anatomical variability maps learnt from atlas training data.

Recently, the NeoBrainS12 challenge (Isgum et al., 2015) provided an opportunity to evaluate the following state-of-the-art neonatal brain segmentation algorithms using the same T_1w and T_2w image sets acquired for preterm infants at 30 and 40 weeks GA:

- Makropoulos et al. (2014) developed a comprehensive segmentation pipeline. First the ALBERTs' labels of 50 brain regions (Gousias et al., 2012) were propagated to the space of each individual subject, and fused to form a probabilistic spatial prior using local-weighted voting (Artaechevarria et al., 2009). The regions that need to be modelled with more than one Gaussian were subdivided using KMC before the subsequent EM-MRF algorithm (Xue et al., 2007). A PV correction step

was inserted between the expectation and maximisation steps using knowledge-based rules and connected component labelling. Finally, the split brain regions were merged once the EM estimation converged.

- Melbourne et al. (2012) simulated a probabilistic brain atlas by averaging two manually annotated training images for incorporation in the method of Cardoso et al. (2011).
- Wu and Avants (2012) obtained spatial priors for the EM-MRF segmentation model by aligning a neonatal brain atlas (Serag et al., 2012) with the target image via symmetric diffeomorphic registration (Avants et al., 2008). Misclassified voxels due to unusual intensity values, such as PV, were removed by reimposing the atlas.
- Gui et al. (2012) developed an alternative approach without relying on any probabilistic or manual atlases, which is based on the watershed algorithm (Lotufo and Falcao, 2000), seeded region growing (Adams and Bischof, 1994) and active contour segmentation (Chan et al., 2000).
- We also participated in the challenge (Wang et al., 2012) where the SPM segmentation tool (Ashburner and Friston, 2005) was adapted to the atlas constructed by Kuklisova-Murgasova et al. (2011). Details of our method and the NeoBrainS12 challenge results will be presented in Chapter 4.

A key observation is that the above neonatal brain segmentation algorithms except Xue et al. (2007) and Gui et al. (2012), all depend on a probabilistic atlas, otherwise a large number of manual annotations are needed for multi-atlas segmentation or to simulate the probabilistic spatial prior. Chapter 3 will therefore focus on the construction of such an atlas that captures the anatomical variability of the developing brain.

Chapter 3

Review of spatio-temporal image analysis methods

3.1 Introduction

Spatio-temporal analysis is becoming more and more important in neuroimaging studies as it offers a framework for characterising changes of the human brain over time in a large population of subjects (Durrleman et al., 2014). The changes may be structural or functional, and can arise due to pathology, therapeutic intervention, neurodevelopment or neurodegeneration.

This chapter focuses on the construction of brain atlases. In Section 3.2, we describe the pairwise affine and deformable registration methods in detail, and briefly introduce groupwise registration. Spatial normalisation is achieved by transforming various images to a common average space found either by pairwise transformation averaging, which is the topic of Section 3.3, or directly by groupwise registration. Section 3.4 presents some candidates for the regression model fitted to the spatially normalised images to generate the age-dependent brain atlas. Lastly in Section 3.5, we explain how the above atlas construction pipeline can be used to estimate gestational ages (GAs) and brain growth.

3.2 Image registration

In this section, we first explain the similarity measures used for image registration. We then review the pairwise affine and deformable registration methods, followed by a brief introduction of groupwise registration.

3.2.1 Similarity measures

Let us denote the intensity of the target image at voxel i as t_i , and the intensity of the source image after transformation at the same location as s_i . Several similarity measures are available, and the simplest is the sum of squared differences (SSD) formulated as:

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

where N is the number of voxels in the target image. The SSD assumes an identity relationship between the intensity distributions of the target and source images, and hence requires both images to be obtained using the same modality and acquisition protocol. If a linear intensity relationship is allowed, the normalised cross correlation (NCC) can then be used:

$$NCC = \frac{1}{N} \cdot \frac{\sum_{i=1}^N (t_i - \bar{t})(s_i - \bar{s})}{\sqrt{\sum_{i=1}^N (t_i - \bar{t})^2 \sum_{i=1}^N (s_i - \bar{s})^2}}$$

where $\bar{t}$ and $\bar{s}$ are the mean intensities in the target and source images respectively.

For images acquired using different modalities, Maes et al. (1997) proposed to use mutual information (MI), or relative entropy, as a matching criterion. The entropy of the target image is defined as:

$$H(I_T) = - \sum_t p(t) \log p(t)$$

Here $p(t)$ denotes the probability distribution of observing intensity t in the target image, which is estimated directly from the histogram. The joint entropy of the target and source images is calculated similarly as:

$$H(I_T, I_S) = - \sum_{t,s} p(t,s) \log p(t,s)$$

Here $p(t, s)$ denotes the joint probability distribution of observing intensity t in the target image and intensity s at the corresponding location in the source image, and is again estimated from the joint histogram. The MI is related to the previously defined entropies according to:

$$MI = H(I_T) + H(I_S) - H(I_T, I_S)$$

Studholme et al. (1999) further proposed an overlap invariant entropy measure, called normalised mutual information (NMI), which is the ratio of the sum of the marginal entropies and the joint entropy:

$$NMI = \frac{H(I_T) + H(I_S)}{H(I_T, I_S)} \quad (3.1)$$

We use the NMI in the later chapters as a similarity measure for image alignment.

3.2.2 Affine registration

An affine transformation that maps a point (x, y, z) in the target image to its corresponding location (x', y', z') in the source image can be written as:

$$\begin{pmatrix} x' \\ y' \\ z' \\ 1 \end{pmatrix} = \begin{pmatrix} \theta_{11} & \theta_{12} & \theta_{13} & \theta_{14} \\ \theta_{21} & \theta_{22} & \theta_{23} & \theta_{24} \\ \theta_{31} & \theta_{32} & \theta_{33} & \theta_{34} \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix} \quad (3.2)$$

The coefficients Θ parameterise the twelve degrees of freedom (DOFs) of the transformation associated with translation, rotation, scaling and skewing.

Mazziotta et al. (2001) applied affine registration when building an adult brain atlas based on the images of normal individuals aged between 18 and 90 years, collected for the International Consortium for Brain Mapping (ICBM) project. Recently, Kuklisova-Murgasova et al. (2011) constructed a neonatal brain atlas using affine registration as well. This is because affine transformations only correct for the global differences in the brain size and shape without losing the local spatial variability, and the linearity of affine

transformations facilitates the computation of the average transformation (Alexa, 2002) during spatial normalisation (see Section 3.3).

3.2.3 Non-rigid registration using free-form deformations

Sometimes it is desirable to incorporate local deformations for better alignment of the detailed brain structures. Below we describe the non-rigid registration method using free-form deformations (FFD) developed by Rueckert et al. (1999). This method is used in our work for modelling the myelination process.

The basic idea of the FFD model is to deform an image by manipulating an underlying mesh of control points based on B-splines (Lee et al., 1997). The spatial transformation $\mathcal{T}$ is given as the addition of a global affine transformation $\mathcal{T}_{affine}$ and a local spline-based displacement field $\mathbf{u}$:

$$\mathcal{T}(x, y, z) = \mathcal{T}_{affine}(x, y, z) + \mathbf{u}(x, y, z) \quad (3.3)$$

where $\mathcal{T}_{affine}$ is characterised by the coefficients Θ in Eq. 3.2.

The deformation is modelled by 3D tensor products of cubic B-splines. Let Φ be a $n_x \times n_y \times n_z$ mesh of control points $\phi_{i,j,k}$ with uniform spacing δ superimposed on the deformed image. The displacement field $\mathbf{u}$ can be written as:

$$\mathbf{u}(x, y, z) = \sum_{p=0}^3 \sum_{q=0}^3 \sum_{r=0}^3 B_p(t_1) B_q(t_2) B_r(t_3) \phi_{i+p, j+q, k+r} \quad (3.4)$$

where $i = \lfloor x/n_x \rfloor - 1$, $j = \lfloor y/n_y \rfloor - 1$, $k = \lfloor z/n_z \rfloor - 1$, $t_1 = x/n_x - \lfloor x/n_x \rfloor$, $t_2 = y/n_y - \lfloor y/n_y \rfloor$, $t_3 = z/n_z - \lfloor z/n_z \rfloor$. B_p represents the p th cubic B-spline basis function:

$$\begin{aligned} B_0(t) &= \frac{1}{6}(1-t)^3 \\ B_1(t) &= \frac{1}{6}(3t^3 - 6t^2 + 4) \\ B_2(t) &= \frac{1}{6}(-3t^3 + 3t^2 + 3t + 1) \\ B_3(t) &= \frac{1}{6}t^3 \end{aligned}$$

The overall cost function to be minimised comprises two competing goals:

$$C(\Theta, \Phi) = -C_{similarity}(I_T, I_S \circ \mathcal{T}) + \lambda C_{smooth}(\mathcal{T})$$

The first term represents the NMI similarity measure defined in Eq. 3.1. The second term is a penalty term that constrains the local displacement field to be smooth, and takes the following form:

$$C_{smooth}(\mathcal{T}) = \frac{1}{V} \int_0^X \int_0^Y \int_0^Z \left[\left(\frac{\partial^2 \mathcal{T}}{\partial x^2} \right)^2 + \left(\frac{\partial^2 \mathcal{T}}{\partial y^2} \right)^2 + \left(\frac{\partial^2 \mathcal{T}}{\partial z^2} \right)^2 + 2 \left(\frac{\partial^2 \mathcal{T}}{\partial xy} \right)^2 + 2 \left(\frac{\partial^2 \mathcal{T}}{\partial xz} \right)^2 + 2 \left(\frac{\partial^2 \mathcal{T}}{\partial yz} \right)^2 \right] dx dy dz$$

where V denotes the volume of the image domain with $0 \leq x < X$, $0 \leq y < Y$ and $0 \leq z < Z$. $C_{smooth}(\mathcal{T})$ equals zero for any affine transformation, and the parameter λ compromises between image alignment and transformation smoothness.

The advantages of the FFD model are that the B-splines are locally controlled, and that the degree of non-rigid deformation can be flexibly adjusted based on the resolution of the control point mesh. The algorithm is implemented in a coarse-to-fine multi-resolution framework, and has been previously used to perform growth analyses for the rapidly developing fetal and neonatal brains (Aljabar et al., 2008; Wright et al., 2015; Makropoulos et al., 2016).

3.2.4 Other deformation models

Apart from B-spline FFDs, there are other deformable image registration methods available that are inspired by different physical models.

3.2.4.1 Elastic body model

Broit (1981) first modelled a deformed image as an elastic material, whose displacement field $\mathbf{u}$ is governed by the Navier-Cauchy partial differential equation:

$$\mu \nabla^2 \mathbf{u} + (\mu + \lambda) \nabla(\nabla \cdot \mathbf{u}) + \mathbf{F} = 0$$

where μ and λ are parameters for the elastic modulus of the material in the context of elasticity. The image grid is deformed under the influence of the force field $\mathbf{F}$ that contains two components. An external force tries to deform the image while an internal force imposes the elastic properties of the material. Matching is achieved when the two components reach an equilibrium.

3.2.4.2 Viscous fluid flow model

Alternatively, Christensen et al. (1996) modelled the deformed image as a viscous fluid flow, whose velocity field $\mathbf{v}$ is governed by the simplified Navier-Stokes equation:

$$\mu \nabla^2 \mathbf{v} + (\mu + \lambda) \nabla(\nabla \cdot \mathbf{v}) + \mathbf{F} = 0$$

Here μ and λ represent viscosity coefficients. Two mechanisms are identified. The first $\mu \nabla^2 \mathbf{v}$ is associated with constant volume viscous flow of the image, and the second $(\mu + \lambda) \nabla(\nabla \cdot \mathbf{v})$ is related to the growth or shrinkage of local regions. The velocity field is integrated to obtain the displacement field.

Compared to the elastic body model, a significant advantage of the viscous fluid flow model is the ability to handle large magnitude non-linear deformations. Miller et al. (1997) applied both the large deformation fluid transformation and the small deformation elasticity transformation in a hierarchical registration scheme for studying computational brain anatomy. This composite approach was employed by Fonov et al. (2011) to build the ICBM152 adult brain atlas.

3.2.4.3 Diffusion model

Most of the algorithms in this category are based on the work of Thirion (1998) who made an analogy between image matching and a diffusion process controlled by the demons in Maxwell's well-known thought experiment. The main idea is to consider the objects' boundaries in one image as semi-permeable membranes, and to let the other image, modelled as a deformable grid, diffuse through the interfaces by the action of demons situated within the membranes.

This demons approach is able to cope with large deformations (Thirion, 1998), and was used by Guimond et al. (2000) to obtain an average T₁w coronal brain template based on 124 adults.

3.2.4.4 Flows of diffeomorphisms

Registration algorithms commonly assume that similar structures are present in both the target and source images, i.e. topology is preserved (Rueckert et al., 2006). Such topology preserving algorithms produce smooth and invertible transformations known as diffeomorphisms. Whether a transformation is diffeomorphic can be checked by calculating a map of Jacobian determinants for which the details will be explained in Section 3.5.2. The Jacobian determinant contains information about the volume change induced by the transformation, and is positive for diffeomorphic transformations.

Flows of diffeomorphisms have been proposed to model deformations (Christensen et al., 1996; Dupuis et al., 1998). In this case, locations in an image are carried via a smooth change of variable into new locations in another image according to the Lagrange transport equation. The key issue is the regularisation term that requires the velocity field to be smooth so that the associated flow is supported on the space of diffeomorphisms:

$$C_{smooth} = \int_0^1 \|v_t\|_V^2 dt$$

Here $\|f\|_V$ is a norm on the space V of smooth velocity fields defined as $\|Df\|_{L_2}$, where D is a differential operator and $\|\cdot\|_{L_2}$ the L_2 norm. Variation of the velocity field over time

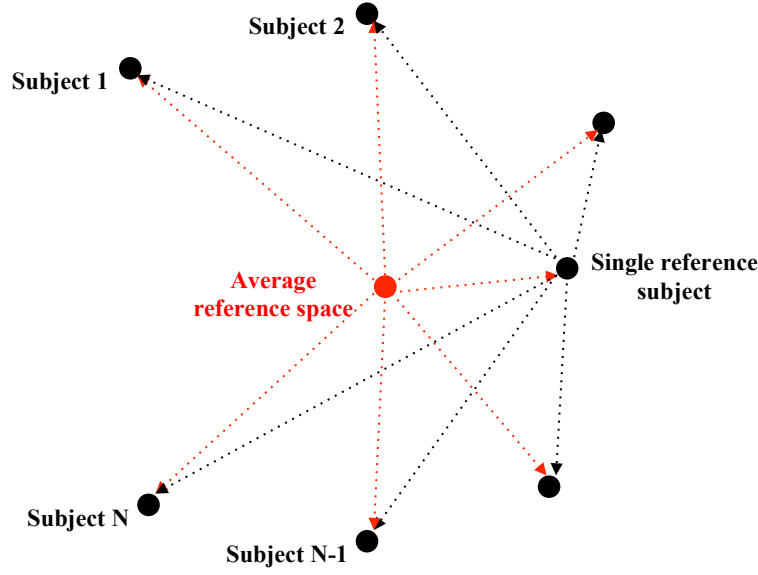


Figure 3.1: Groupwise registration for spatial co-alignment of various images in an average reference space as compared to pairwise image registration towards a single reference subject.

allows the estimation of large deformations, and hence this framework is also known as large deformation diffeomorphic metric mapping (LDDMM).

3.2.5 Groupwise registration

An alternative approach for spatial co-alignment of images in different spaces is groupwise registration (Avants and Gee, 2004; Bhatia et al., 2004; Studholme and Cardenas, 2004), illustrated schematically in Fig. 3.1. Fonov et al. (2011) used groupwise registration to construct the ICBM152 adult brain atlas shown in Fig. 2.2 in Chapter 2. The images in the project database were first registered with an initial estimate of the average intensity template $\bar{I}$ using the deformation model proposed by Miller et al. (1997). The template was then updated iteratively until the following conditions were satisfied:

- The sum of squared intensity differences between the template and the individual images after transformation is minimised:

$$\bar{I}^* = \arg \min_{\bar{I}} \left[\sum_{n=1}^N \int_{\mathbf{x}} (\bar{I}(\mathbf{x}) - I_n \circ \mathcal{T}_{n,\bar{I}}(\mathbf{x}))^2 d\mathbf{x} \right]$$

where $\bar{I}(\mathbf{x})$ is the template intensity at location $\mathbf{x}$, and $I_n \circ \mathcal{T}_{n,\bar{I}}(\mathbf{x})$ the intensity in the

individual image after transformation $\mathcal{T}_{n,\bar{I}}$ that maps location $\mathbf{x}$ in the average space to the space of subject n .

- The total magnitude of all the deformations is minimised:

$$\bar{I}^* = \arg \min_{\bar{I}} \left[\sum_{n=1}^N \int_{\mathbf{x}} |\mathcal{T}_{n,\bar{I}}(\mathbf{x}) - \mathbf{x}|^2 d\mathbf{x} \right]$$

3.3 Spatial normalisation

Spatial normalisation refers to the transformation of images from different spaces to a common average space found either by pairwise transformation averaging or by group-wise registration directly, and we focus on the former in this section.

When the pairwise registration methods are used for spatial co-alignment, the images of different subjects are typically registered with a single reference subject (Evans et al., 2012), such as the Talairach space based on a single 60-year-old female brain (Talairach and Tournoux, 1988). The final atlas is inevitably biased towards the chosen reference subject. This can cause problems because such an atlas is usually used to serve as a template for assessing brain development, or to provide spatial priors for tissue segmentation. Transformation averaging has thus been proposed to solve this issue.

Alexa (2002) derived a geometrically meaningful definition for a linear combination of transformations that can be represented as symmetric, positive-definite matrices (Arsigny et al., 2007), hence affine transformations:

$$\bar{\mathcal{A}} = \exp \left(\sum_{n=1}^N w_n \log \mathcal{A}_n \right)$$

The symbol $\bar{\mathcal{A}}$ denotes the average transformation, $\mathcal{A}_n$ the pairwise affine transformation between the reference subject and subject n , and N the number of subjects. Davis et al. (2007) calculated the weight w_n using a Gaussian kernel, and the average transformation

$\bar{\mathcal{A}}$ becomes a function of age t :

$$\bar{\mathcal{A}}(t) = \exp \left(\frac{\sum_{n=1}^N w(t_n, t) \log \mathcal{A}_n}{\sum_{n'=1}^N w(t_{n'}, t)} \right) \quad (3.5)$$

where

$$w(t_n, t) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left(-\frac{(t_n - t)^2}{2\sigma^2} \right) \quad (3.6)$$

Here t_n is the age of subject n , and σ the standard deviation (SD) of the Gaussian kernel. The images are then transformed from the spaces of individual subjects to the average space using the composed transformations $\mathcal{A}_n \circ \bar{\mathcal{A}}(t)^{-1}$.

Rueckert et al. (2003) developed a statistical deformation model by extending the idea of averaging to the non-linear FFD deformations. Let $\mathbf{u}_n(x, y, z)$ be the pairwise spline-based displacement field that warps the image of the reference subject to match the image of subject n , as described in Eq. 3.4. The average displacement field $\bar{\mathbf{u}}$ can be obtained by averaging the control points for B-spline representation (Kuklisova-Murgasova et al., 2009):

$$\bar{\mathbf{u}}(x, y, z) = \frac{1}{N} \sum_{n=1}^N \mathbf{u}_n(x, y, z)$$

which is further modified by Ericsson et al. (2008) to incorporate the Gaussian kernel weights calculated using Eq. 3.6:

$$\bar{\mathbf{u}}(x, y, z, t) = \frac{\sum_{n=1}^N w(t_n, t) \mathbf{u}_n(x, y, z)}{\sum_{n'=1}^N w(t_{n'}, t)} \quad (3.7)$$

While the inversion of the average affine transformation is straightforward, the average displacement field is, however, not analytically invertible. Aljabar et al. (2008) computed $\bar{\mathbf{u}}^{-1}$ using the inverse filtering method proposed by Unser (1999).

3.4 Regression models

A variety of voxelwise regression models can be fitted to the spatially normalised images to generate an age-dependent brain atlas.

3.4.1 Kernel smoothing

Kernel smoothing is a standard statistical technique for interpolating between samples and filtering out high-frequency noise from signals (Hutton et al., 2003). This method has been applied to construct an anatomical template $\bar{I}$ at any given age t through voxelwise weighted intensity averaging (Ericsson et al., 2008; Kuklisova-Murgasova et al., 2011):

$$\bar{I}(t) = \frac{\sum_{n=1}^N w(t_n, t) I_n \circ \mathcal{T}_n \circ \bar{\mathcal{T}}^{-1}}{\sum_{n'=1}^N w(t_{n'}, t)} \quad (3.8)$$

Here N is the number of subjects, and $\bar{\mathcal{T}}$ the average transformation, be it affine (Eq. 3.5) or spline-based (Eq. 3.7). The images of different subjects are co-aligned in the average reference space at age t using the transformations $\mathcal{T}_n \circ \bar{\mathcal{T}}^{-1}$.

The weight $w(t_n, t)$ is again calculated using the Gaussian kernel defined in Eq. 3.6, and heavy weights are assigned to the subjects with ages t_n close to the atlas age t . The SD σ of the Gaussian kernel is determined by the age distribution of the input images. Small values of σ will introduce high-frequency noise, whereas large values over-smooth the variation over time. A suitable kernel width is found based on experience with specific datasets. For instance, Ericsson et al. (2008) tested $3 < \sigma < 5$ years for modelling the aging brain. Kuklisova-Murgasova et al. (2011) constructed the neonatal brain atlas based on affinely registered T₂w images using an SD of two weeks.

Tissue probability maps such as those in Fig. 2.2 in Chapter 2 can be created similarly. The posterior probability map (PPM) $P_{k,n}$ of tissue k for subject n is transformed accordingly using $\mathcal{T}_n \circ \bar{\mathcal{T}}^{-1}$ into the average reference space at age t , followed by voxelwise weighted probability averaging:

$$P_k(t) = \frac{\sum_{n=1}^N w(t_n, t) P_{k,n} \circ \mathcal{T}_n \circ \bar{\mathcal{T}}^{-1}}{\sum_{n'=1}^N w(t_{n'}, t)} \quad (3.9)$$

The probabilities in the PPMs are already in the same range between zero and one, whereas the intensities of the individual images need to be normalised using the z-statistic analysis before fitting the regression model (Aljabar et al., 2008).

A problem of fixed-width kernel smoothing is that it is affected by the non-uniform

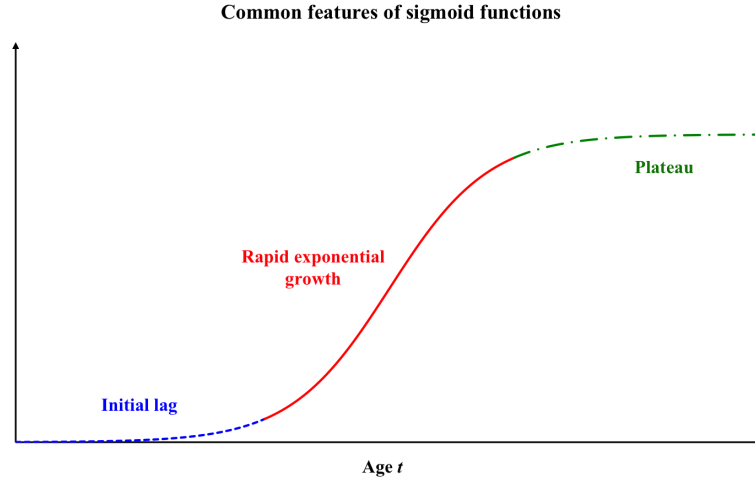


Figure 3.2: Common features of sigmoid functions. Despite different parameterisations, these functions all experience an initial lag or a period of slow increase, followed by a transitional period of rapid exponential growth, and finally reach a plateau of little change in the modelled quantity.

distribution of subjects at different ages, and we will demonstrate this point when modelling the myelination process in Chapter 6. Serag et al. (2012) thus proposed adaptive kernel regression to address uneven sampling distributions with a time-varying Gaussian SD.

3.4.2 Group of sigmoid functions

A group of parametric growth models, called sigmoid functions, can be fitted voxelwise to the spatially normalised images in the average space to describe the variation of signal intensities or tissue probabilities at each voxel over time. Despite different parameterisations, these functions all experience an initial lag or a period of slow increase, followed by a transitional period of rapid exponential growth, and finally reach a plateau of little change in the modelled quantity, as illustrated in Fig. 3.2. These properties make the sigmoid functions particularly suitable for modelling myelination because we expect such a behaviour in myelin growth.

Dean III et al. (2014) measured the myelin water fraction (MWF) in 122 healthy male children from 3 to 60 months of age using an advanced MRI technique based on T_1 and T_2 relaxometry. As listed in Table 3.1, the aligned MWF maps with values between zero and

one were fitted with the logistic and hyperbolic tangent functions as well as the growth models of Gompertz (1825), modified Gompertz (Mogensen and Vieth, 1973), Bleasdale and Nelder (1960), Stannard et al. (1985) and Richards (1959). These models were compared using the Bayesian Information Criterion (BIC) proposed by Schwarz (1978) that evaluates the fit residuals with a penalty for a large number of model parameters.

Model	Number of parameters	Functional form of age t
Logistic	3	$\frac{\alpha}{1 + e^{\beta - \gamma t}}$
Gompertz	3	$\alpha e^{-e^{\beta - \gamma t}}$
Modified Gompertz	4	$\alpha e^{-e^{\beta - \gamma t} + \delta t}$
Bleasdale and Nelder	4	$\frac{1}{(\alpha + \beta t^\delta)^{1/\gamma}}$
Stannard	4	$\frac{\alpha}{\left(1 + e^{-\frac{\beta + \gamma t}{\delta}}\right)^\delta}$
Richards	4	$\frac{\alpha}{(1 + \beta e^{\gamma(\delta - t)})^{1/\delta}}$
Hyperbolic tangent	4	$\alpha + \beta \tanh(\gamma x - \delta)$

Table 3.1: Age-dependent sigmoid functions. The parameters written in Greek letters are estimated using a Levenberg-Marquardt non-linear least squares algorithm (Levenberg, 1944).

3.4.3 Polynomial functions

Another group of commonly used parametric growth models are polynomial functions, as used by Habas et al. (2010) to develop an MR intensity template and tissue probability maps for the fetal brain. The images were co-aligned in an average reference space using elastic groupwise registration (Broit, 1981; Studholme and Cardenas, 2004), and a temporal polynomial of degree D was fitted voxelwise to model the MR intensity y_i at voxel i in the average space:

$$y_i = \sum_{d=0}^D \alpha_{id} t^d \quad (3.10)$$

with the coefficients α_{id} of the d th degree term at voxel i estimated again through least squares fitting.

However, if the above polynomial function is applied directly to generate tissue probability maps, the resulting model estimates may not be in the range of valid probabilities between zero and one. Habas et al. (2010) thus transformed the probabilities to the logarithm of the odds (LogOdds) (Pohl et al., 2006b) before fitting the polynomial model:

$$\mathcal{O}(p_{ik}) = l_{ik} = \log \left(\frac{p_{ik}}{p_{iNB}} \right)$$

Here p_{ik} is the probability that voxel i belongs to one of the following tissues k : grey matter (GM), white matter (WM), germinal matrix (GMAT) and ventricles (VEN). The corresponding LogOdds value l_{ik} is defined as the logarithm of the ratio between p_{ik} and p_{iNB} which is the probability of voxel i being the non-brain structures. The LogOdds of class $k \in \{\text{GM}, \text{WM}, \text{GMAT}, \text{VEN}\}$ can then be modelled using a voxelwise temporal polynomial similar to Eq. 3.10.

3.5 Applications of the atlas construction pipeline

Apart from atlas-based EM brain segmentation, below we explain how the above atlas construction pipeline can be applied for morphological age estimation (Dittrich et al., 2014) and deformation-based morphometry (Ashburner et al., 1998; Aljabar et al., 2008).

3.5.1 Morphological age estimation

Dittrich et al. (2014) built a spatio-temporal latent atlas upon 32 healthy fetuses between 20 and 30 weeks GA for semi-supervised learning of fetal brain segmentations and morphological age estimation.

Firstly, the atlas was initialised using fixed-width kernel regression with $\sigma = 2$ weeks based on the manual segmentations available for a subset of images. The brain segmentations for all the images were initialised by mapping the manual labels to the unannotated images. The spatio-temporal atlas was then re-estimated based on the current segmen-

tations, and in an alternating manner, the segmentations were refined using a level set framework that incorporates the current atlas as spatial prior.

The morphological age, rather than the nominal age, of a given subject was estimated by assessing the similarity between its brain segmentation and the normal subjects' atlas. Dittrich et al. (2014) integrated the age estimation problem into the whole iterative algorithm by optimising the following expression:

$$\hat{t}^* = \arg \max_{t^*} (-\log p(\Gamma \mid \eta, t^*))$$

where t^* is an age estimate for a brain segmentation Γ based on the atlas η .

3.5.2 Deformation-based morphometry

A different type of atlases have been constructed to describe brain growth based on deformation fields (Ashburner et al., 1998; Aljabar et al., 2008), and the analysis is known as deformation-based morphometry (DBM).

Aljabar et al. (2008) performed DBM on the children born prematurely with longitudinal T₁w scans at one and two years of age. An individual growth map was first computed for each subject as follows. Intra-subject FFD registration (Rueckert et al., 1999) was carried out using a baseline image in year one and a follow-up scan in year two as the target and source images respectively. The Jacobian operator D of the resultant transformation $\mathcal{T}$ is written as:

$$D\mathcal{T} = \begin{pmatrix} \frac{\partial \mathcal{T}_x}{\partial x} & \frac{\partial \mathcal{T}_x}{\partial y} & \frac{\partial \mathcal{T}_x}{\partial z} \\ \frac{\partial \mathcal{T}_y}{\partial x} & \frac{\partial \mathcal{T}_y}{\partial y} & \frac{\partial \mathcal{T}_y}{\partial z} \\ \frac{\partial \mathcal{T}_z}{\partial x} & \frac{\partial \mathcal{T}_z}{\partial y} & \frac{\partial \mathcal{T}_z}{\partial z} \end{pmatrix}$$

The volume change due to the transformation at target location $\mathbf{x}$ is given by the Jacobian determinant $J(\mathbf{x})$. From Eq. 3.3, we have:

$$J(\mathbf{x}) = \det(D\mathcal{T}(\mathbf{x})) = \det(D\mathcal{T}_{affine}(\mathbf{x}) + D\mathbf{u}(\mathbf{x}))$$

where $D\mathcal{T}_{affine}$ for the global affine transformation and $D\mathbf{u}$ for the local B-spline displace-

ment field can be both calculated analytically. Such an image of the Jacobian determinants constitutes the growth map of an individual subject.

Aljabar et al. (2008) then generated an atlas based on the growth maps for a large population of subjects by estimating the average inter-subject transformation at each time point. Similar to the intensity template, the Jacobian determinants need to be normalised first. Aljabar et al. (2008) chose the geometric mean of the Jacobian determinants at location $\mathbf{x}$ over all the N subjects as the voxelwise normalisation constant:

$$J_{norm}(\mathbf{x}) = \left(\prod_{n=1}^N J_n(\mathbf{x}) \right)^{\frac{1}{N}}$$

3.6 Summary

In this chapter, we described a pipeline for constructing spatio-temporal brain atlases. Firstly, the images of a large population of subjects are co-aligned using either pairwise or groupwise registration. Among the pairwise registration methods, affine transformations only correct for the global differences in the brain size and shape without losing the local structural variability (Kuklisova-Murgasova et al., 2011). Moreover, the linearity of affine transformations facilitates the computation of the average transformation (Alexa, 2002) during spatial normalisation. When local matching is required, the FFD model has the advantage that the resolution of the control point mesh represented as B-splines can be flexibly adjusted (Rueckert et al., 1999), as compared to the other deformable registration methods described in Section 3.2.4. Large control point spacing allows global non-rigid deformations to be modelled, whereas small control point spacing results in highly local non-rigid deformations. Secondly, spatial normalisation is performed that involves transformation of the images into an average space found either by pairwise transformation averaging, or directly by groupwise registration. Lastly, a voxelwise regression model is fitted to the spatially normalised images to describe the average intensity change over time at each location. Fixed-width kernel smoothing is straightforward to apply, but is affected by uneven sampling distributions (Serag et al., 2012). The parametric models have

the advantage of allowing the growth rates and turning points to be extracted analytically, which is not possible using the kernel-based method.

The above pipeline can be applied to brain segmentations and Jacobian determinant maps to generate a tissue probability map and a brain growth atlas, respectively (Aljabar et al., 2008; Habas et al., 2010; Kuklisova-Murgasova et al., 2011). The intensity template, tissue probability map and brain growth atlas correlate different aspects of brain development with GA, and hence are able to predict the morphological age, as compared to the nominal age, based on the actual brain maturation status (Dittrich et al., 2014).

Later in Chapter 6, we will apply FFD non-rigid registration and the logistic regression model to investigate the myelination process in preterm infants. Because spatial co-alignment may be disrupted due to the image contrast change caused by ongoing myelination and the decreasing water content (Huppi and Inder, 2001), we use B-spline FFDs with large control point spacing, so that local contrast variations would not affect alignment of the anatomical structures. The shape of the logistic function depicted in Fig. 3.4.2, that contains a relatively small number of parameters (see Table 3.4.2), is consistent with the expected behaviour of myelin growth.

Chapter 4

Atlas-based segmentation of neonatal brain MRI

4.1 Introduction

Automated segmentation of neonatal brain MRI have great significance in understanding brain maturation patterns in large-scale studies of the newborns (Boardman et al., 2006; Ball et al., 2012; Young et al., 2015). However, challenges arise due to the partial volume (PV) effect, low tissue contrast and rapid brain growth. As a result, spatial prior information becomes helpful when applying adult brain segmentation methods to neonatal data (Kuklisova-Murgasova et al., 2007).

In this chapter, we develop a neonatal brain MRI segmentation method by adapting the Statistical Parametric Mapping (SPM) software¹ (version SPM8) (Ashburner and Friston, 2005) to use a dedicated neonatal brain atlas² (Kuklisova-Murgasova et al., 2011). The SPM segmentations are post-processed to correct PV misclassifications and to segment myelinated white matter which is not included in the atlas. We participated in the MICCAI NeoBrainS12 challenge³ (Isgum et al., 2015) where the proposed method was applied to different image sets of preterm infants, and evaluated together with four other methods.

The work in this chapter has been published in *Medical Image Analysis* as part of the MICCAI NeoBrainS12 challenge (Isgum et al., 2015).

¹<http://www.fil.ion.ucl.ac.uk/spm>

²<http://brain-development.org/brain-atlases>

³<http://neobrain12.isi.uu.nl>

Image set		TR (ms)	TE (ms)	In-plane voxel size (mm ²)	Slice thickness (mm)
Axial (40 weeks GA)	T ₁ w	9.4	4.6	0.35×0.35	2.0
	T ₂ w	6293	120	0.35×0.35	2.0
Coronal (30 weeks GA)	T ₁ w	9.4	4.6	0.34×0.34	2.0
	T ₂ w	10085	120	0.34×0.34	2.0
Coronal (40 weeks GA)	T ₁ w	9.5	4.6	0.78×0.78	1.2
	T ₂ w	48475	150	0.35×0.35	1.2

Table 4.1: Acquisition parameters for three sets of T₁w and T₂w brain MR images of preterm infants provided in the NeoBrainS12 challenge. Abbreviations: GA—gestational age, TR—repetition time, TE—echo time.

The chapter is organised as follows. Section 4.2 introduces the challenge data and the 4D probabilistic neonatal brain atlas (Kuklisova-Murgasova et al., 2011). In Section 4.3, we briefly describe the SPM segmentation tool developed by Ashburner and Friston (2005), followed by a detailed explanation of our approach for refining the initial SPM segmentations. In Sections 4.4 and 4.5, we present the final segmentations and discuss the evaluation results.

4.2 Materials

4.2.1 NeoBrainS12 data

We received three image sets of preterm infants without brain injuries from the NeoBrainS12 challenge: 1) axial scans acquired at 40 weeks gestational age (GA), 2) coronal scans acquired at 30 weeks GA, and 3) coronal scans acquired at 40 weeks GA. Each set consists of T₁w and T₂w brain images for three subjects acquired on a Philips 3 T system using the sequence parameters listed in Table 4.1.

4.2.2 Segmentation protocol for manual annotations

The task of the challenge is to segment eight tissue classes in the neonatal brain, including the cortical grey matter (CoGM), basal ganglia and thalami (BGT), unmyelinated white

matter (UWM), myelinated white matter (MWM), brainstem (BS), cerebellum (CRB), ventricles (VEN) and cerebrospinal fluid in the extracerebral space (ECSF). Manual annotations, which are not publicly available, were performed on T_2w images according to the following segmentation protocol.

The CoGM, UWM and ECSF can be distinguished from the surrounding tissues based on their intensities. The hippocampi and amygdalae in the sub-cortex are included in the CoGM. The borders of BGT, BS and CRB are clearly visible using the T_2w sequence due to the compact shapes. The lateral, third and fourth ventricles form the VEN. The manually annotated MWM includes regions that show higher T_1w intensities and lower T_2w intensities than the surrounding tissues in the regions of BGT and BS. Thus both T_1w and T_2w images are used for labelling MWM.

4.2.3 Probabilistic neonatal brain atlas

The 4D probabilistic neonatal brain atlas constructed by Kuklisova-Murgasova et al. (2011) provides brain templates as well as individual probability maps for six tissue types from 29 to 44 weeks GA, including the CoGM, deep grey matter (DGM), UWM, BS, CRB and cerebrospinal fluid (CSF). We manually modify the atlas at the required 30 and 40 weeks GA in order to match the segmentation protocol described in Section 4.2.2.

The CSF probability map is divided into separate probability maps for the VEN and ECSF. The DGM in the atlas incorporates the basal ganglia, thalami, hippocampi and amygdalae. Therefore, we manually extract the probability map of BGT from that of DGM, and add the remaining hippocampi and amygdalae to the probability map of CoGM. Fig. 4.1 shows the T_2w brain template and the probability maps after amendments at 40 weeks GA. Currently, there is no probability map available for MWM.

4.3 Method

The overall pipeline proposed for automated segmentation of T_2w neonatal brain MRI is shown in Fig. 4.2. T_1w images are only used for segmenting MWM. We first exclude

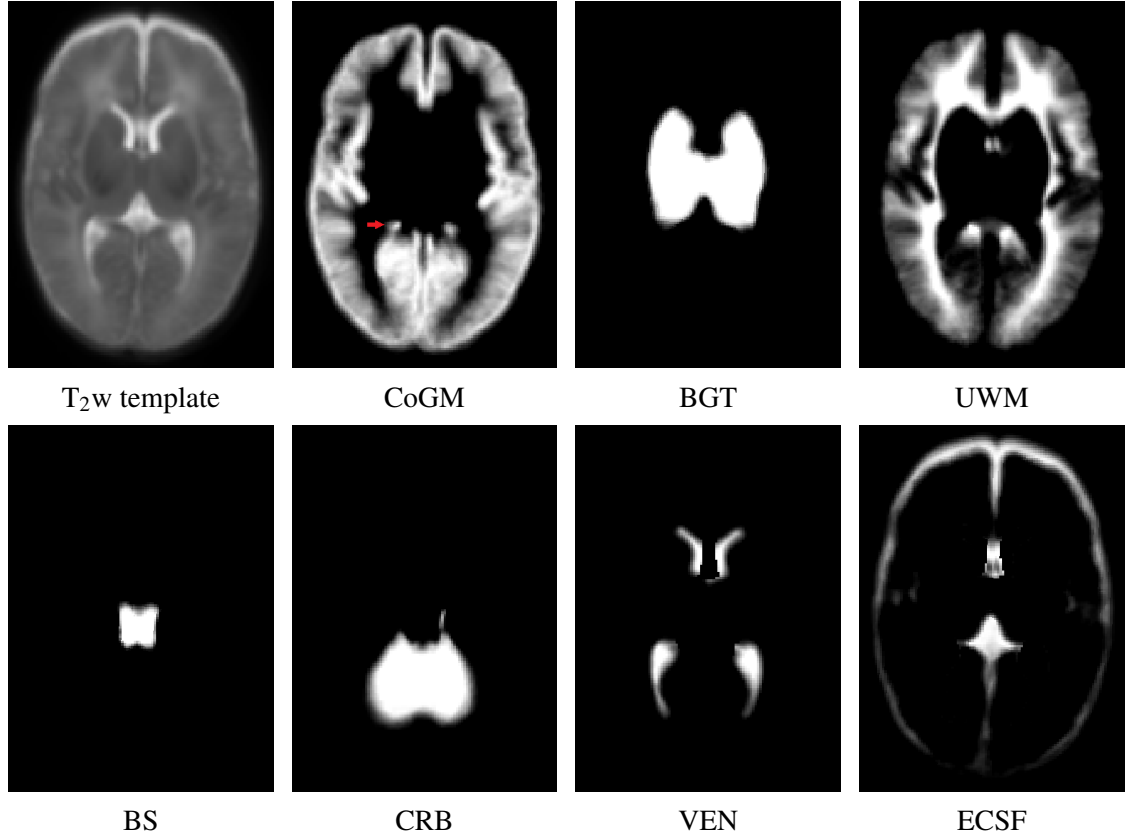


Figure 4.1: T₂w brain template and the probability maps of all the tissue classes except MWM after amendments at 40 gestational weeks. The red arrow points out the hippocampus moved manually from the original probability map of deep grey matter to the CoGM according to the challenge segmentation protocol. Currently, there is no probability map available for MWM. Abbreviations: CoGM—cortical grey matter, BGT—basal ganglia and thalami, UWM—unmyelinated white matter, MWM—myelinated white matter, BS—brainstem, CRB—cerebellum, VEN—ventricles, ECSF—cerebrospinal fluid in the extracerebral space.

the non-brain tissues during image pre-processing. We then adapt the SPM software (Ashburner and Friston, 2005) to the neonatal brain atlas (Kuklisova-Murgasova et al., 2011), and obtain the initial segmentations of all the tissue classes except MWM. Lastly, we refine the SPM segmentations by correcting misclassified PV voxels and segmenting MWM from the regions of BGT and BS using a multivariate Gaussian mixture model (GMM).

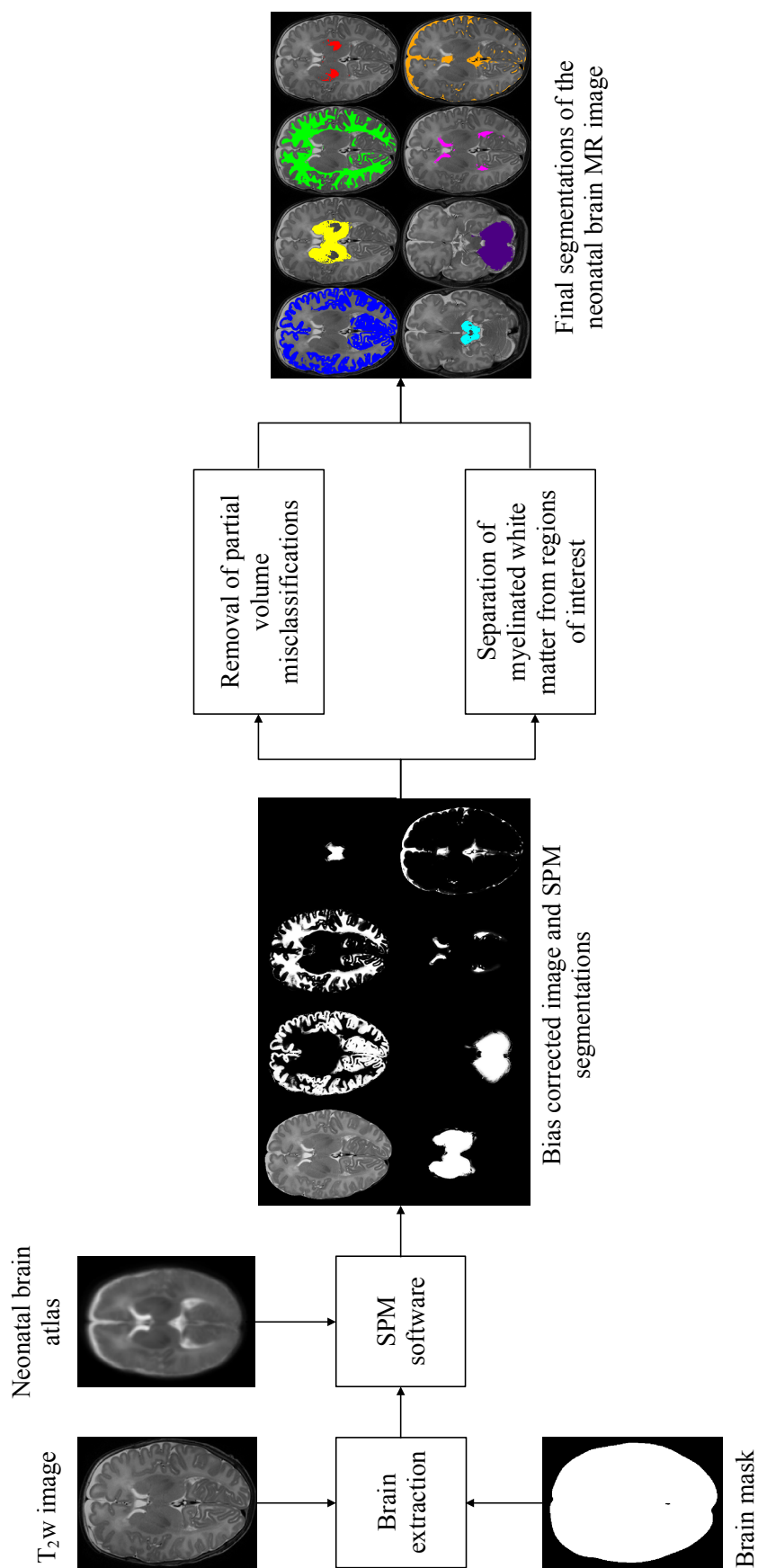


Figure 4.2: Schematic flow of the proposed method for automated segmentation of neonatal brain MRI.

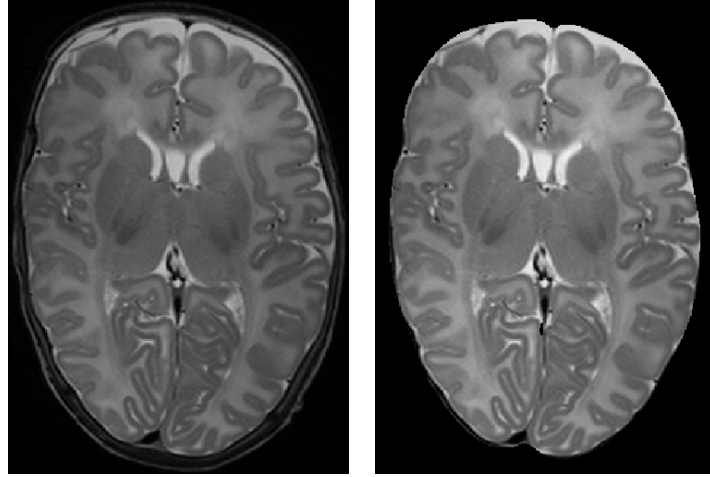


Figure 4.3: T₂w axial brain image for a subject at 40 gestational weeks before and after brain masking.

4.3.1 Image pre-processing

The T₂w brain template (see Fig. 4.1) at the required GA is first brought into approximate alignment with the T₂w image of each individual subject using the SPM affine registration tool *Coregister: Estimate & Reslice*. We choose normalised mutual information (NMI) as the objective function and two sampling distances of 1 mm and 0.5 mm to allow a coarse registration followed by a finer one. The affinely registered template is then deformed with respect to the T₂w image of each subject using the SPM non-rigid registration tool *Normalise: Estimate & Write*, where small deformations are parameterised using a linear combination of approximately 1000 3D low-frequency cosine transformation bases. Only the bases of periods longer than the cut-off, which we set to 25 mm, are used to describe the deformations.

Lastly, the resultant affine and non-rigid transformations are applied to transform the brain mask provided with the atlas from the standard space to each subject's space in order to exclude the non-brain tissues for the subsequent segmentation algorithm, as illustrated in Fig. 4.3.

4.3.2 Unified segmentation in SPM

As explained in Section 2.2.4.1 in Chapter 2, the SPM segmentation tool *Segment* developed by Ashburner and Friston (2005) enables tissue classification, bias field correction and non-rigid atlas registration to be achieved within a single probabilistic framework. The log-likelihood objective function is optimised locally using an iterated conditional modes (ICM) approach (Besag, 1986). At each iteration, three parameter groups for the GMM, bias field and atlas deformations are estimated one group at a time, while fixing the others at their current best values. The GMM parameters are updated using the EM algorithm (Dempster et al., 1977), and the Levenberg-Marquardt algorithm (Levenberg, 1944) is used to estimate the bias field and atlas deformations.

We amend the SPM segmentation tool to take the skull-stripped image and the probability maps of seven tissue classes at the required GA as inputs, including the CoGM, BGT, UWM, BS, CRB, VEN and ECSF. As explained in Section 4.2.3, the original probability maps have been modified to comply with the challenge segmentation protocol. MWM will be segmented separately from the regions of BGT and BS as no probabilistic spatial prior is currently available. We use two Gaussians to represent the intensity distribution of each tissue class and three Gaussians for the background (BKG). This is because the real intensity distributions become non-Gaussian in the presence of PV voxels that contain signal contributions from more than one type of tissue. In order to distinguish the intensity variations that arise due to artefacts of the MR scanning process from those due to different tissue properties, prior knowledge about the distribution of the bias field can be incorporated in the Bayesian framework. For example, because the intensity non-uniformity in the NeoBrainS12 challenge data is relatively smooth, we select a large full width at half maximum (FWHM) cut-off of 80 mm and extremely light regularisation of 0.00001 for the bias field model. Lastly, we set the sampling distance to 0.5 mm and keep the remaining settings at their default values.

We then obtain the bias-corrected image as well as the posterior probability maps of the seven tissue classes that are further converted into hard segmentations using the maximum-vote rule. Even though multiple Gaussians have been used to model the PV

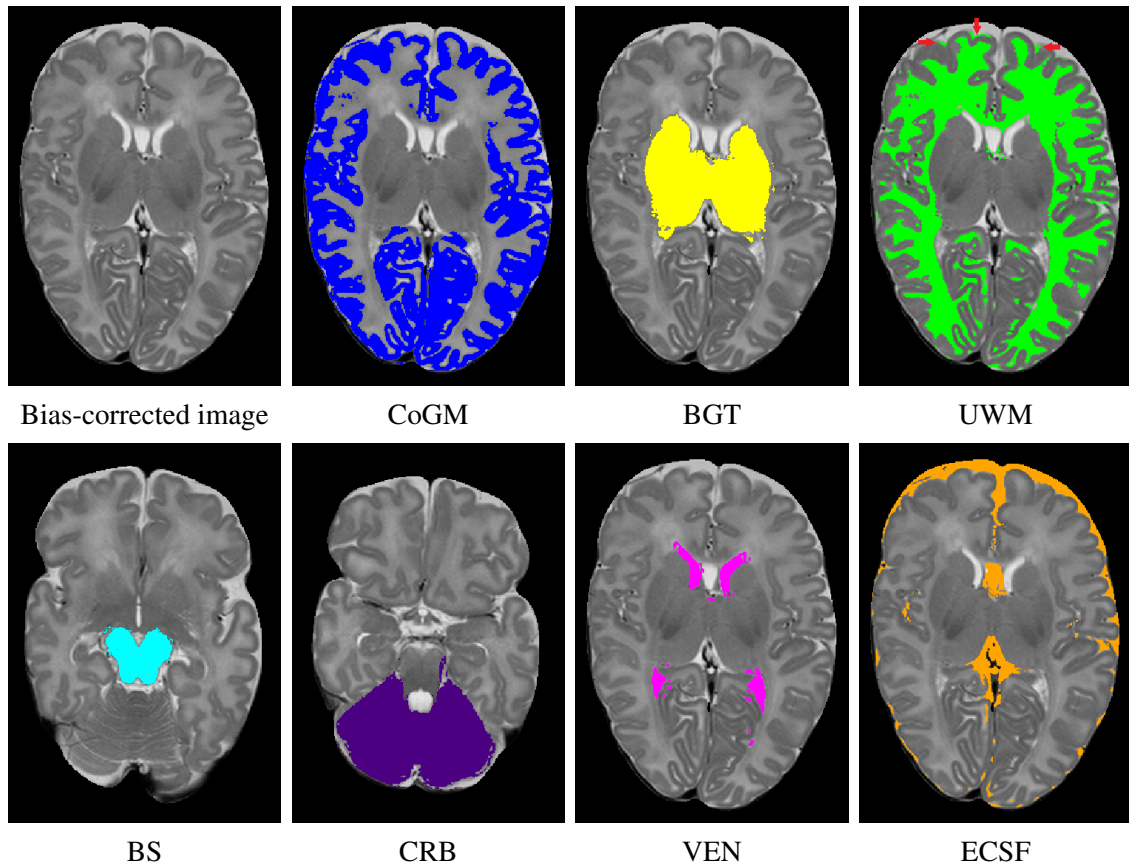


Figure 4.4: Bias-corrected T₂w image and the initial SPM segmentations of all the tissue classes except MWM for a subject in the axial image set scanned at 40 gestational weeks. Substantial misclassifications occur across tissue boundaries due to partial volume effects, especially those mislabelled as UWM at the CoGM/ECSF interface (pointed out by the red arrows). The current BGT and BS segmentations still contain MWM which will be segmented separately using a purely intensity-based approach. Abbreviations: CoGM—cortical grey matter, BGT—basal ganglia and thalami, UWM—unmyelinated white matter, MWM—myelinated white matter, BS—brainstem, CRB—cerebellum, VEN—ventricles, ECSF—cerebrospinal fluid in the extracerebral space.

effect, substantial misclassifications still occur across tissue boundaries, especially those mislabelled as UWM at the CoGM/ECSF interface. Moreover, the BGT and BS segmentations shown in Fig. 4.4 still contain MWM which will be segmented using a purely intensity-based multivariate GMM in the next section.

4.3.3 Segmentation refinements

4.3.3.1 Connected component labelling

Voxels containing mixed tissues at the boundary between CoGM and ECSF have intensities close to those of UWM, and are hence frequently misclassified as UWM by the intensity-based segmentation model in SPM (Ashburner and Friston, 2005) despite the incorporation of probabilistic spatial priors. We use the MATLAB function *bwconncomp*, whose 3D connectivity is restricted to six connected neighbours, to identify the largest connected component in the initial SPM segmentation of UWM as the true segmentation. The remaining voxels in the initial UWM segmentation are reclassified as ECSF. Similarly, we find the largest connected component in the initial CoGM segmentation, and reclassify the remaining PV voxels as ECSF at the boundary between ECSF and BKG.

4.3.3.2 Segmentation of myelinated white matter

According to the anatomical prior knowledge (Counsell et al., 2002; Rutherford, 2002), myelination of the newborn brain originates in the inferior cerebellar peduncles as early as 25 weeks GA, followed by the inferior collicoli, posterior BS and thalamic nuclei. Therefore, we choose the SPM binary segmentations of BGT and BS to form individual regions of interest (ROIs) for assessing MWM during the perinatal period.

We define two classes in each ROI: MWM and BKG. The MATLAB (R2013b)⁴ functions *fitgmdist* and *cluster* are used to fit a two-class multivariate GMM to the T_1w and T_2w intensities in each individual ROI of the aligned images through the EM algorithm (Dempster et al., 1977), as illustrated in Fig. 4.6. The initial clusters are assigned based on the fact that MWM shows higher signal intensities than the surrounding tissues on T_1w images, and lower in the T_2w contrast.

The conditional probability density function (PDF) of class k in the multivariate GMM

⁴<http://uk.mathworks.com/>

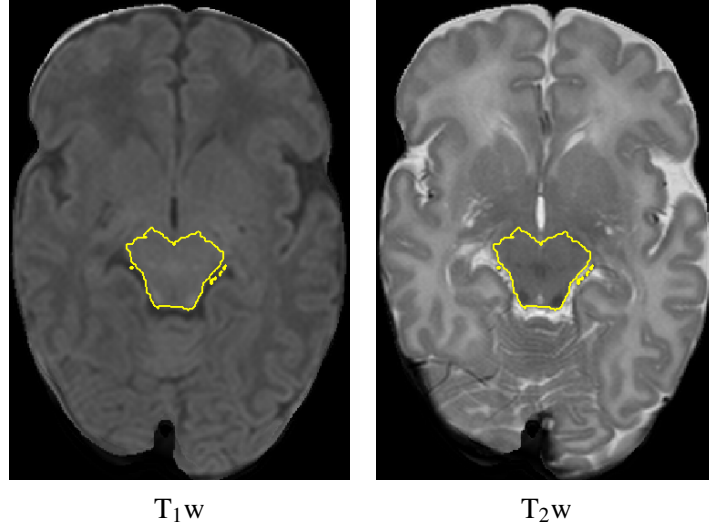


Figure 4.5: T₁w and T₂w axial brain images for a subject at 40 gestational weeks. The SPM segmentation of the brainstem (yellow label) is used as the binary mask of a region of interest for segmenting myelinated white matter.

can be written as:

$$G(\mathbf{y}_i, \boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k) = \frac{1}{(2\pi)^{\frac{R}{2}} |\boldsymbol{\Sigma}_k|^{\frac{1}{2}}} e^{-\frac{1}{2}(\mathbf{y}_i - \boldsymbol{\mu}_k)^T \boldsymbol{\Sigma}_k^{-1} (\mathbf{y}_i - \boldsymbol{\mu}_k)} \quad (4.1)$$

The symbol R is the number of modalities, and $\boldsymbol{\Sigma}_k$ indicates the covariance matrix of class k for all the modalities. We denote $\mathbf{y}_i = \{y_{i1}, y_{i2}, \dots, y_{iR}\}$ where y_{ir} is the observed intensity at voxel i in modality r , and $\boldsymbol{\mu}_k = \{\mu_{k1}, \mu_{k2}, \dots, \mu_{kR}\}$ with μ_{kr} representing the mean intensity of class k in modality r . We apply the EM algorithm as follows:

- E-step: Calculate the probability p_{ik} that voxel i belongs to class k :

$$p_{ik}^{(m+1)} = \frac{G(\mathbf{y}_i, \boldsymbol{\mu}_k^{(m)}, \boldsymbol{\Sigma}_k^{(m)}) c_k^{(m)}}{\sum_{k'=1}^K G(\mathbf{y}_i, \boldsymbol{\mu}_{k'}^{(m)}, \boldsymbol{\Sigma}_{k'}^{(m)}) c_{k'}^{(m)}} \quad (4.2)$$

where c_k is the mixing proportion of class k , K the number of classes, and m the current EM iteration number.

- M-step: Maximum likelihood estimation yields the following equations for updating

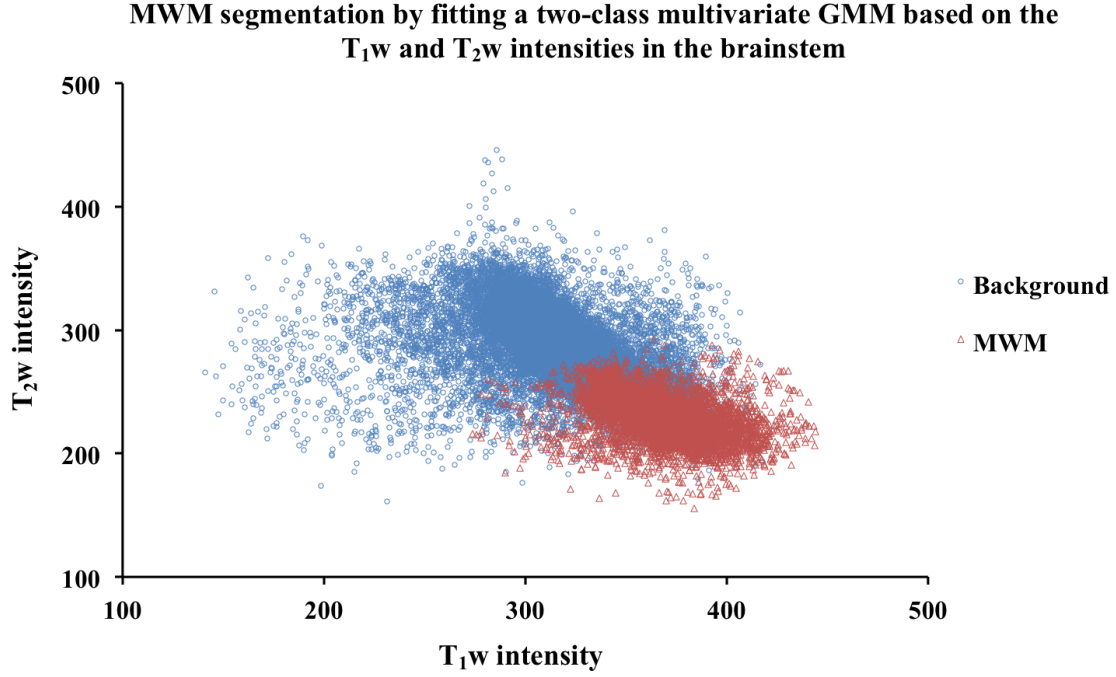


Figure 4.6: Clusters of myelinated white matter (MWM) and the background tissue in the brainstem for a subject at 40 gestational weeks obtained by fitting a two-class multivariate Gaussian mixture model (GMM) based on the T₁w and T₂w intensities.

the GMM parameters:

$$\mu_{kr}^{(m+1)} = \frac{\sum_{i=1}^N p_{ik}^{(m+1)} y_{ir}}{\sum_{i=1}^N p_{ik}^{(m+1)}} \quad (4.3)$$

$$\left(\Sigma_k^{(m+1)} \right)_{rs} = \frac{\sum_{i=1}^N p_{ik}^{(m+1)} (y_{ir} - \mu_{kr}^{(m+1)}) (y_{is} - \mu_{ks}^{(m+1)})}{\sum_{i=1}^N p_{ik}^{(m+1)}} \quad (4.4)$$

$$c_k^{(m+1)} = \frac{\sum_{i=1}^N p_{ik}^{(m+1)}}{N} \quad (4.5)$$

where N is the number of voxels in the ROI, and $\left(\Sigma_k^{(m+1)} \right)_{rs}^2$ the element in the covariance matrix between modalities r and s .

Below we summarise the algorithm for segmenting MWM on neonatal brain MR images:

Step 1. Align the T₁w image with the corresponding T₂w image using the SPM rigid registration tool *Coregister: Estimate & Reslice*.

Step 2. Extract the ROI masks from the atlas-based SPM segmentations of BGT and BS, as explained in Section 4.3.2.

Step 3. Calculate the mean T_1w and T_2w intensities in the region of BGT.

Step 4. Initialise the MWM segmentation in the BGT with voxels that satisfy the following criteria:

(i) The T_1w intensity is higher than the mean T_1w intensity.

(ii) The T_2w intensity is lower than the mean T_2w intensity.

Step 5. Perform a cluster analysis using the MATLAB (R2013b) functions *fitgmdist* and *cluster* that fit a two-class multivariate GMM to the T_1w and T_2w intensities in the BGT of the aligned images.

Step 6. Repeat Steps 3 to 5 in the region of BS, and combine MWM segmented in both ROIs to form the final segmentation.

4.4 Results

We processed all the NeoBrainS12 challenge images described in Section 4.2.1 using the proposed algorithm, and the resultant segmentations of the eight tissue classes for a representative subject in each of the three image sets are shown in Figs. 4.7, 4.8 and 4.9 respectively.

The automatic brain segmentations were evaluated with respect to the manually annotated ground truth using the Dice coefficient (DC) (Dice, 1945) and Hausdorff distance (HD) (Huttenlocher et al., 1993). Our method was compared with four other neonatal brain segmentation algorithms in the NeoBrainS12 challenge (Isgum et al., 2015). Figs. 4.10 and 4.11 show the DCs and HDs, respectively, obtained using different methods for each tissue class in all three image sets.

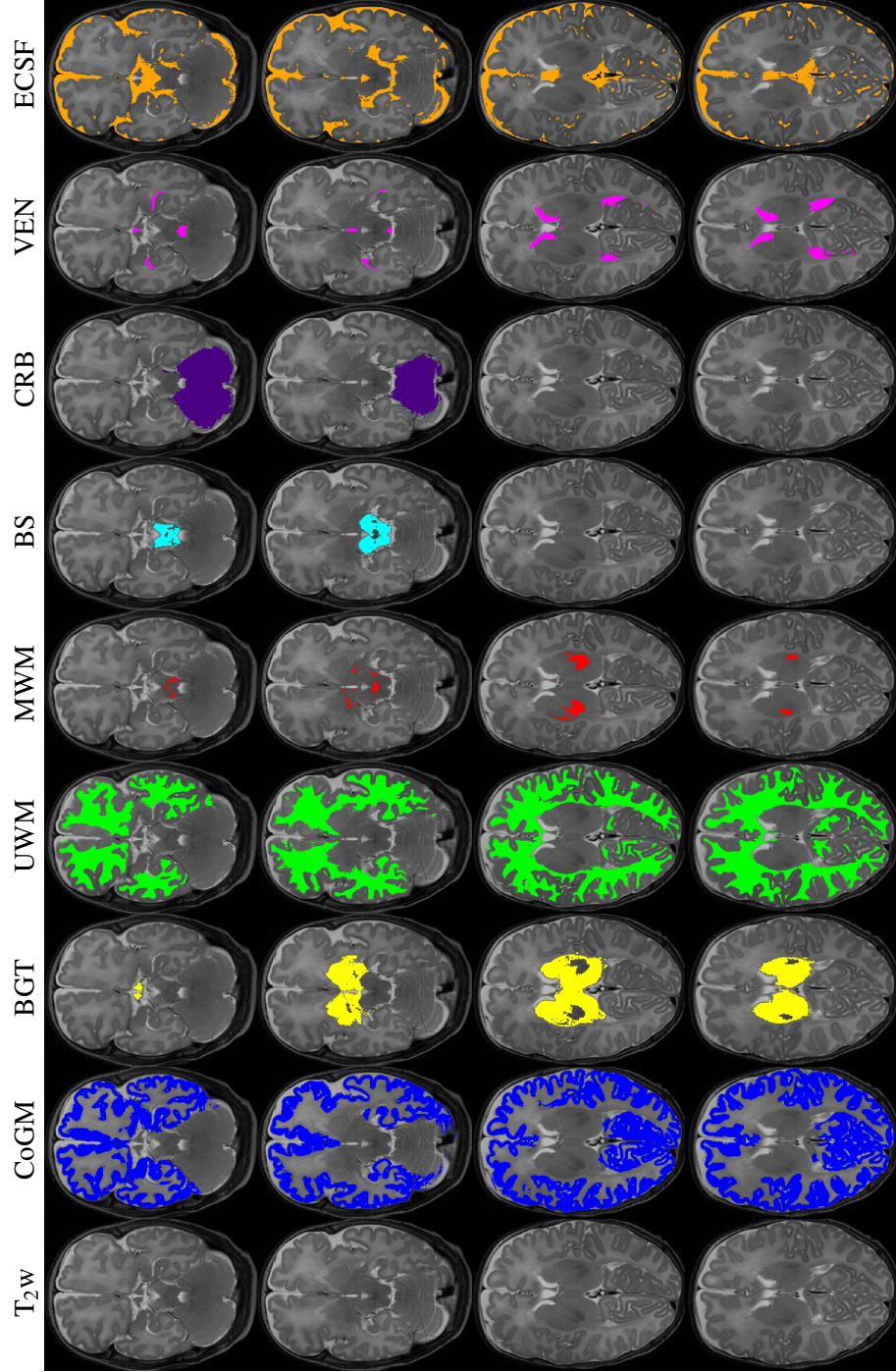


Figure 4.7: Automatic brain segmentations obtained using the proposed algorithm for a subject in the axial image set acquired at 40 gestational weeks. The columns from left to right show the eight tissue classes, and the rows from top to bottom correspond to different axial positions. Abbreviations: CoGM–cortical grey matter, BGT–basal ganglia and thalami, UWM–unmyelinated white matter, MWM–myelinated white matter, BS–brainstem, CRB–cerebellum, VEN–ventricles, ECSF–cerebrospinal fluid in the extracerebral space.

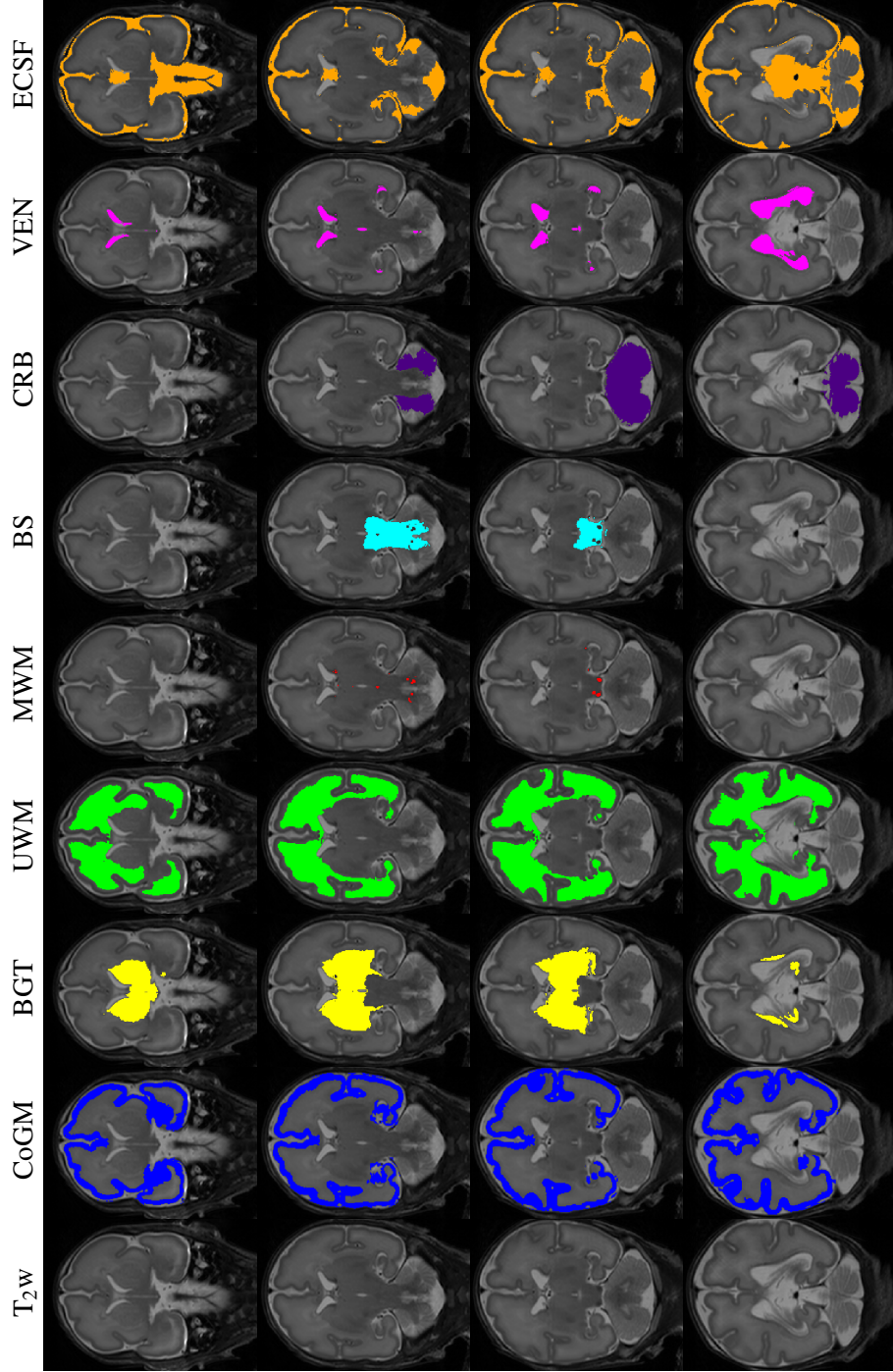


Figure 4.8: Automatic brain segmentations obtained using the proposed algorithm for a subject in the coronal image set acquired at 30 gestational weeks. The columns from left to right show the eight tissue classes, and the rows from top to bottom correspond to different coronal positions. The abbreviations of the tissue classes can be found in Fig. 4.7.

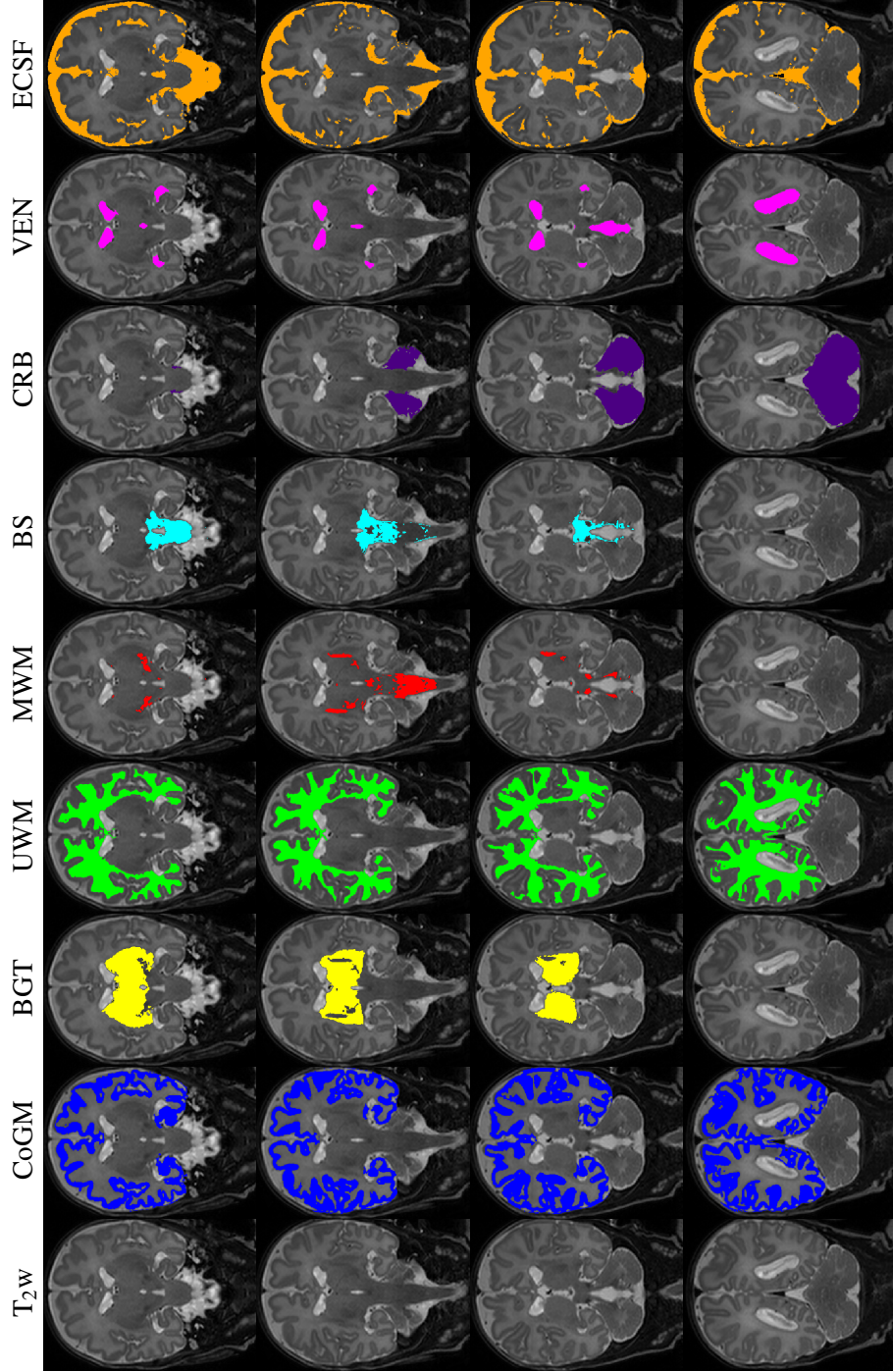


Figure 4.9: Automatic brain segmentations obtained using the proposed algorithm for a subject in the coronal image set acquired at 40 gestational weeks. The columns from left to right show the eight tissue classes, and the rows from top to bottom correspond to different coronal positions. The abbreviations of the tissue classes can be found in Fig. 4.7.

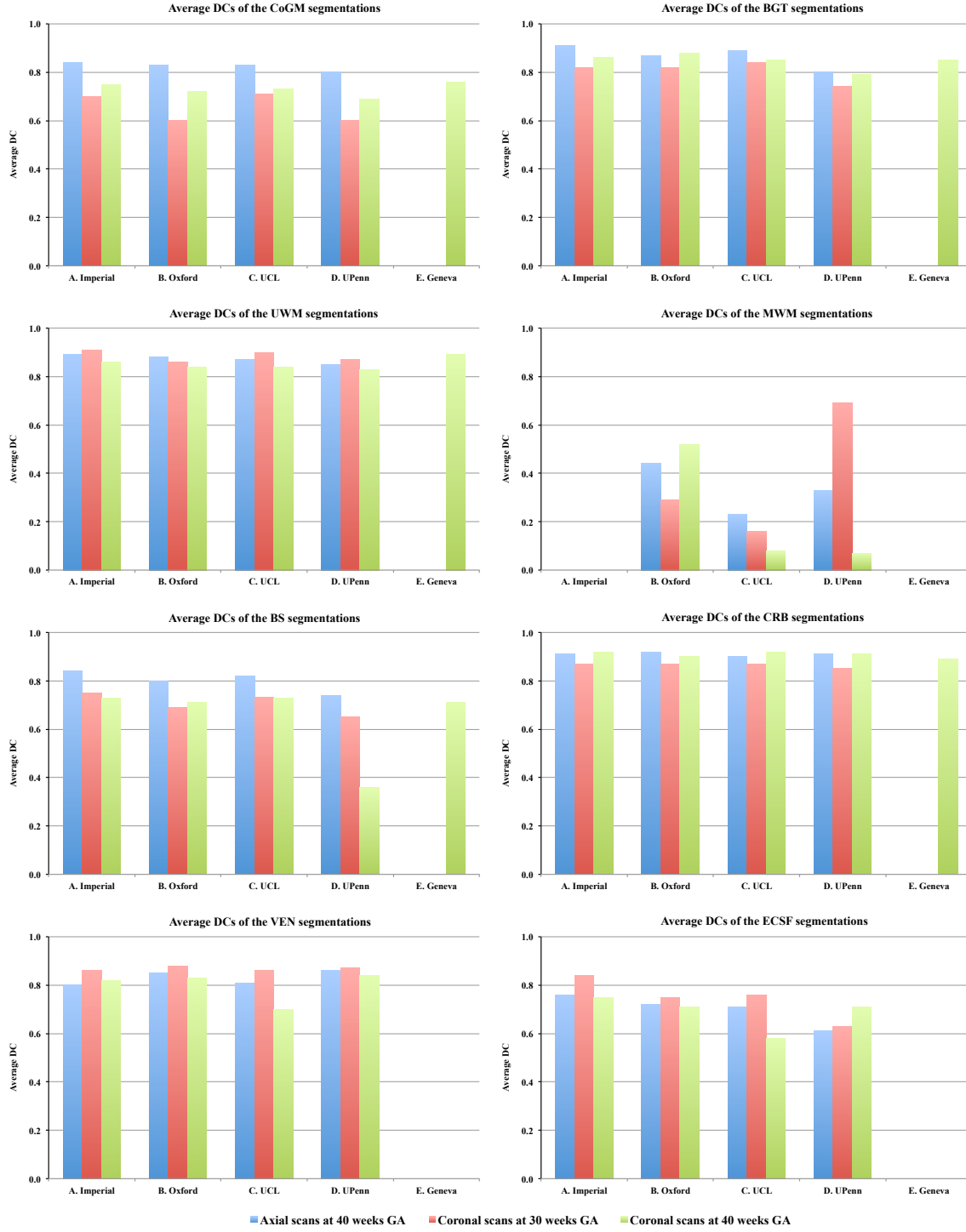


Figure 4.10: Dice coefficients (DCs) of the automatic segmentations for eight tissue classes averaged over three preterm infants in each image set. The participating methods in the NeoBrainS12 challenge achieve high DCs for general neonatal brain segmentation using well-developed spatial prior knowledge. In contrast, myelinated white matter (MWM) is poorly segmented due to insufficient spatial prior information. The error bars cannot be plotted as the DCs of individual subjects are not publicly available. The algorithms are briefly described in Section 4.5.1 and the abbreviations of the tissue classes can be found in Fig. 4.7.

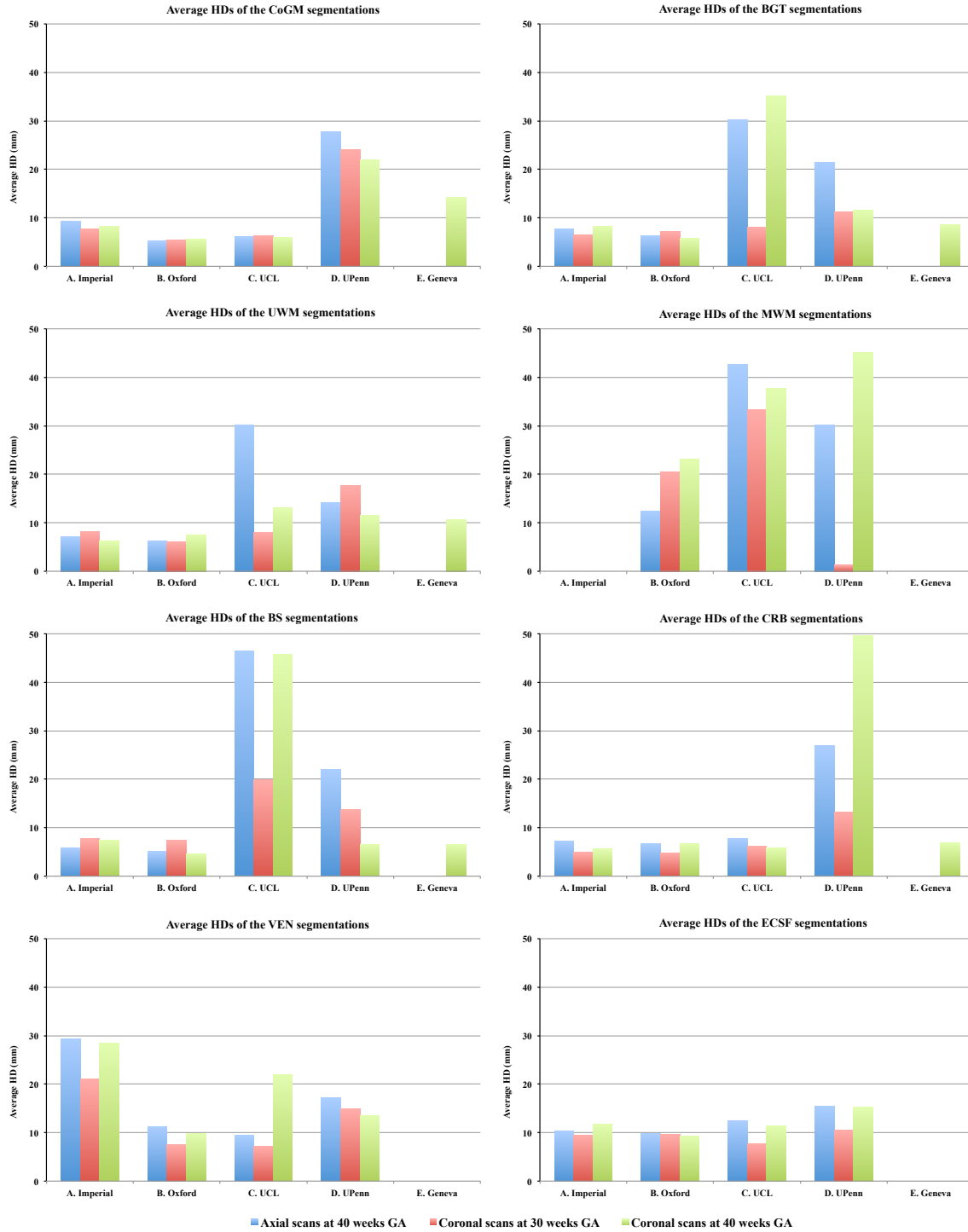


Figure 4.11: Hausdorff distances (HDs) of the automatic segmentations for eight tissue classes averaged over three preterm infants in each image set. Besides the underperformance for myelinated white matter (MWM) in all the methods in the NeoBrainS12 challenge, Methods C and D yield large HDs and hence large local shape variations for the other general brain structures despite the high Dice overlap. The error bars cannot be plotted as the HDs of individual subjects are not publicly available. The algorithms are briefly described in Section 4.5.1 and the abbreviations of the tissue classes can be found in Fig. 4.7.

4.5 Discussion and conclusion

4.5.1 NeoBrainS12 challenge

The performance of our method, called Method B in Isgum et al. (2015), was stable and accurate with respect to the manual annotations across different image sets and different tissue classes except MWM. The best segmentations were achieved for the axial image set at 40 weeks GA, where the highest DC was 0.92 for the CRB and the lowest 0.72 for the ECSF excluding MWM. Similar results were observed in the other methods because of the compact shape of the CRB and the complex morphology of the ECSF. Since none of the currently available neonatal brain atlases includes myelin (Kuklisova-Murgasova et al., 2011; Serag et al., 2012), we segmented MWM by fitting a two-class multivariate GMM to the T_1w and T_2w intensities after alignment in the ROIs of BGT and BS separately. However, the best DC of the MWM only reached 0.52 for the coronal image set at 40 weeks GA, where the HD was over 20 mm. Despite the incorporation of multi-spectral MRI data, the MWM was considerably over-estimated, and we will show in Chapter 5 that this can be rectified by PV modelling.

As described in Chapter 2, Method A was developed by Makropoulos et al. (2014) at the Imperial College London. Spatial priors for the EM-MRF segmentation algorithm (Van Leemput et al., 1999b) were obtained by non-rigid registration (Rueckert et al., 1999) of the ALBERTs' atlas (Gousias et al., 2012) to the subject's T_2w image. This method achieved consistent DCs well above 0.7 for all the tissue classes, except that it did not segment MWM which is not included in the atlas.

Method C proposed by Melbourne et al. (2012) at the University College London used T_2w images only. A probabilistic spatial prior for incorporation in the method of Cardoso et al. (2011) was simulated by averaging the manual annotations of two training subjects that were provided additionally by the NeoBrainS12 challenge for each image set. Due to the small number of training manual annotations, this approach poorly segmented MWM and yielded large HDs for the BGT, UWM and BS.

Method D presented by Wu and Avants (2012) from the University of Pennsylva-

nia obtained the spatial priors using symmetric diffeomorphic registration (Avants et al., 2008) of the probabilistic neonatal brain atlas constructed by Serag et al. (2012). Both T_1w and T_2w images were used in the subsequent EM-MRF algorithm (Van Leemput et al., 1999b). MWM segmented by manually annotating the atlas template was subject to extremely large variations across different image sets.

Method E was developed by Gui et al. (2012) at Geneva University Hospital. A series of techniques, such as the watershed algorithm (Lotufo and Falcao, 2000), seeded region growing (Adams and Bischof, 1994) and active contour segmentation (Chan et al., 2000), were applied to both T_1w and T_2w images. Even though no probabilistic atlas or manual annotations were required, this method had very limited applicability and only segmented five brain structures in the coronal image set acquired at 40 weeks GA.

Therefore, these state-of-the-art methods in the NeoBrainS12 challenge all showed great promise in general neonatal brain segmentation, such as the CoGM, UWM, CSF, BGT, BS and CRB, for which probabilistic spatial priors or manual annotations are usually available in practice. In contrast, MWM segmentation was seriously underperformed due to insufficient spatial prior information on neonatal brain MR images. Our attempt to identify MWM using a two-class multivariate GMM proved to be of insufficient accuracy without PV estimation.

4.5.2 Segmentation protocol for myelinated white matter

Lastly, we would like to call attention to the MWM segmentation protocol and the manual annotations with respect to which the automatic MWM segmentations were evaluated in the NeoBrainS12 challenge. As defined in Section 4.2.2, MWM in the perinatal period includes the regions that show higher T_1w intensities and lower T_2w intensities than the surrounding tissues in the ROIs of BGT and BS (Isgum et al., 2015). This definition, however, completely depends on image intensities without incorporating prior knowledge of myelination that has been obtained from histological studies (Yakolev and Lecours, 1967; Rorke and Riggs, 1969) and clinical MRI observations (Counsell et al., 2002; Rutherford, 2002).

Fig. 4.12 shows the T_1w and T_2w images of a training subject provided for the axial image set acquired at 40 weeks GA. According to the challenge segmentation protocol, the densely organised globus pallidus (GP) was labelled as MWM despite the fact that there is no evidence of myelination in the GP at this age (Yakolev and Lecours, 1967; Rorke and Riggs, 1969; Counsell et al., 2002; Rutherford, 2002). On T_1w and T_2w images, myelination becomes evident in three central brain structures in the region of BGT by the time of 40 weeks GA (Counsell et al., 2002), including the posterior limbs of the internal capsule (PLIC), ventrolateral nuclei (VLN) and subthalamic nuclei (STN). Also, the entire region of the posterior BS was manually annotated as MWM in the training example, but we can actually identify the myelinated BS structures in detail, particularly on the T_2w image.

In reality, the signal properties of the developing brain tissues on the conventional MR images reflect several maturational processes that occur simultaneously, including chemical composition changes, increased densities of neurons and synapses as well as increased myelination (Korogi et al., 1996). In Chapter 5, we will thus first define a segmentation protocol for the tissue that is likely to contain myelin, which we will refer to as myelin-like signals (MLS), based on both image intensities and anatomical knowledge. We will then develop a new segmentation method for MLS that takes the PV effect into account without relying on any probabilistic atlas or manual annotations of myelin.

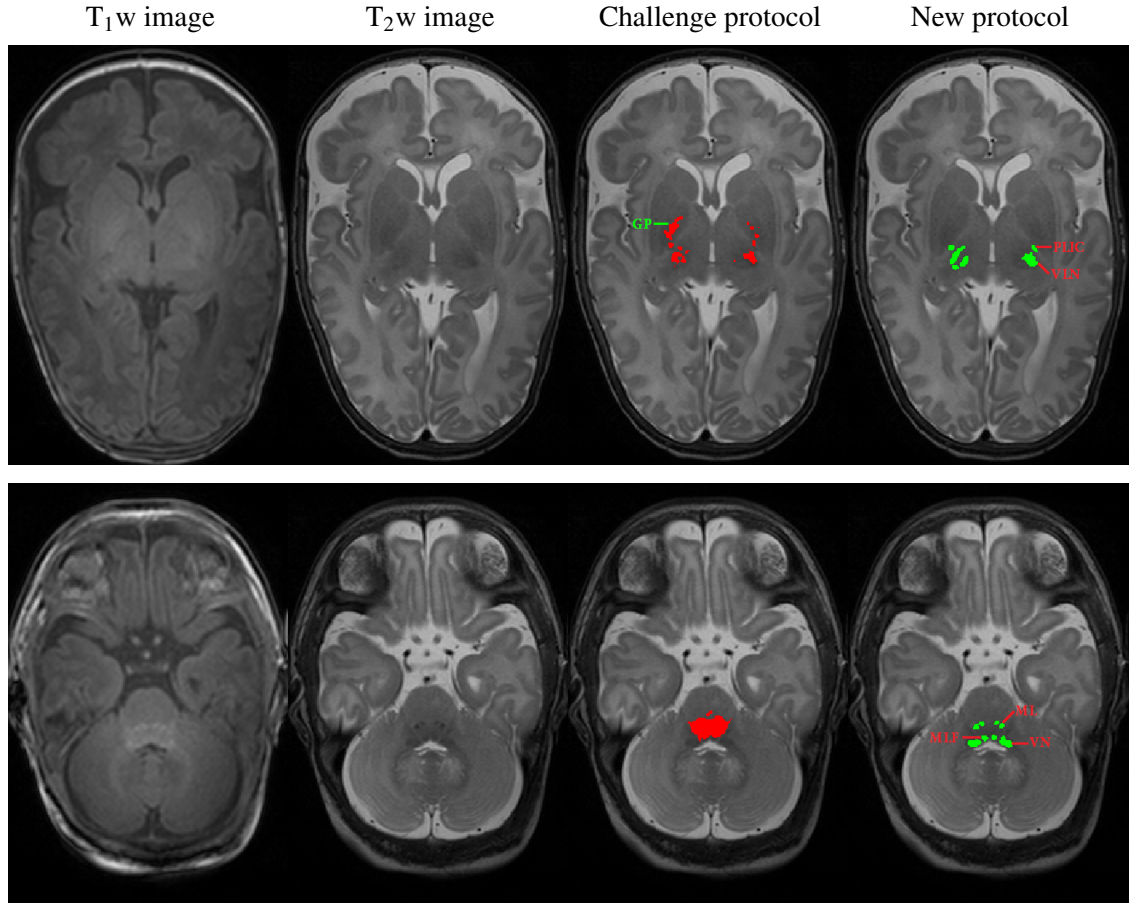


Figure 4.12: Manual annotations of the myelinated tissue delineated according to different segmentation protocols for a training subject at 40 gestational weeks. The segmentation protocol in the NeoBrainS12 challenge (the second column from right) completely depends on image intensities without considering anatomical prior knowledge. As a result, the densely organised globus pallidus (GP) in the basal ganglia and thalami (BGT), which does not contain myelin at this age, and the entire region of the posterior brainstem (BS) have been labelled for myelination. In Chapter 5, we define a new segmentation protocol (the first column from right) based on both image intensities and anatomical knowledge, that would include the posterior limbs of the internal capsule (PLIC) and ventrolateral nuclei (VLN) in the region of BGT, and the medial lemnisci (ML), medial longitudinal fasciculi (MLF) and vestibular nuclei (VN) in the region of BS.

Chapter 5

Partial volume modelling via second-order MRFs: Application to segmentation of myelin-like signals

5.1 Introduction

Myelination is considered to be an important developmental process during human brain maturation and closely correlated with gestational age (GA). As discussed in Chapter 1, direct assessment of myelin requires dedicated imaging, but the conventional MR images routinely acquired during clinical imaging of neonates carry signatures that are thought to be associated with myelination (Barkovich, 1998; Counsell et al., 2002), which we define as myelin-like signals (MLS).

As demonstrated in Chapter 4, while the state-of-the-art methods (Gui et al., 2012; Melbourne et al., 2012; Wang et al., 2012; Wu and Avants, 2012; Makropoulos et al., 2014) in the NeoBrainS12 challenge (Isgum et al., 2015) all show great promise in general neonatal brain segmentation, none of them are directly applicable to MLS due to insufficient spatial prior knowledge of myelination on the conventional brain MR images. We proposed to segment myelinated white matter in the challenge by fitting a two-class multivariate Gaussian mixture model (GMM) to the T_1w and T_2w intensities of the aligned

The work described in this chapter is in preparation for publication.

images in predefined regions of interest (ROIs). However, this approach resulted in significant over-estimation without considering partial volume (PV) voxels. We discussed in Chapter 2 that the current PV estimation methods require prior knowledge of the spatial locations of the composing pure tissues for robust PV modelling (Shattuck et al., 2001; Van Leemput et al., 2003; Tohka et al., 2004; Manjon et al., 2010; Cardoso et al., 2011; Ledig et al., 2012). However, there is very limited spatial prior information available for MLS due to the absence of any probabilistic atlas or manual annotation database.

In this chapter, we propose a novel segmentation method with explicit PV modelling for MLS in a predefined ROI on T₂w neonatal brain MR images. We choose T₂w images because the T₂w contrast is considered better for segmentation in general in neonates. We first obtain the ROI using the atlas-based method proposed in Chapter 4. The subsequent MLS segmentation in the ROI is atlas-free. We introduce an explicit PV class, and generate the spatial priors by configuring the PV locations in relation to MLS and the background (BKG) tissue via second-order MRFs. This approach allows us to distinguish the small proportion of MLS in the presence of substantial PV voxels. We achieved automatic MLS segmentations of high Dice coefficients (DCs) (Dice, 1945) with respect to the manual annotations consistently for 16 preterm infants at one-week intervals between 29 and 44 weeks GA.

This chapter is structured as follows. In Section 5.2, we define an MLS segmentation protocol in two ROIs, namely the thalami and brainstem. Manual annotations are delineated according to this protocol. Section 5.3 presents the proposed MLS segmentation method, followed by the experiments and results in Section 5.4. We conclude this chapter with the discussion in Section 5.5.

5.2 Materials

5.2.1 Subjects and MR acquisition

The image sets acquired for research purposes in the NeoBrainS12 challenge only include nine preterm infants aged at either 30 or 40 weeks GA, and the manual annotations are

not publicly available.

In order to investigate the correlation between myelination and GA, we use a new dataset that consists of 114 preterm infants without brain injuries, scanned between 29 and 44 weeks GA at Hammersmith Hospital, London, UK with ethical approval obtained. The GAs in weeks are converted from the number of days between the conception date and the date when the image was acquired. T₂w fast spin-echo brain images were acquired on a 3 T Philips Intera system with repetition time = 8700 ms, echo time = 160 ms and voxel sizes = 0.86 mm×0.86 mm×1 mm.

5.2.2 Segmentation protocol and manual annotations

We choose the thalami and brainstem as the ROIs for assessing MLS because these are the regions where the majority of myelination occurs between 29 and 44 weeks GA (Counsell et al., 2002; Rutherford, 2002). The automatic procedure through which we obtain the binary ROI masks of the thalami and brainstem for individual subjects is explained in Section 5.3.1.

The MLS segmentation protocol is defined within the ROIs based on both image intensities and anatomical knowledge (Counsell et al., 2002). We delineate three central brain structures in the thalami, including the posterior limbs of the internal capsule (PLIC), ventrolateral nucleus (VLN) and subthalamic nucleus (STN). Particularly the PLIC tracts in the preterm brain, which show MLS at approximately 40 weeks GA on T₂w images, is an important signature that can be used to assess neonatal brain maturation (Bird et al., 1989; Counsell et al., 2002). We define the segmentation protocol in the brainstem mainly based on intensities because the image resolution is insufficient to identify the detailed anatomical structures.

We select 16 preterm infants, each at a different time point, between 29 and 44 weeks GA at one-week intervals for method validation. Manual annotations were performed for the 16 test subjects in the ROIs according to the defined segmentation protocol. An example is shown in Fig. 5.1. Eight of the 16 subjects, from 30 weeks GA onward at two-week intervals, have a repeated manual segmentation delineated by the same rater

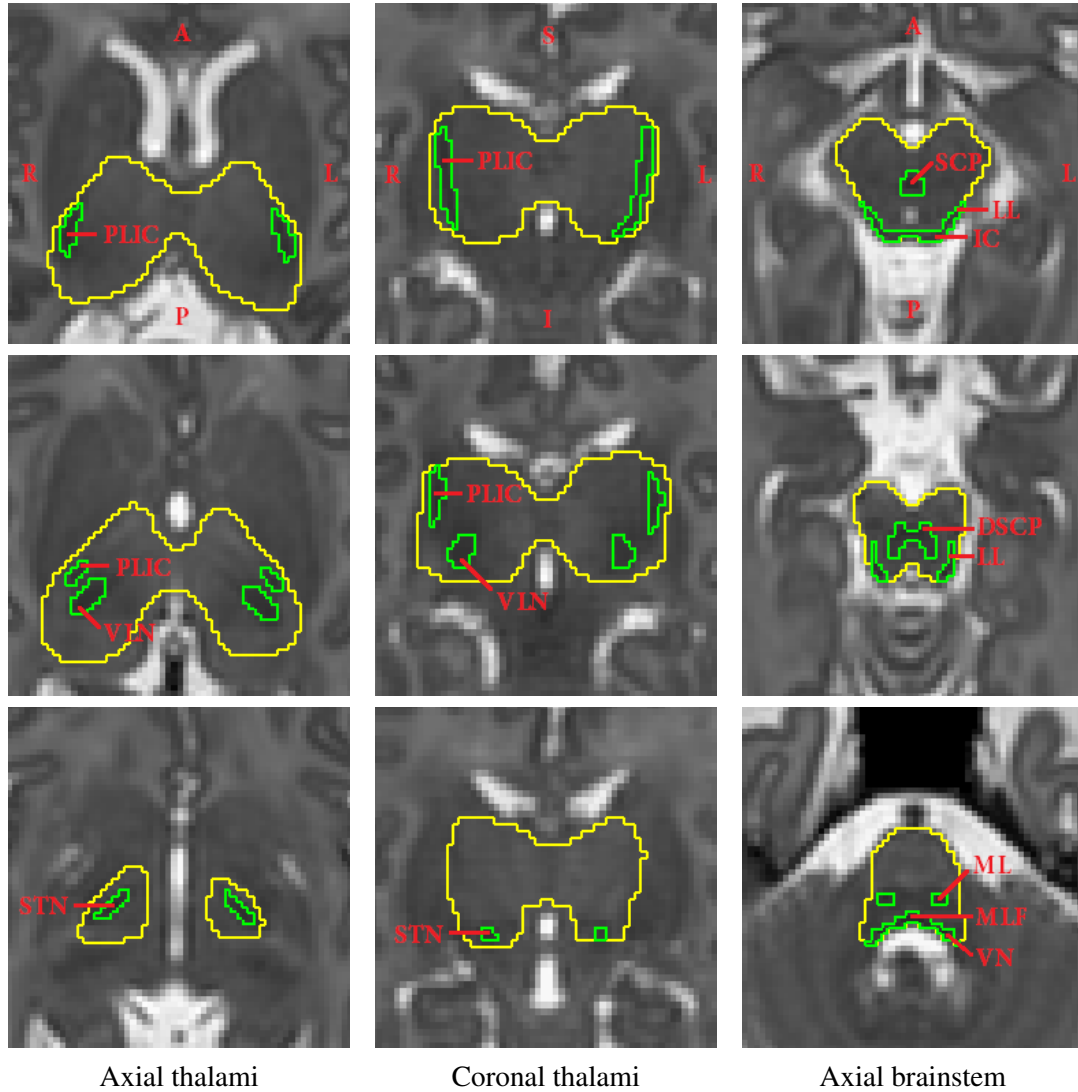


Figure 5.1: Manual annotations of myelin-like signals (green label) in the thalami and brainstem of a subject at 42 gestational weeks, delineated according to the defined segmentation protocol. The regions of interest are labelled in yellow. The columns from left to right show the thalami in the axial and coronal views, and the brainstem in the axial view. Abbreviations: PLIC–posterior limb of the internal capsule, VLN–ventrolateral nucleus, STN–subthalamic nucleus, SCP–superior cerebellar peduncle, DSCP–decussation of the superior cerebellar peduncle, IC–inferior colliculus, LL–lateral lemniscus, ML–medial lemniscus, MLF–medial longitudinal fasciculus, VN–vestibular nucleus.

in order to assess the intra-rater reliability. The quality of these manual annotations was confirmed by a clinical expert.

5.3 Method

The overall pipeline proposed for MLS segmentation on T₂w neonatal brain MR images is shown in Fig. 5.2. In this section, we first explain the procedure of image pre-processing. We then describe a purely intensity-based GMM without PV modelling, followed by the incorporation of an explicit PV class and spatial regularisation in the EM framework. Lastly, we explain the implementation details of the proposed method for obtaining the final binary MLS segmentations.

5.3.1 Image pre-processing

The aim of image pre-processing is to achieve brain extraction and bias field correction, and to obtain the binary ROI masks of the thalami and brainstem for individual subjects.

We remove the non-brain tissues using label propagation (Heckemann et al., 2006) of manually annotated brain masks, and segment the skull-stripped images using the Statistical Parametric Mapping (SPM) software¹ (version SPM8) (Ashburner and Friston, 2005). The spatial priors are provided by a 4D probabilistic neonatal brain atlas² (Kuklisova-Murgasova et al., 2011). Subsequently, we obtain the bias-corrected images as well as the segmentations of the cortical GM, WM, deep grey matter (DGM), brainstem, cerebellum and cerebrospinal fluid (CSF).

We use the automatic SPM segmentations of the brainstem directly as the ROI masks. The DGM segmentations include the basal ganglia, thalami, hippocampi and amygdalae. Since the myelinated thalamic nuclei can have similar intensities to the densely organised GM in the basal ganglia which does not contain myelin between 29 and 44 weeks GA (Counsell et al., 2002), we extract the ROI masks for the thalami from the DGM segmentations in order to prevent misclassifications by the intensity-based segmentation model.

¹<http://www.fil.ion.ucl.ac.uk/spm>

²<http://brain-development.org/brain-atlases>

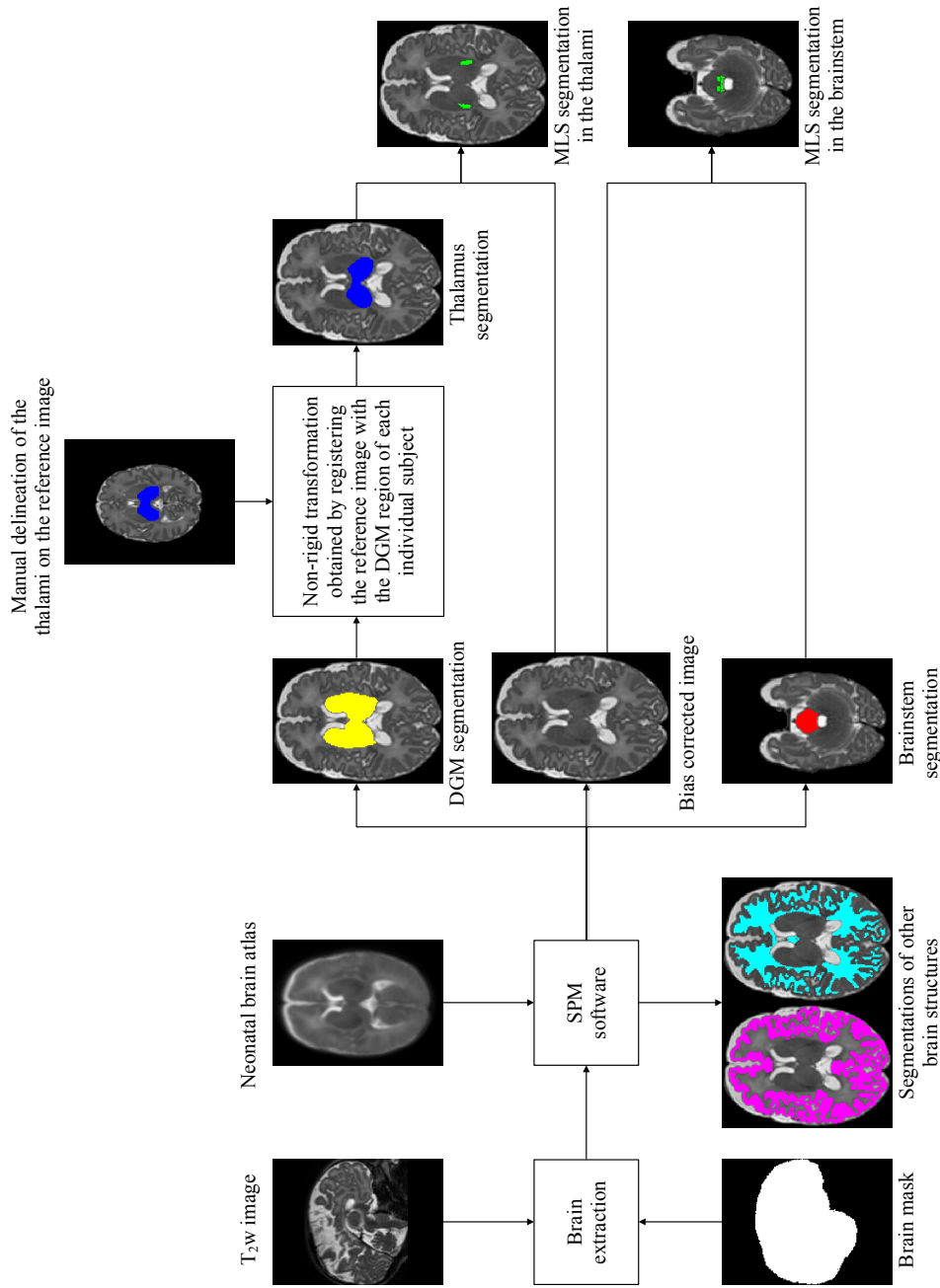


Figure 5.2: Schematic flow of the proposed segmentation approach for myelin-like signals (MLS) on T₂w neonatal brain images. Through image pre-processing, we achieve brain extraction and bias field correction, and obtain the binary masks of the thalami and brainstem for individual subjects. The automatic segmentations of the deep grey matter (DGM), obtained using the Statistical Parametric Mapping (SPM) software, contain the basal ganglia, thalami, hippocampi and amygdalae. We extract the masks for the thalami from the DGM segmentations via label propagation.

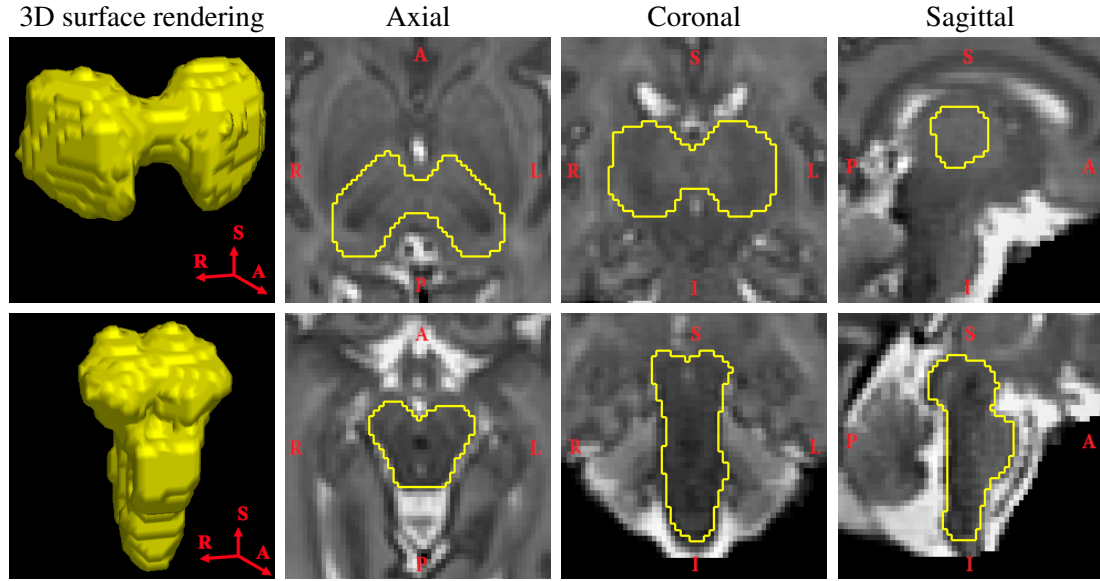


Figure 5.3: Two separate regions of interest (ROIs) labelled in yellow for assessing myelin-like signals in the perinatal period. The first and second rows show the thalami and brainstem, respectively, for a subject at 34 gestational weeks. The columns from left to right show the ROIs in 3D surface rendering, axial, coronal and sagittal views.

We first manually delineate the thalami on the T_2w image of a single reference subject at 36 weeks GA. We register the reference image with the dilated DGM region of each subject using free-form deformation (FFD) non-rigid registration (Rueckert et al., 1999), and then transform the manual delineation of the thalami from the reference space to each individual subject’s space. The ROI masks of the thalami and brainstem are illustrated in Fig. 5.3.

5.3.2 Purely intensity-based GMM without PV modelling

First we segment MLS using a standard GMM without PV modelling. We define two classes in the ROI: MLS and BKG. As explained in Chapter 2, MR images can be represented using a parameterised model where the observed intensities y are generated by the hidden labels z with Gaussian probability density functions (PDFs). The Gaussian parameters Φ_y consist of mean μ_k and standard deviation (SD) σ_k of class k . The prior probability that voxel i belongs to class k is modelled by the spatially constant mixing proportion c_k .

We apply the EM algorithm (Dempster et al., 1977) as follows:

- E-step: Calculate the probability p_{ik} that voxel i belongs to class k :

$$p_{ik}^{(m+1)} = \frac{G(y_i, \mu_k^{(m)}, \sigma_k^{(m)}) c_k^{(m)}}{\sum_{k'=1}^K G(y_i, \mu_{k'}^{(m)}, \sigma_{k'}^{(m)}) c_{k'}^{(m)}} \quad (5.1)$$

where y_i is the observed intensity, K the number of classes, and m the current EM iteration number.

- M-step: We assume that the SDs of the MLS and BKG classes are equal:

$$\sigma_{\text{MLS}} = \sigma_{\text{BKG}} \quad (5.2)$$

Ideally the MLS and BKG tissues in the ROI have uniform intensities. Variations arise mainly due to noise and can be considered constant across the MR image. Therefore, the two classes may be regarded as Gaussians with equal SDs provided that their mean intensities are distant from zero.

We showed in Chapter 2 that the conditional expectation of log likelihood for the complete data $\{\mathbf{y}, \mathbf{z}\}$ can be written as:

$$\begin{aligned} Q(\Phi_y | \Phi_y^{(m)}) &= E \left[\ln p(\mathbf{y}, \mathbf{z} | \Phi_y) | \mathbf{y}, \Phi_y^{(m)} \right] \\ &= \sum_{l=1}^K \sum_{i=1}^N p_{il}^{(m+1)} \ln G(y_i, \mu_l, \sigma_l) + \sum_{l=1}^K \sum_{i=1}^N p_{il}^{(m+1)} \ln c_l \end{aligned} \quad (5.3)$$

and

$$\ln G(y_i, \mu_l, \sigma_l) = -\frac{1}{2} \ln(2\pi\sigma_l^2) - \frac{(y_i - \mu_l)^2}{2\sigma_l^2} \quad (5.4)$$

where N is the number of voxels in the ROI. By substituting Eq. 5.2 into Eq. 5.4, maximisation of $Q(\Phi_y | \Phi_y^{(m)})$ with respect to mean μ_k , SD σ_k^2 and mixing proportion c_k of class k yields the following equations:

$$\mu_k^{(m+1)} = \frac{\sum_{i=1}^N p_{ik}^{(m+1)} y_i}{\sum_{i=1}^N p_{ik}^{(m+1)}} \quad (5.5)$$

$$\left(\sigma_k^{(m+1)}\right)^2 = \left(\sigma^{(m+1)}\right)^2 = \frac{1}{N} \sum_{l=1}^K \sum_{i=1}^N p_{il}^{(m+1)} \left(y_i - \mu_l^{(m+1)}\right)^2 \quad (5.6)$$

$$c_k^{(m+1)} = \frac{1}{N} \sum_{i=1}^N p_{ik}^{(m+1)} \quad (5.7)$$

5.3.3 Explicit PV modelling

As demonstrated in Chapter 1, the intermediate PV intensities cause a heavy tail between the intensity distributions of the composing pure tissues in the image histogram. As a result, when fitting the simple two-class GMM described in Section 5.3.2, the mean of the smaller MLS class is shifted towards the larger BKG class, which leads to over-estimation of MLS. Therefore, we introduce an explicit PV class in addition to the previous MLS and BKG classes to model the intensities of mixed tissues.

The E-step (Eq. 5.1) remains unchanged at this stage. In the M-step, we approximate the PV class mean as the arithmetic mean of the composing pure tissue means, and assume that the SDs of the MLS, PV and BKG classes are identical for the reason explained in Section 5.3.2:

$$\mu_{PV} = \frac{1}{2}(\mu_{MLS} + \mu_{BKG}) \quad (5.8)$$

$$\sigma_{MLS} = \sigma_{PV} = \sigma_{BKG} \quad (5.9)$$

By substituting Eq. 5.9 into Eq. 5.4, maximisation of $Q(\Phi_y | \Phi_y^{(m)})$ in Eq. 5.3 with respect to σ_k^2 yields Eq. 5.6 for updating the identical SD. By substituting both Eqs. 5.8 and 5.9 into Eq. 5.4, we then differentiate $Q(\Phi_y | \Phi_y^{(m)})$ with respect to μ_{MLS} and μ_{BKG} , respectively, and obtain the following system of linear equations:

$$\begin{aligned} \frac{1}{4}\mu_{\text{MLS}} \sum_{i=1}^N p_{i\text{PV}} + \mu_{\text{BKG}} \sum_{i=1}^N \left(p_{i\text{BKG}} + \frac{1}{4}p_{i\text{PV}} \right) &= \sum_{i=1}^N y_i \left(p_{i\text{BKG}} + \frac{1}{2}p_{i\text{PV}} \right) \\ \mu_{\text{MLS}} \sum_{i=1}^N \left(p_{i\text{MLS}} + \frac{1}{4}p_{i\text{PV}} \right) + \frac{1}{4}\mu_{\text{BKG}} \sum_{i=1}^N p_{i\text{PV}} &= \sum_{i=1}^N y_i \left(p_{i\text{MLS}} + \frac{1}{2}p_{i\text{PV}} \right) \end{aligned}$$

We omit the iteration number in the posterior probabilities for clarity. Lastly, we solve the linear equations via matrix inversion:

$$\begin{bmatrix} \mu_{\text{MLS}}^{(m+1)} \\ \mu_{\text{BKG}}^{(m+1)} \end{bmatrix} = \begin{bmatrix} \frac{1}{4}B_{\text{PV}} & B_{\text{BKG}} + \frac{1}{4}B_{\text{PV}} \\ B_{\text{MLS}} + \frac{1}{4}B_{\text{PV}} & \frac{1}{4}B_{\text{PV}} \end{bmatrix}^{-1} \begin{bmatrix} A_{\text{BKG}} + \frac{1}{2}A_{\text{PV}} \\ A_{\text{MLS}} + \frac{1}{2}A_{\text{PV}} \end{bmatrix} \quad (5.10)$$

where

$$A_k = \sum_{i=1}^N y_i p_{ik}^{(m+1)} \quad \text{and} \quad B_k = \sum_{i=1}^N p_{ik}^{(m+1)}$$

5.3.4 Spatial regularisation via second-order MRFs

Van Leemput et al. (2003) pointed out that introducing an additional PV class without any spatial prior can disrupt the process of EM convergence. Therefore, robust PV estimation methods all require prior knowledge of the spatial locations of the composing pure tissues to guide the search for PV voxels (Shattuck et al., 2001; Van Leemput et al., 2003; Tohka et al., 2004; Manjon et al., 2010; Cardoso et al., 2011; Ledig et al., 2012). However, MLS is neither included in any of the existing probabilistic neonatal brain atlases (Kuklisova-Murgasova et al., 2011; Serag et al., 2012), nor any manual annotation database. To overcome this difficulty, we incorporate spatial regularisation in the EM framework via second-order MRFs (Ledig et al., 2012) where we configure the dependencies among a triplet of classes in the neighbourhood using a 3D connectivity tensor (Fig. 5.4).

	BKG	PV	MLS		BKG	PV	MLS		BKG	PV	MLS
BKG	t_1	0	t_2	BKG	t_3	0	0	BKG	0	0	t_2
PV	0	0	0	PV	0	0	0	PV	0	0	0
MLS	t_2	0	0	MLS	0	0	t_3	MLS	t_2	0	t_1
	$k = \text{MLS}$				$k = \text{PV}$				$k = \text{BKG}$		

Figure 5.4: Matrix $\mathbf{T}_k$ in the 3D connectivity tensor for $k = \text{MLS}$, PV and BKG . Each element $\mathbf{T}_k(k_1, k_2)$ indicates the penalty when classes k_1 and k_2 are both present in the neighbourhood of class k . We will determine the values of penalties t_1 , t_2 and t_3 empirically in Section 5.4.1. Abbreviations: MLS–myelin-like signals, PV–partial volume, BKG–background.

First we modify the E-step (Eq. 5.1) by replacing the spatially constant mixing proportion c_k with the spatial prior $p(\mathbf{z}_i = \mathbf{e}_k | \mathbf{p}_{\mathcal{N}_i}^{(m)}, \Phi_z)$ acquired from the neighbourhood as described in Van Leemput et al. (1999b):

$$p_{ik}^{(m+1)} = \frac{G(y_i, \mu_k^{(m)}, \sigma_k^{(m)}) p(\mathbf{z}_i = \mathbf{e}_k | \mathbf{p}_{\mathcal{N}_i}^{(m)}, \Phi_z)}{\sum_{k'=1}^K G(y_i, \mu_{k'}^{(m)}, \sigma_{k'}^{(m)}) p(\mathbf{z}_i = \mathbf{e}_{k'} | \mathbf{p}_{\mathcal{N}_i}^{(m)}, \Phi_z)} \quad (5.11)$$

where

$$p(\mathbf{z}_i = \mathbf{e}_k | \mathbf{p}_{\mathcal{N}_i}^{(m)}, \Phi_z) = \frac{e^{-U_{\text{MRF}}(\mathbf{e}_k | \mathbf{p}_{\mathcal{N}_i}^{(m)}, \Phi_z)}}{\sum_{k''=1}^K e^{-U_{\text{MRF}}(\mathbf{e}_{k''} | \mathbf{p}_{\mathcal{N}_i}^{(m)}, \Phi_z)}} \quad (5.12)$$

The symbol $\mathbf{z}_i$ indicates to which class voxel i belongs, and $\mathbf{e}_k$ is a unit vector of length K whose components are equal to zero except the k th component. The probabilities in the 26-neighbourhood $\mathcal{N}_i$ of voxel i are denoted as $\mathbf{p}_{\mathcal{N}_i}$, and the MRF parameters as Φ_z . We compute the energy function $U_{\text{MRF}}(\mathbf{e}_k | \mathbf{p}_{\mathcal{N}_i}, \Phi_z)$ for second-order MRFs as follows:

$$U_{\text{MRF}}(\mathbf{e}_k | \mathbf{p}_{\mathcal{N}_i}^{(m)}, \Phi_z) = \sum_{k_1=1}^K v_{k_1}^{(m)} \sum_{k_2=1}^K \mathbf{T}_k(k_1, k_2) v_{k_2}^{(m)} \quad (5.13)$$

where

$$v_k^{(m)} = \sum_{j \in \mathcal{N}_i} \frac{p_{jk}^{(m)}}{d_{ij}} \quad (5.14)$$

Here p_{jk} is the probability that neighbour j belongs to class k , and d_{ij} the Euclidean distance between voxel i and neighbour j . The element of the connectivity tensor, denoted as $\mathbf{T}_k(k_1, k_2)$, indicates the penalty when classes k_1 and k_2 are both present in the neighbourhood of class k .

Eqs. 5.13 and 5.14 can also be written in matrix notation:

$$\begin{aligned}
 U_{\text{MRF}}(\mathbf{e}_k | \mathbf{p}_{\mathcal{N}_i}, \Phi_z) &= \mathbf{v} \mathbf{T}_k \mathbf{v}^T \\
 &= \begin{bmatrix} v_1 & v_2 & \cdots & v_K \end{bmatrix} \begin{bmatrix} T_k(1,1) & T_k(1,2) & \cdots & T_k(1,K) \\ T_k(2,1) & T_k(2,2) & \cdots & T_k(2,K) \\ \vdots & \vdots & \ddots & \vdots \\ T_k(K,1) & T_k(K,2) & \cdots & T_k(K,K) \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_K \end{bmatrix} \quad (5.15)
 \end{aligned}$$

where

$$\mathbf{v} = \begin{bmatrix} v_1 & v_2 & \cdots & v_K \end{bmatrix} = \begin{bmatrix} \frac{1}{d_{i,1}} & \frac{1}{d_{i,2}} & \cdots & \frac{1}{d_{i,26}} \end{bmatrix} \begin{bmatrix} p_{1,1} & p_{1,2} & \cdots & p_{1,K} \\ p_{2,1} & p_{2,2} & \cdots & p_{2,K} \\ \vdots & \vdots & \ddots & \vdots \\ p_{26,1} & p_{26,2} & \cdots & p_{26,K} \end{bmatrix} \quad (5.16)$$

The iteration number m has been omitted for clarity.

We can see from Eq. 5.15 that by assigning a large value to $\mathbf{T}_k(k_1, k_2)$ in the connectivity tensor, second-order MRFs enable us to penalise a particular configuration when the probabilities of classes k_1 and k_2 are both high when summed over all the neighbours. This would result in a large value of the energy function $U_{\text{MRF}}(\mathbf{e}_k | \mathbf{p}_{\mathcal{N}_i}, \Phi_z)$, making voxel i less probable to belong to class k .

Let us first consider the construction of matrices $\mathbf{T}_{\text{MLS}}$ and $\mathbf{T}_{\text{BKG}}$. In order to penalise direct adjacency between MLS and BKG and to insert a thin layer of PV voxels at the tissue boundary, we assign penalty t_1 to $\mathbf{T}_{\text{MLS}}(\text{BKG}, \text{BKG})$ and $\mathbf{T}_{\text{BKG}}(\text{MLS}, \text{MLS})$, and penalty t_2 to $\mathbf{T}_{\text{MLS}}(\text{MLS}, \text{BKG})$ and $\mathbf{T}_{\text{BKG}}(\text{MLS}, \text{BKG})$. Note that every matrix $\mathbf{T}_k$ is symmetric. There is no penalty for homogeneous regions containing a single pure tissue class. We thus set $\mathbf{T}_{\text{MLS}}(\text{MLS}, \text{MLS})$ and $\mathbf{T}_{\text{BKG}}(\text{BKG}, \text{BKG})$ to zero. To encourage the desirable configuration

when MLS, PV and BKG are all included in the triplet of classes k , k_1 and k_2 irrespective of the order of appearance, we set $\mathbf{T}_{\text{MLS}}(\text{BKG}, \text{PV})$ and $\mathbf{T}_{\text{BKG}}(\text{MLS}, \text{PV})$ to zero. Lastly, we set $\mathbf{T}_{\text{MLS}}(\text{PV}, \text{PV})$, $\mathbf{T}_{\text{MLS}}(\text{MLS}, \text{PV})$, $\mathbf{T}_{\text{BKG}}(\text{PV}, \text{PV})$ and $\mathbf{T}_{\text{BKG}}(\text{BKG}, \text{PV})$ to zero as the PV class is likely to adjoin any of the pure tissue classes.

We then consider the construction of matrix $\mathbf{T}_{\text{PV}}$. If t_1 and t_2 remain the only non-zero penalties in the connectivity tensor with all the elements of $\mathbf{T}_{\text{PV}}$ equal to zero, the PV posterior probability map (PPM) would be significantly over-estimated, as illustrated in Fig. 5.5(b). Based on this empirical observation, we assign penalty t_3 to $\mathbf{T}_{\text{PV}}(\text{MLS}, \text{MLS})$ and $\mathbf{T}_{\text{PV}}(\text{BKG}, \text{BKG})$ to prevent the PV class from “invading” the MLS and BKG regions where the probabilities of the MLS and BKG classes respectively are high in the neighbourhood. Later experiments in Section 5.4.2 will demonstrate that penalty t_3 provides sufficient spatial regularisation of the PV class, and there is no need to introduce another parameter for $\mathbf{T}_{\text{PV}}(\text{MLS}, \text{PV})$ or $\mathbf{T}_{\text{PV}}(\text{BKG}, \text{PV})$. Again to encourage the configuration when MLS and BKG are both present in the neighbourhood of PV, we set $\mathbf{T}_{\text{PV}}(\text{MLS}, \text{BKG})$ to zero. Lastly, we set $\mathbf{T}_{\text{PV}}(\text{PV}, \text{PV})$ to zero, otherwise the PPM of the PV class would diminish, as shown in Fig. 5.5(e). The values of penalties t_1 , t_2 and t_3 will be determined empirically in Section 5.4.1.

Denoting $\Phi = \{\Phi_y, \Phi_z\}$, Van Leemput et al. (1999b) showed that the conditional expectation of the complete log likelihood incorporating MRFs becomes:

$$\begin{aligned} Q(\Phi | \Phi^{(m)}) &= E \left[\ln p(\mathbf{y}, \mathbf{z} | \Phi) | \mathbf{y}, \Phi^{(m)} \right] \\ &= \sum_{l=1}^K \sum_{i=1}^N p_{il}^{(m+1)} \ln G(y_i, \mu_l, \sigma_l) \\ &\quad - E \left[U_{\text{MRF}}(\mathbf{z} | \Phi_z) + \ln f(\Phi_z) | \mathbf{y}, \Phi^{(m)} \right] \end{aligned} \quad (5.17)$$

where $U_{\text{MRF}}(\mathbf{z} | \Phi_z)$ is the energy function, and $f(\Phi_z)$ denotes a normalisation constant summed over all possible configurations of the hidden labels. We again assume that the PV class mean equals the arithmetic mean of the composing pure tissue means and that the SDs of the MLS, PV and BKG classes are identical. Provided that the penalties remain constant, the SD σ_k and mean μ_k of class k in the EM framework with MRFs

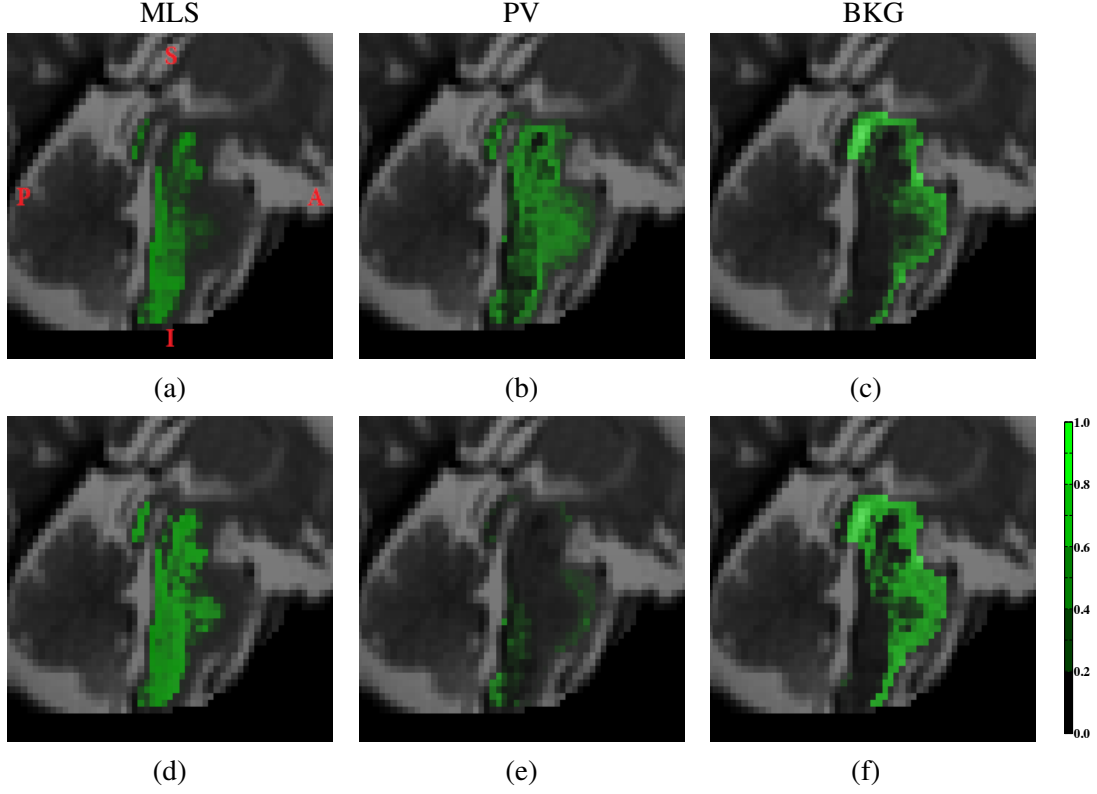


Figure 5.5: Empirical effects on the posterior probability maps (PPMs) (a-c) when all the elements of matrix $\mathbf{T}_{PV}$ are set to zero, and (d-f) when all the elements of $\mathbf{T}_{PV}$ are set to zero except $\mathbf{T}_{PV}(PV, PV)$. Matrices $\mathbf{T}_{MLS}$ and $\mathbf{T}_{BKG}$ are constructed as shown in Fig. 5.4. The columns from left to right show the PPMs of myelin-like signals (MLS), partial volume voxels (PV) and background (BKG) in the sagittal brainstem for a subject at 37 gestational weeks. The PV class is either (b) over-estimated or (e) diminished. Neither of the settings is able to constrain the PV class as a thin layer of voxels at the boundary between MLS and BKG.

can be estimated using Eqs. 5.6 and 5.10, respectively, in the same way as the purely intensity-based GMM with an explicit PV class.

5.3.5 Implementation

The EM algorithm is initialised as follows. We first sort the T_2w intensities in the thalami of individual subjects in ascending order and choose the 6th percentile as the threshold. This is because the average volume fraction of manually annotated MLS in the thalami is 6% over the 16 test subjects at one-week intervals between 29 and 44 weeks GA. Voxels with intensities below the threshold constitute the initial MLS segmentation, and the remaining voxels in the ROI form the initial BKG segmentation. We will verify in

Section 5.4.3 that the percentile estimated this way optimises the MLS segmentations that could possibly be achieved for all the 16 subjects using a single threshold. We assign the initial MLS and BKG segmentations to the corresponding PPMs, and set the PPM of the PV class equal to zero.

Similarly, we threshold the T_2w intensities in the brainstem of individual subjects at the 25th percentile to obtain the initial segmentation estimates in this region.

Tissue classification is achieved by interleaving the E-step and M-step as described in Section 5.3.4 until the relative change in the objective function of log likelihood (Van Leemput et al., 1999b) is less than 0.01%. The resultant PPMs are converted into hard segmentations using the maximum-vote rule. In addition, we reclassify the PV voxels as one of the composing pure tissues by calculating the fraction of the MLS class, denoted as f_i , at PV voxel i :

$$f_i = \frac{\mu_{\text{BKG}} - y_i}{\mu_{\text{BKG}} - \mu_{\text{MLS}}} \quad (5.18)$$

where μ_{MLS} and μ_{BKG} are the means of the MLS and BKG classes, and y_i the observed intensity. PV voxels with fractions above 0.5 are reclassified as MLS and combined with the voted hard segmentation of the MLS class to form the final segmentation.

5.4 Results and discussion

In this section, we first determined the optimal values of the penalties in the connectivity tensor for the proposed method with explicit PV modelling via second-order MRFs (GMM-PV-MRF). We then evaluated GMM-PV-MRF in terms of Dice overlap with manually annotated MLS, as compared to two other segmentation models in Table 5.1. The initial MLS segmentations generated by optimal thresholding exploited manual annotations of MLS that are usually unavailable in practice. We thus further investigated the dependence of each method on the initial threshold. Lastly, we assessed the intra-rater reliability compared to the accuracies of the automated model-based methods.

Segmentation model	Description	Partial volume (PV) modelling	Spatial regularisation
GMM	Purely intensity-based Gaussian mixture model without PV modelling	×	×
GMM-PV	GMM incorporating a spatially unconstrained PV class	✓	×
GMM-PV-MRF	GMM with explicit PV modelling via second-order MRFs	✓	✓

Table 5.1: Segmentation models for myelin-like signals.

5.4.1 Optimisation of the penalties in the connectivity tensor

As explained in Section 5.3.4, we acquired the MLS, PV and BKG priors by configuring their spatial dependencies in the connectivity tensor (Fig. 5.4) via second-order MRFs. For each individual ROI, we optimised the penalties t_1 , t_2 and t_3 using a coarse-to-fine grid search method. We first determined the coarse value of each penalty from the set:

$$\{0.003, 0.007, 0.03, 0.07, 0.3, 0.7\}$$

Two values were included for the lower and upper halves of each tenfold scale. We computed MLS segmentations using all 216 combinations of the coarse t_1 , t_2 and t_3 values. We chose the combination that yielded the highest average DC over the 16 test subjects with respect to the manual annotations.

We then performed a fine grid search through ten different values for each penalty with the selected coarse value at the center. For instance, we examined the set:

$$\{0.007, 0.008, 0.009, 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07\}$$

for t_1 in the thalami, whose best value at the coarse level was found to be 0.03. The test sets and optimal values of all three penalties at the coarse and fine levels are summarised in Table 5.2 for both ROIs. The final (t_1, t_2, t_3) combination that yielded the highest average DC was $(0.05, 0.03, 0.01)$ in the thalami, and $(0.05, 0.03, 0.009)$ in the brainstem, which were highly consistent even though optimised separately.

In Figs. 5.6 and 5.7, we plotted the average DCs against the test values of each penalty at the coarse and fine levels in the thalami and brainstem, respectively, while holding the other two penalties at their optimal values. The resultant DCs were particularly sensitive to the values of t_3 in both ROIs. As explained in Fig. 5.4, t_3 penalised the PV class if only one of the composing pure tissues was found in the neighbourhood, and thus played a key role in preventing over-estimation of the PV class.

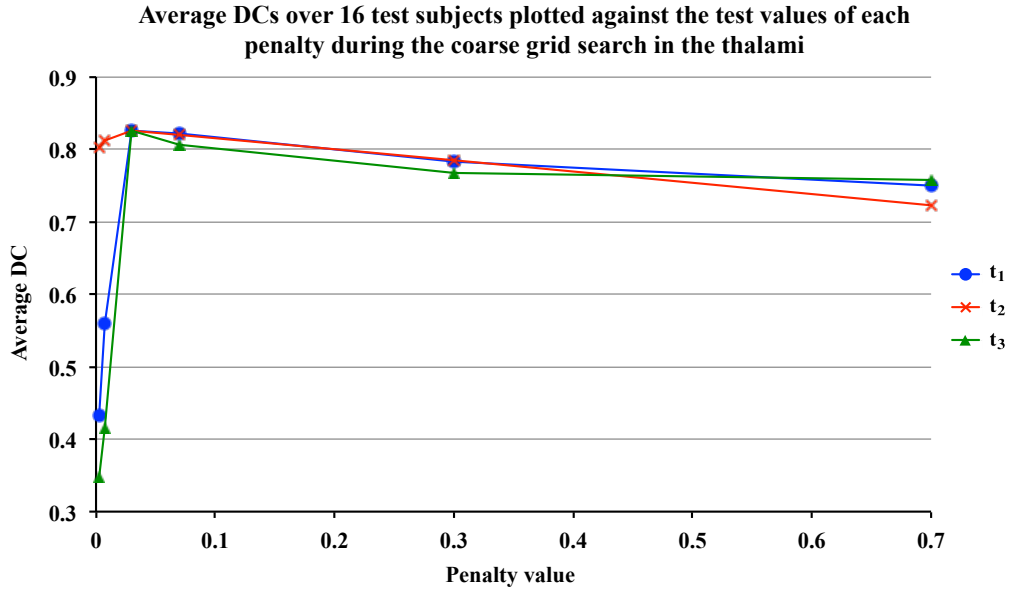
Search level	Penalty	Set of test values	Optimal value
Coarse	t_1	$\{0.003, 0.007, 0.03, 0.07, 0.3, 0.7\}$	0.03
	t_2	Same as t_1 at the coarse level	0.03
	t_3	Same as t_1 at the coarse level	0.03
Fine	t_1	$\{0.007, 0.008, 0.009, 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07\}$	0.05
	t_2	Same as t_1 at the fine level	0.03
	t_3	Same as t_1 at the fine level	0.01

(a) Grid search results in the thalami

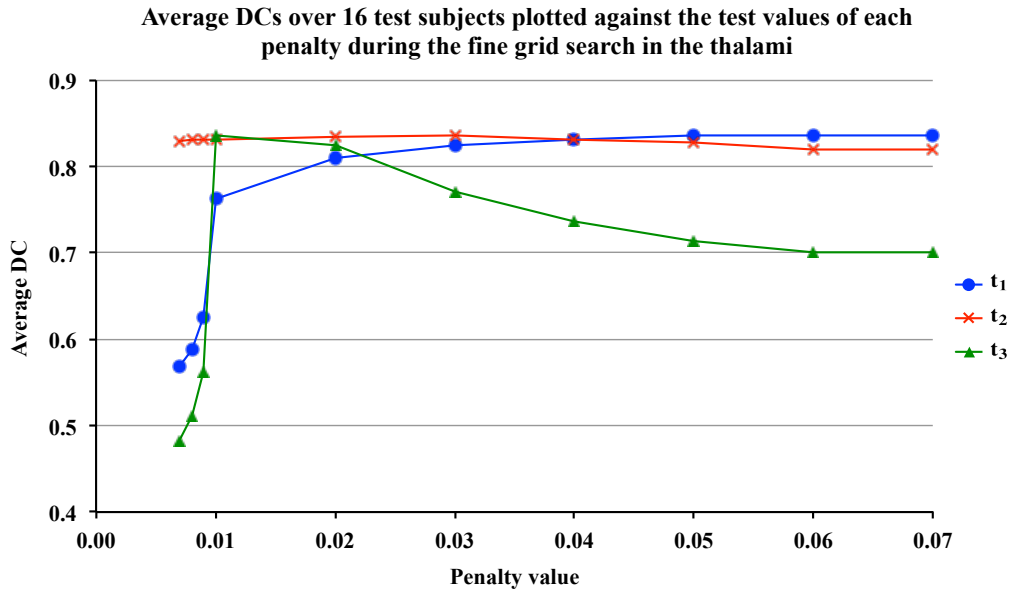
Search level	Penalty	Set of test values	Optimal value
Coarse	t_1	$\{0.003, 0.007, 0.03, 0.07, 0.3, 0.7\}$	0.03
	t_2	Same as t_1 at the coarse level	0.03
	t_3	Same as t_1 at the coarse level	0.007
Fine	t_1	$\{0.007, 0.008, 0.009, 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07\}$	0.05
	t_2	Same as t_1 at the fine level	0.03
	t_3	$\{0.003, 0.004, 0.005, 0.006, 0.007, 0.008, 0.009, 0.01, 0.02, 0.03\}$	0.009

(b) Grid search results in the brainstem

Table 5.2: Optimal values of the penalties t_1 , t_2 and t_3 in different regions of interest (ROIs) determined using a coarse-to-fine grid search method. These penalties are assigned to the connectivity tensor (Fig. 5.4) for configuring spatial dependencies of myelin-like signals, partial volume voxels and the background in each ROI.

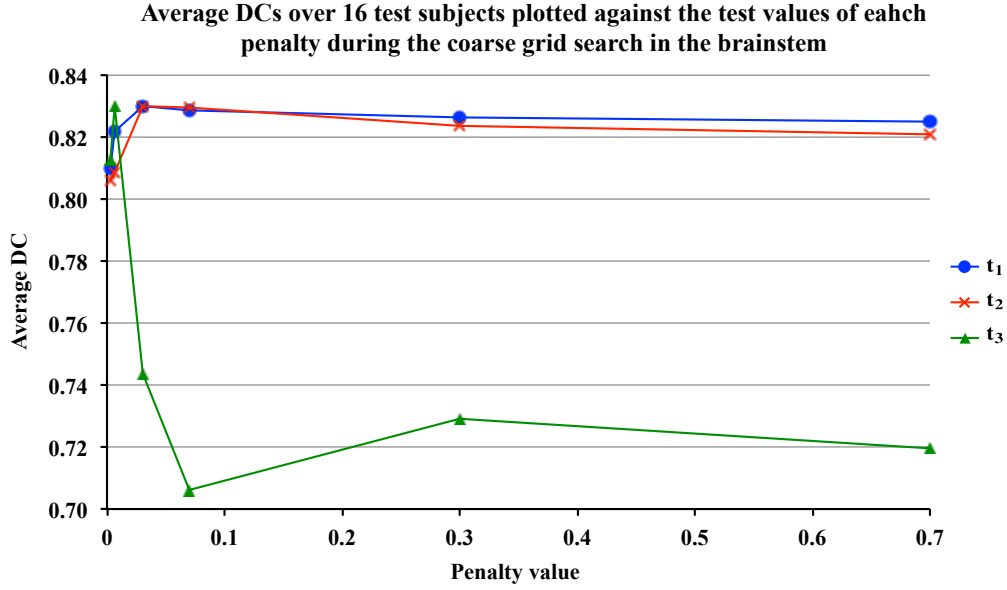


(a) Coarse search in the thalami

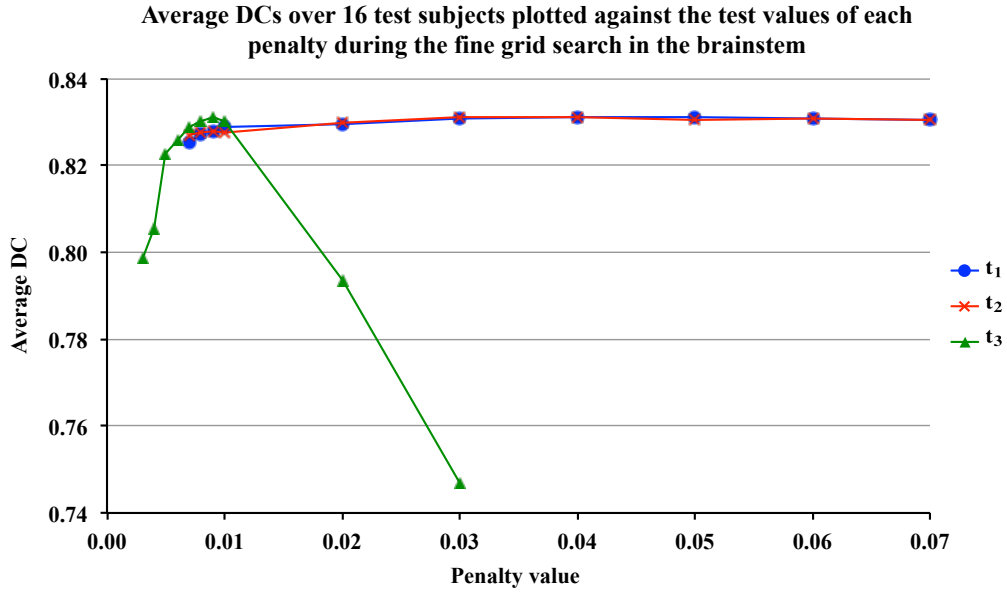


(b) Fine search in the thalami

Figure 5.6: Average Dice coefficients (DCs) over 16 test subjects plotted against the test values of each penalty at the coarse and fine levels of grid search in the thalami. The plots are generated by varying one of the penalties through the values listed in Table 5.2(a) while the other two remain at their optimal values. The (t_1, t_2, t_3) combination that yields the highest average DC is $(0.03, 0.03, 0.03)$ at the coarse level, and $(0.05, 0.03, 0.01)$ at the fine level.



(a) Coarse search in the brainstem



(b) Fine search in the brainstem

Figure 5.7: Average Dice coefficients (DCs) over 16 test subjects plotted against the test values of each penalty at the coarse and fine levels of grid search in the brainstem. The plots are generated by varying one of the penalties through the values listed in Table 5.2(b) while the other two remain at their optimal values. The (t_1, t_2, t_3) combination that yields the highest average DC is $(0.03, 0.03, 0.007)$ at the coarse level, and $(0.05, 0.03, 0.009)$ at the fine level.

5.4.2 Evaluation of the MLS segmentation methods

We segmented MLS using the model-based methods GMM, GMM-PV and GMM-PV-MRF (see Table 5.1) in the thalami and brainstem for the 16 test subjects at one-week intervals between 29 and 44 weeks GA. From here on, we used the optimal values of t_1 , t_2 and t_3 determined in Section 5.4.1 for GMM-PV-MRF. We implemented GMM as described in Sections 5.3.2 and 5.3.5, and GMM-PV as in Sections 5.3.3 and 5.3.5.

The DCs of the resultant segmentations with respect to the manual annotations are summarised for both ROIs in Figs. 5.8 and 5.9. We also evaluated the initial MLS segmentations obtained by optimal thresholding as a baseline method. We compared GMM-PV-MRF with all the other methods using two-tailed Student's t -tests at the 5% significance level, and the p -values are shown in Table 5.3.

Method	Average DC	p -value
Thresholding at the 6th percentile	0.789 ± 0.091	0.043
GMM	0.762 ± 0.092	0.001
GMM-PV	0.662 ± 0.192	0.002
GMM-PV-MRF	0.837 ± 0.057	—

(a) Evaluation results in the thalami

Method	Average DC	p -value
Thresholding at the 25th percentile	0.827 ± 0.035	0.020
GMM	0.675 ± 0.206	0.010
GMM-PV	0.762 ± 0.057	5.449×10^{-4}
GMM-PV-MRF	0.831 ± 0.038	—

(b) Evaluation results in the brainstem

Table 5.3: Average Dice coefficients (DCs) ($\pm$ standard deviations) of myelin-like signals segmented by optimal thresholding and the model-based methods (see Table 5.1 for method descriptions) for 16 test subjects aged between 29 and 44 gestational weeks. GMM-PV-MRF is compared with all the other methods using two-tailed Student's t -tests at the 5% significance level. It can be seen that GMM-PV-MRF achieves the highest average DC and significantly outperforms all the other methods in both regions of interest.

Thresholding at the optimal level resulted in a lower average DC and a larger SD in the thalami than in the brainstem. The 6th percentile threshold was derived from the average

volume fraction of manually annotated MLS in the thalami over the 16 subjects, and hence decreased the DCs at the early and late GAs due to the appearance of new myelination in this region (Counsell et al., 2002). In contrast, the myelinated brainstem grew with the overall increase in the brain size with relatively little new myelination appearing between 29 and 44 weeks GA (Chi et al., 1977; Rutherford, 2002), which implied an approximately constant MLS volume fraction.

The purely intensity-based GMM without PV modelling displayed large segmentation variations in the brainstem compared to the thalami. In Figs. 5.10 and 5.11, we studied the Gaussian distributions of the PPMs estimated using different segmentation models for a subject at 37 weeks GA. With respect to the manual annotation histograms, GMM overestimated the MLS class in the thalami with a shift of the mean towards BKG, and failed to separate the two classes in the brainstem. This is caused by PV voxels represented as the overlapping area between the histograms of manually annotated MLS and BKG in Figs. 5.10(e) and 5.11(e).

The extra Gaussian in GMM-PV caused large segmentation variations in the thalami due to the absence of spatial regularisation. In the brainstem, GMM-PV appeared to reduce the variations previously shown by GMM, but nonetheless yielded significantly lower DCs ($p < 0.05$) than the baseline method of optimal thresholding. Fig. 5.13(c) revealed that, without any spatial constraints, the PPM of the PV class in GMM-PV for the same subject at 37 weeks GA spread across the brainstem with little variation in space.

GMM-PV-MRF achieved high DCs consistently across all the GAs in both ROIs. The extra Gaussian modelled the intermediate PV intensities, and the spatial constraints imposed via second-order MRFs confined the PV locations at the boundary between MLS and BKG. This approach prevented the PV Gaussian distribution from shifting towards BKG and hence the spread of the PV class in space, as demonstrated in Figs. 5.10(c), 5.11(c), 5.12(f) and 5.13(f). The histograms of the binary segmentations obtained using GMM-PV-MRF in Figs. 5.10(d) and 5.11(d) highly resembled those of the manual annotations. Student's t -tests in Table 5.3 confirmed that GMM-PV-MRF outperformed all the other methods in both ROIs.

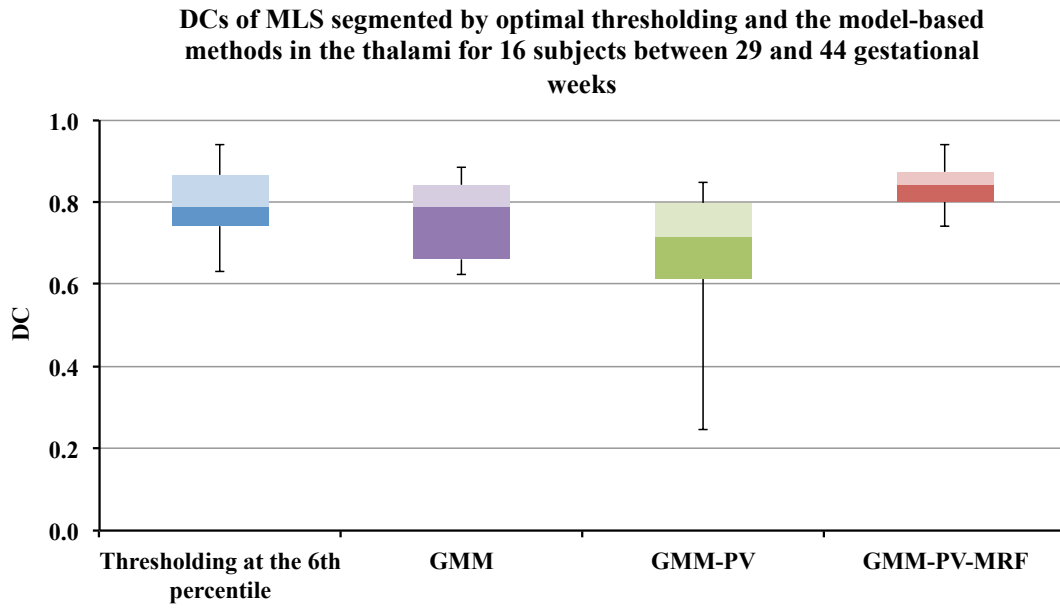


Figure 5.8: Dice coefficients (DCs) of myelin-like signals (MLS) segmented by optimal thresholding and the model-based methods (see Table 5.1 for method descriptions) in the thalami for 16 test subjects aged between 29 and 44 gestational weeks. The dark/light color interface inside each box indicates the median. The bottom and top of the box represent the first and third quartiles, and the ends of the whiskers the maximum and minimum DCs. The DCs obtained from GMM become lower than those of the baseline initial segmentations estimated by optimal thresholding. The extra Gaussian in GMM-PV causes large segmentation variations due to the absence of spatial regularisation. GMM-PV-MRF outperforms all the other methods with the highest average DC and the smallest standard deviation.

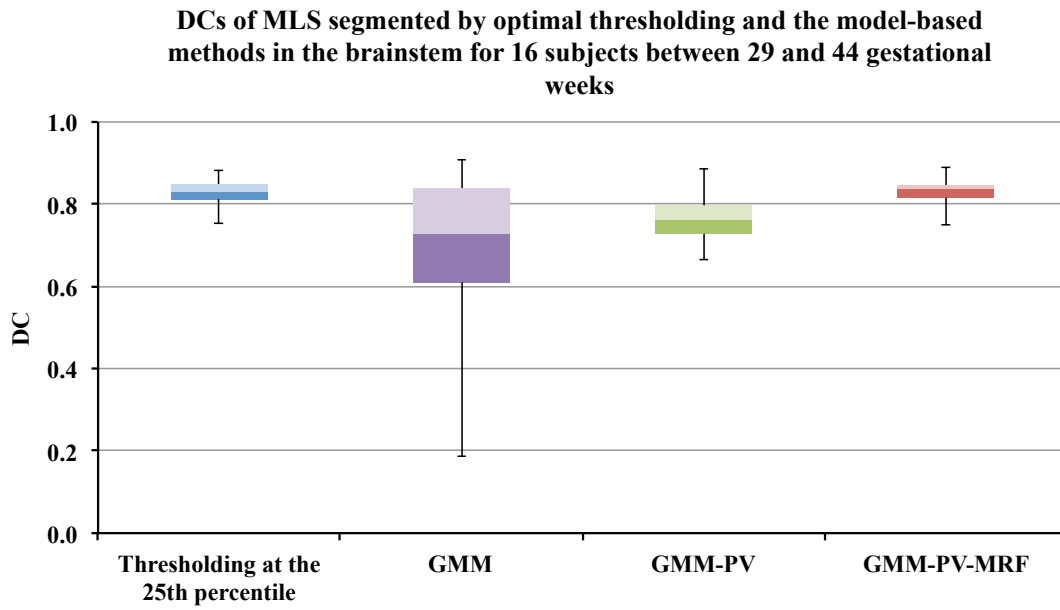


Figure 5.9: Dice coefficients (DCs) of myelin-like signals (MLS) segmented by optimal thresholding and the model-based methods (see Table 5.1 for method descriptions) in the brainstem for 16 test subjects aged between 29 and 44 gestational weeks. The box plot is constructed in the same way as Fig. 5.8. Despite the excellent performance, thresholding at the optimal level is not possible in practice without manual annotations. GMM-PV appears to reduce the segmentation variations shown by GMM, but nonetheless yields lower DCs than the baseline method of optimal thresholding. Only GMM-PV-MRF significantly improves the DCs of the initial MLS estimates.

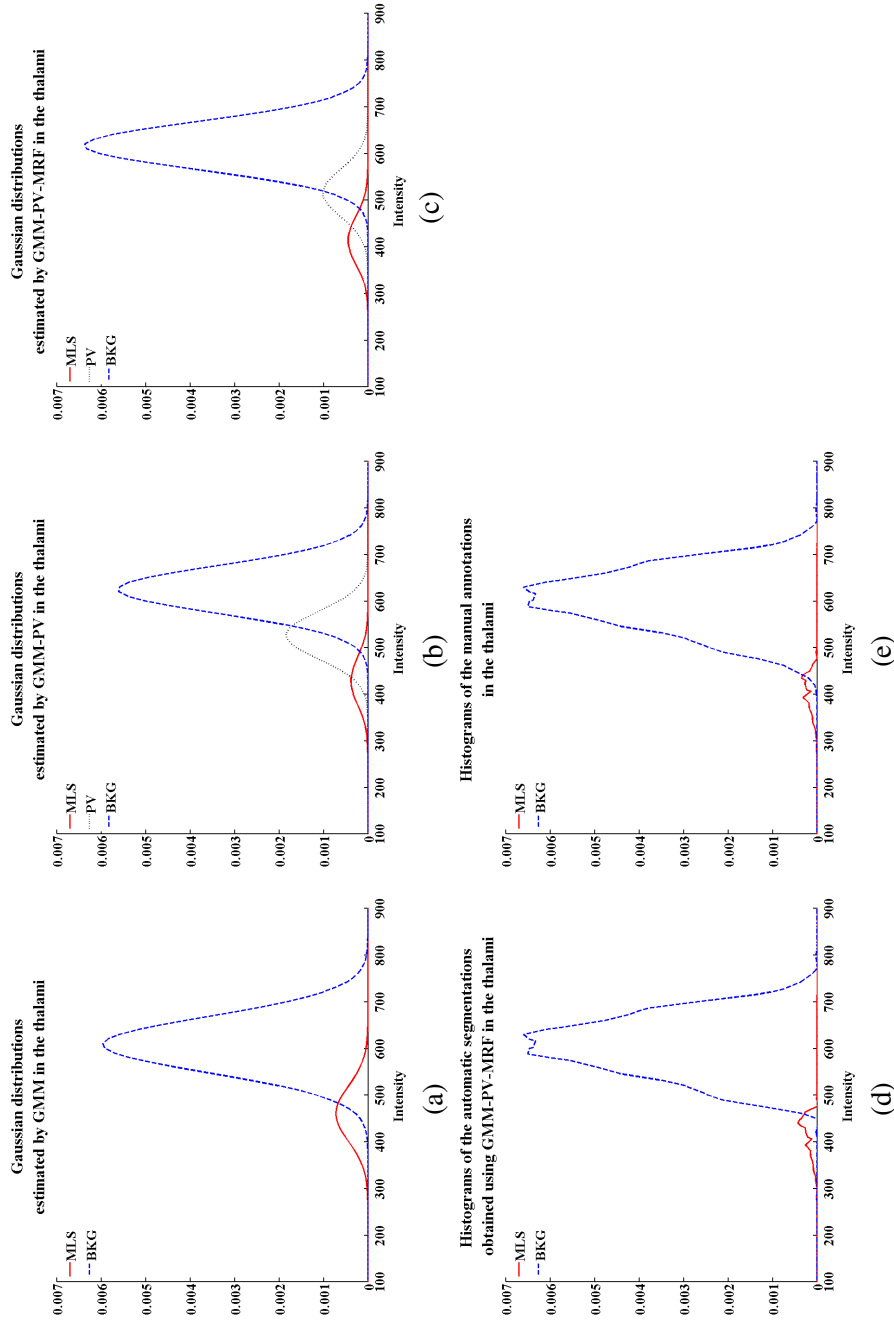


Figure 5.10: Gaussian distributions of the posterior probability maps estimated using the segmentation models in Table 5.1 for myelin-like signals (MLS), background (BKG) and partial volume voxels (PV) in the thalami of a subject at 37 gestational weeks. The sum of the graph areas for all the classes equals one. GMM over-estimates the MLS class with a shift of the mean towards BKG compared to the manual annotation histograms. GMM-PV over-estimates the PV class compared to GMM-PV-MRF due to the absence of spatial regularisation, as confirmed in Fig. 5.12(c). The histograms of the binary segmentations obtained using GMM-PV-MRF are almost the same as those of the manual annotations.

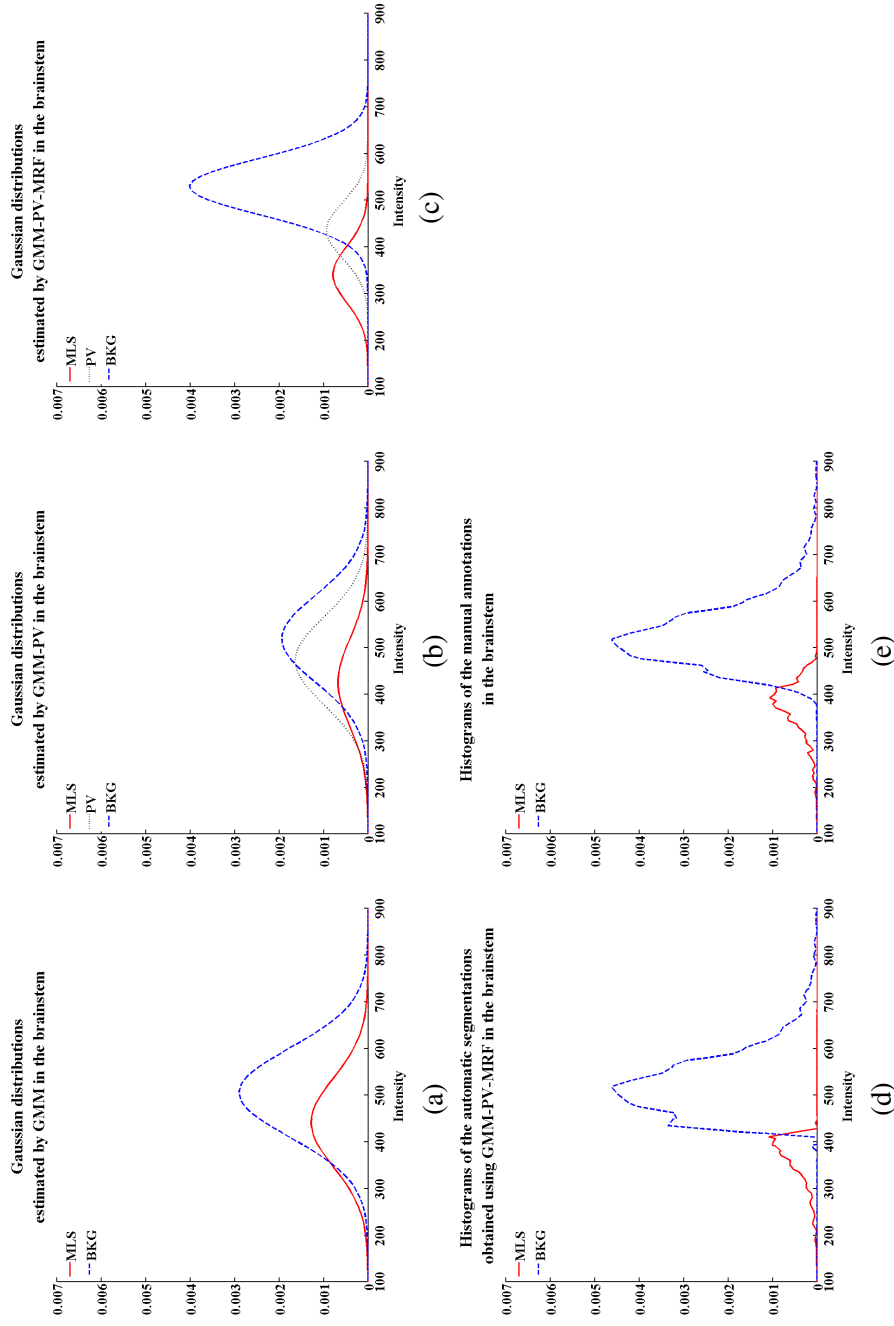


Figure 5.11: Gaussian distributions of the posterior probability maps estimated using the segmentation models in Table 5.1 for myelin-like signals (MLS), background (BKG) and partial volume voxels (PV) in the brainstem of a subject at 37 gestational weeks. The distributions and histograms are shown in the same scale as Fig. 5.10. The brainstem contains more PV voxels than the thalami, which can be seen from the larger overlapping area between the histograms of manually annotated MLS and BKG. EM classification only succeeds in GMM-PV-MRF, and the histograms of the resultant binary segmentations highly resemble those of the manual annotations.

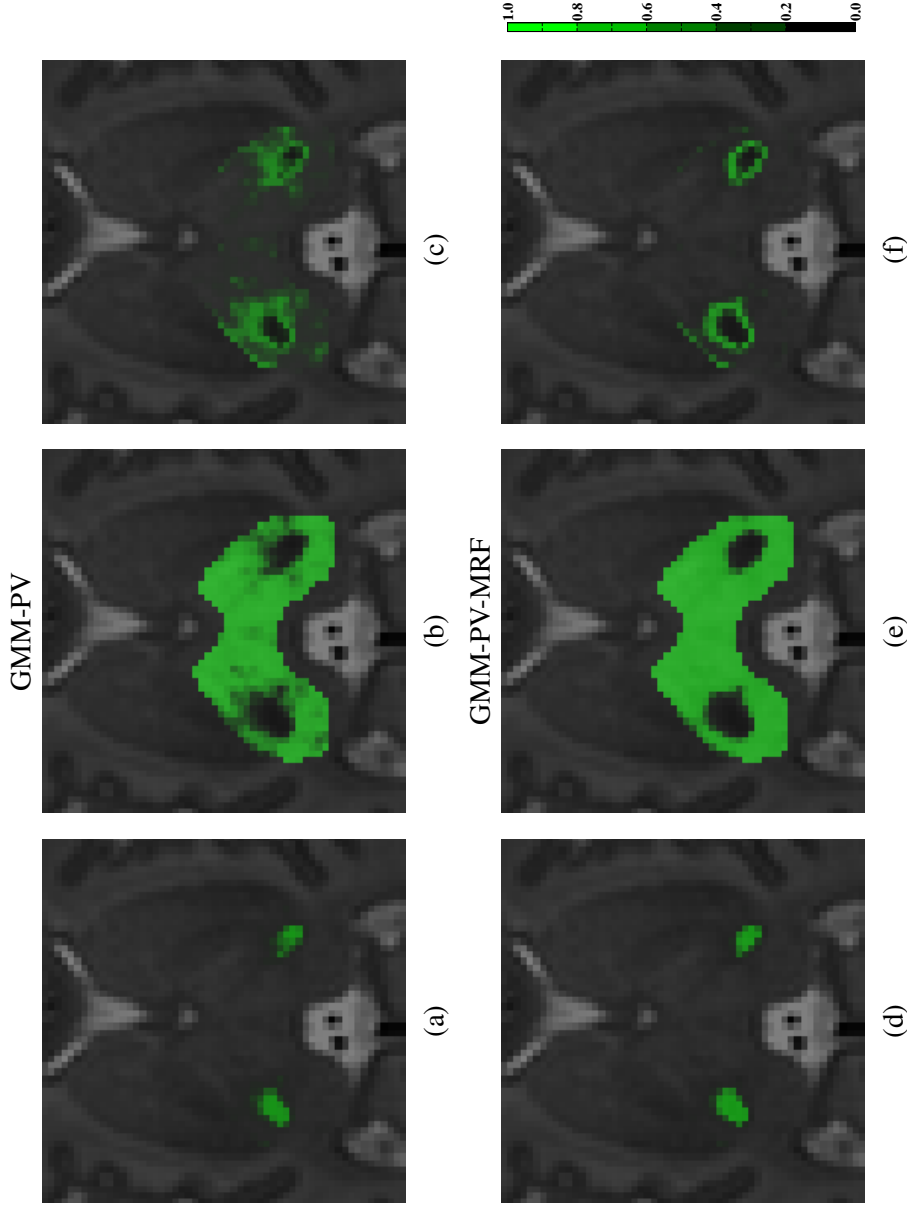


Figure 5.12: Posterior probability maps (PPMs) estimated using GMM-PV and GMM-PV-MRF (see Table 5.1 for method descriptions) in the thalami of a subject at 37 gestational weeks. The columns from left to right show the PPMs in the axial view for myelin-like signals (MLS), background (BKG) and partial volume voxels (PV). The PV class in GMM-PV spreads across the thalami due to the absence of spatial regularisation. In contrast, GMM-PV-MRF configures spatial dependencies of all the classes via second-order MRFs and indicates the PV locations accurately at the boundary between MLS and BKG.

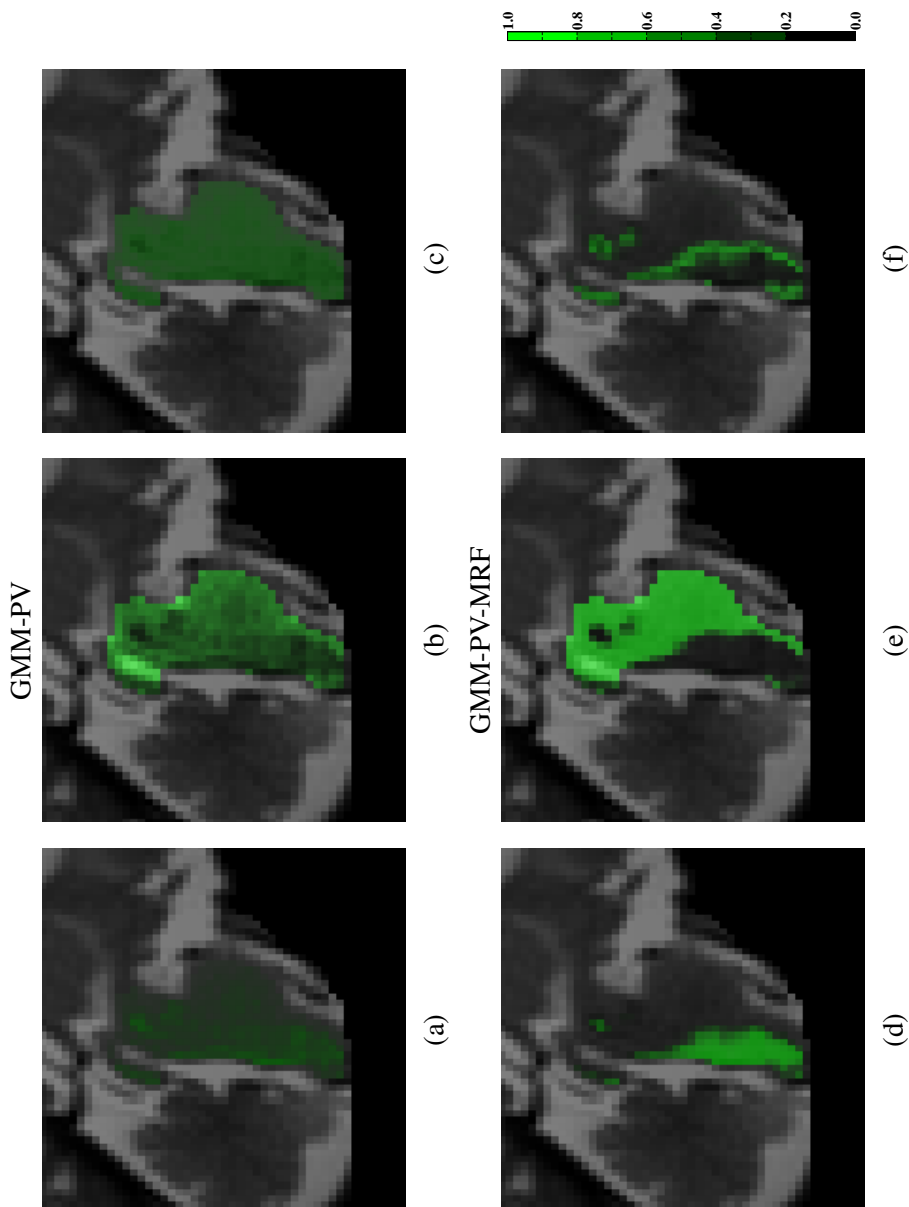


Figure 5.13: Posterior probability maps (PPMs) estimated using GMM-PV and GMM-PV-MRF (see Table 5.1 for method descriptions) in the brainstem of a subject at 37 gestational weeks. The columns from left to right show the PPMs in the sagittal view for myelin-like signals (MLS), background (BKG) and partial volume voxels (PV). GMM-PV-MRF performs accurately and consistently as in the thalami, whereas PV estimation using GMM-PV becomes much worse than in the thalami and shows little variation in space. This is consistent with the nearly overlaid Gaussians of the BKG and PV classes in Fig. 5.11(b).

5.4.3 Dependence on the initial threshold

We initialised MLS segmentations for all the methods in Table 5.1 using an intensity percentile threshold estimated from the volume fraction of manually annotated MLS in the ROI. However, manual reference is usually unavailable in practice, without which it is difficult to estimate the MLS content as it can vary with many factors, such as brain regions and GAs. Therefore, we continued to investigate the dependence of each method on the initial threshold.

We segmented MLS in the thalami and brainstem for the 16 test subjects by thresholding, GMM, GMM-PV and GMM-PV-MRF based on a range of initial thresholds. At each threshold, we calculated the average DC of the resultant segmentations over the subjects for each method, which was then plotted against the threshold in Figs. 5.14 and 5.15. In addition, we displayed the MLS segmentations at selected thresholds for two individual subjects in Figs. 5.16 and 5.17.

GMM-PV-MRF performed most stably and most accurately as the initial threshold varied from the 1st to 20th percentile in the thalami, and from the 21st to 40th percentile in the brainstem. The segmentation robustness towards the threshold is a significant advantage as the choice of the initial threshold is difficult without manual annotations, and moreover, the need for choosing the initial threshold is undesirable for practical applications. Especially when comparing the automatic MLS segmentation of an individual subject against the average growth trajectory for assessing brain maturation, it is of great importance that the volume fraction of the initial MLS estimate does not influence the final segmentation.

Thresholding, GMM and GMM-PV were particularly sensitive to the initial threshold, and performed inconsistently in different ROIs. Compared to the brainstem, the DCs obtained by thresholding underwent dramatic changes in the thalami. The additional Gaussian in GMM-PV lowered the DCs of GMM in the thalami due to the absence of spatial regularisation. In the brainstem with substantial PV voxels, GMM-PV helped to stabilise the variations shown by GMM, but was nonetheless unable to outperform the baseline method of thresholding.

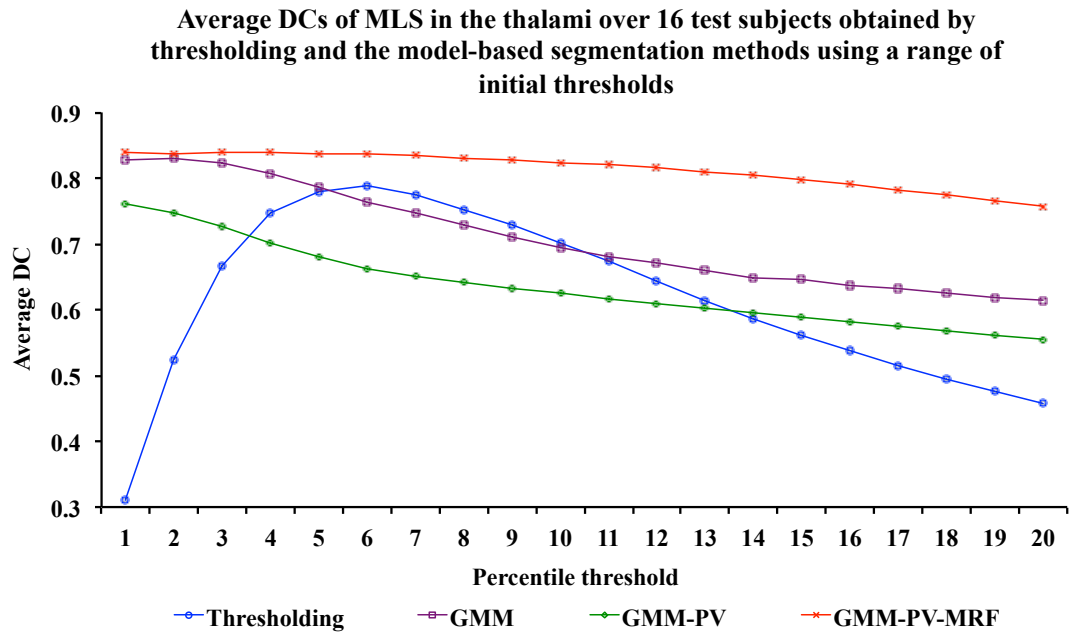


Figure 5.14: Average Dice coefficients (DCs) of myelin-like signals (MLS) in the thalami over 16 test subjects obtained by thresholding and the model-based methods (see Table 5.1 for method descriptions) using a range of initial thresholds. GMM-PV-MRF performs most robustly with the highest DCs as the threshold varies from the 1st to 20th percentile of the T_2w intensities in this region. Thresholding shows dramatic changes of the DC values, and the additional Gaussian in GMM-PV lowers the DCs compared to GMM due to the absence of spatial regularisation.

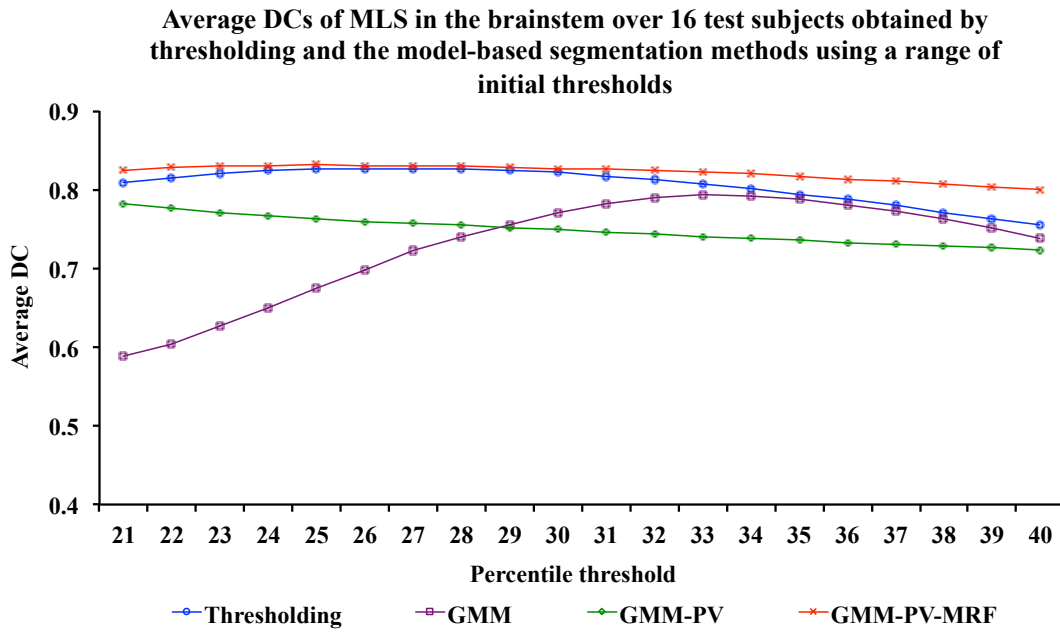


Figure 5.15: Average Dice coefficients (DCs) of myelin-like signals (MLS) in the brainstem over 16 test subjects obtained by thresholding and the model-based methods (see Table 5.1 for method descriptions) using a range of initial thresholds. GMM-PV-MRF performs most stably and most accurately with nearly constant DCs throughout the threshold range from the 21st to 40th percentile. The DCs of thresholding are much less affected by the varying threshold because the brainstem contains more myelinated tissue than the thalami. GMM-PV helps to stabilise the variations shown by GMM in the presence of substantial partial volume voxels, but nonetheless lowers the DCs of the baseline initial segmentations obtained from thresholding.

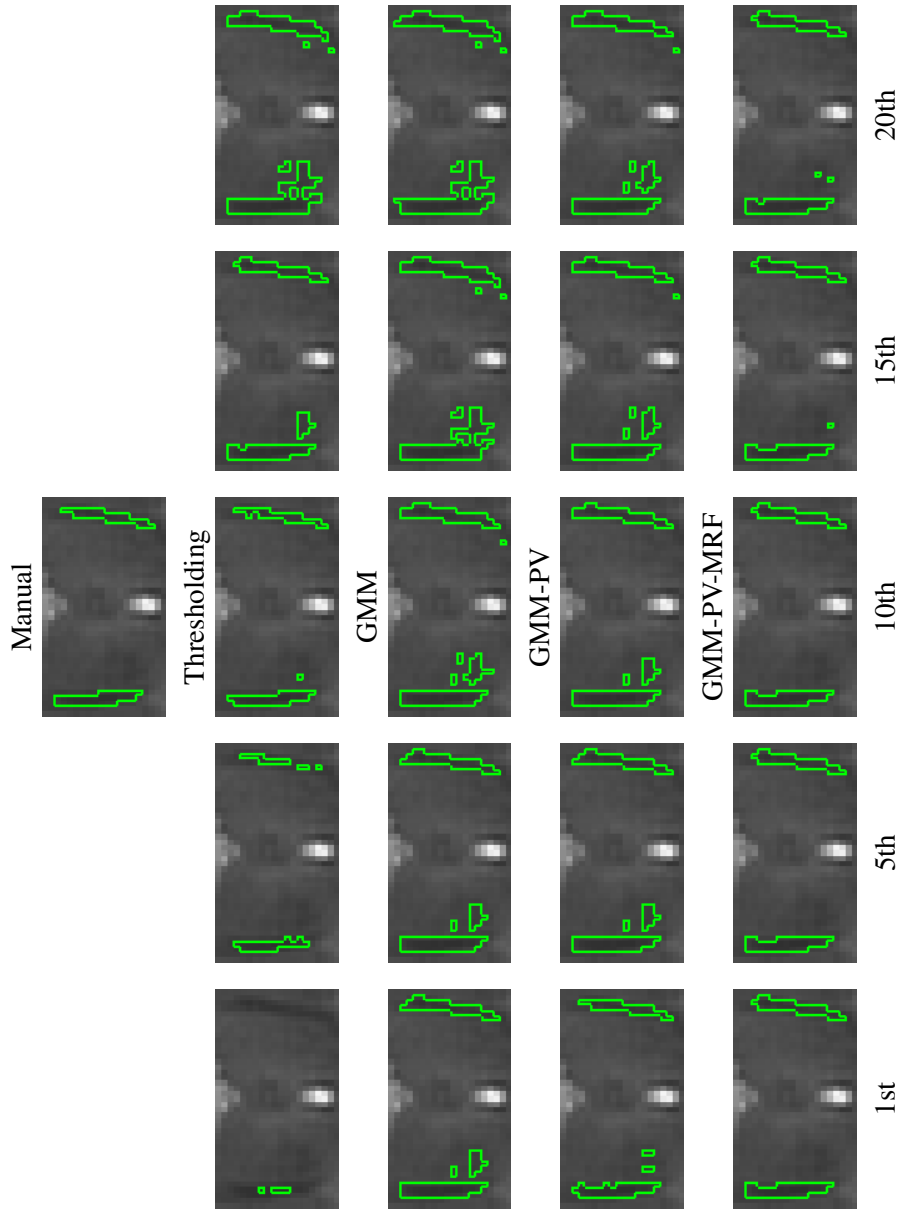


Figure 5.16: Coronal images of myelin-like signals (MLS) in the thalami segmented by thresholding and the model-based methods (see Table 5.1 for method descriptions) at the 1st, 5th, 10th, 15th and 20th percentile thresholds for a subject at 43 gestational weeks. GMM-PV-MRF segments the posterior limbs of the internal capsule (see Fig. 5.1) accurately with respect to the manual annotation regardless of the MLS volume fraction used for initialisation. Both GMM and GMM-PV are unable to remove the unmyelinated, low-intensity regions included in the initial MLS segmentations as the threshold increases above the optimal 6th percentile.

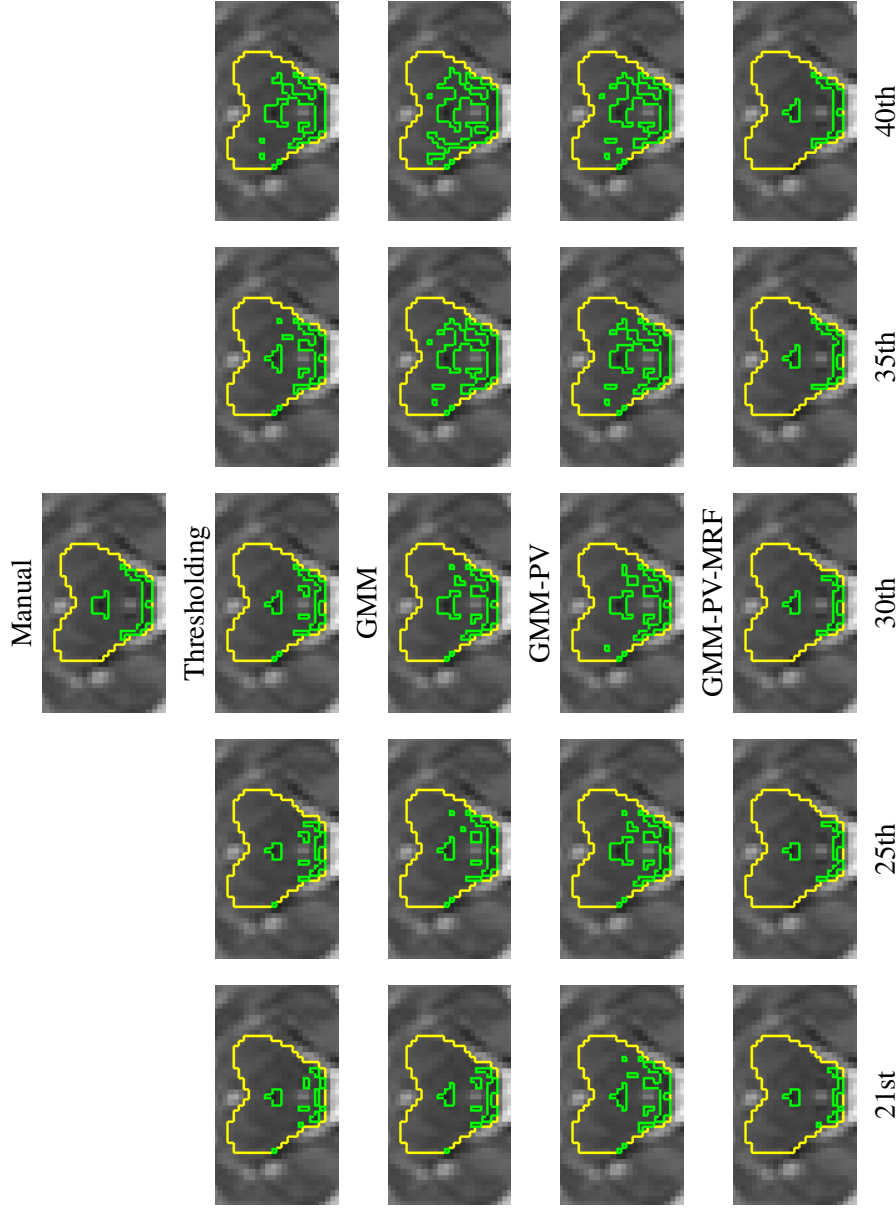


Figure 5.17: Axial images of myelin-like signals (MLS) in the brainstem segmented by thresholding and the model-based methods (see Table 5.1 for method descriptions) at the 21st, 25th, 30th, 35th and 40th percentile thresholds for a subject at 34 gestational weeks. The MLS segmentations produced by thresholding, GMM and GMM-PV all vary greatly with the volume fraction of the initial MLS segmentation and contain many mislabelled voxels due to the absence of spatial regularisation, whereas GMM-PV-MRF segments well-defined MLS structures that remain nearly invariant over the threshold range.

Lastly, we verified that the 6th and 25th percentiles of the T_2w intensities in the thalami and brainstem, respectively, optimised the DCs that could possibly be achieved using a single threshold for all the 16 subjects.

5.4.4 Comparison to the intra-rater reliability

We assessed the intra-rater reliability as Dice overlap between two sets of manual annotations that were available for eight of the 16 test subjects at two-week intervals between 30 and 44 weeks GA. We evaluated the automated model-based methods with respect to each manual set. The average DCs over the eight subjects are shown in Table 5.4 along with the p -values of two-tailed Student's t -tests with the intra-rater DCs at the 5% significance level. It can be seen that only GMM-PV-MRF performed equally well as a human rater in both ROIs, and demonstrated no significant differences with respect to the repeated manual annotations.

5.4.5 Volume growths of myelin-like signals

Based on the results in Section 5.4, we applied GMM-PV-MRF as proposed to segment MLS in the ROIs for 114 preterm infants aged between 29 and 44 weeks GA. The volumes of the resultant segmentations in the thalami and brainstem were plotted against GA in Figs. 5.18 and 5.19 respectively.

We found that the MLS volume appeared to grow exponentially with GA in the thalami, and this trend persisted in the MLS volume fraction after we corrected for the different thalamus sizes in individual subjects. In contrast, the MLS volume appeared to grow linearly with GA in the brainstem, which resulted in an approximately constant MLS volume fraction after correcting for the different brainstem sizes in individual subjects.

This is because most of the brainstem appears to have completed myelination before 29 weeks GA, whereas new brain structures become myelinated in the thalami between 29 and 44 weeks GA (Chi et al., 1977; Counsell et al., 2002; Rutherford, 2002). Consequently, the MLS volume increased in the brainstem mainly due to the overall brain growth, whereas the increase in the thalami was attributed to both brain growth and ap-

Method	Manual set 1		Manual set 2		Intra-rater reliability
	Average DC	<i>p</i> -value	Average DC	<i>p</i> -value	
GMM	0.785 ± 0.085	0.257	0.754 ± 0.054	0.023	0.834 ± 0.043
GMM-PV	0.662 ± 0.210	0.060	0.626 ± 0.217	0.028	
GMM-PV-MRF	0.843 ± 0.051	0.624	0.816 ± 0.046	0.413	

(a) Comparison results in the thalami

Method	Manual set 1		Manual set 2		Intra-rater reliability
	Average DC	<i>p</i> -value	Average DC	<i>p</i> -value	
GMM	0.702 ± 0.171	0.046	0.695 ± 0.156	0.026	0.842 ± 0.034
GMM-PV	0.763 ± 0.047	0.002	0.766 ± 0.077	0.031	
GMM-PV-MRF	0.832 ± 0.034	0.608	0.815 ± 0.024	0.087	

(b) Comparison results in the brainstem

Table 5.4: Average Dice coefficients (DCs) ($\pm$ standard deviations) over eight test subjects obtained by thresholding and the model-based segmentation methods (see Table 5.1 for method descriptions) with respect to two sets of manual annotations compared to the intra-rater reliability. The *p*-values indicate the outcomes of two-tailed Student's *t*-tests with the intra-rater DCs at the 5% significance level. Only GMM-PV-MRF performs equally well as a human rater in both regions of interest.

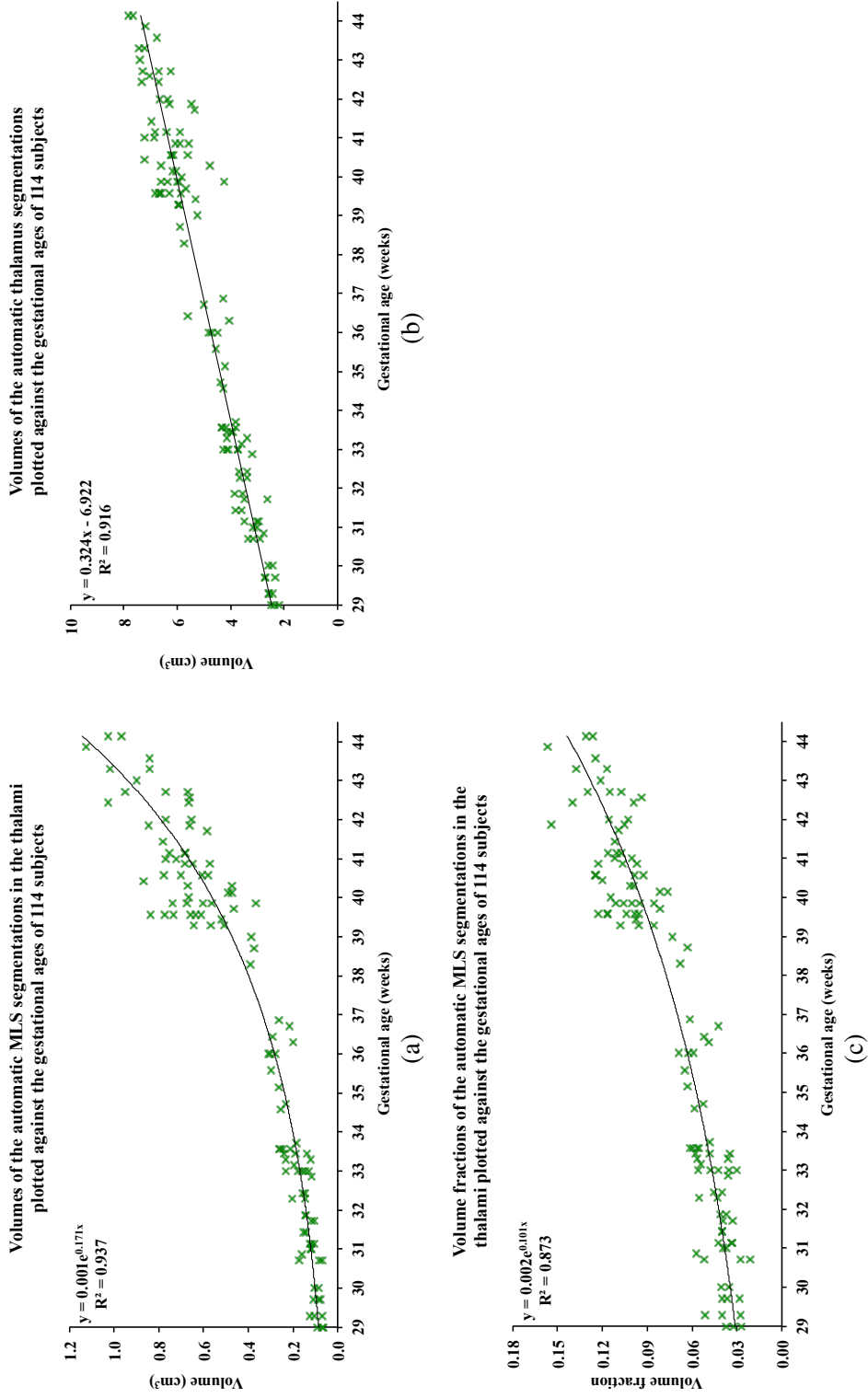


Figure 5.18: (a) Volumes of the automatic segmentations for myelin-like signals (MLS) in the thalami obtained using the proposed method GMM-PV-MRF, (b) volumes of the automatic thalamus segmentations obtained during image pre-processing, and (c) volume fractions of MLS plotted against the gestational ages (GAs) of 114 subjects between 29 and 44 weeks GA. The MLS volume appears to grow exponentially with GA, and this trend persists in the MLS volume fraction after correcting for the different sizes of the thalami in individual subjects.

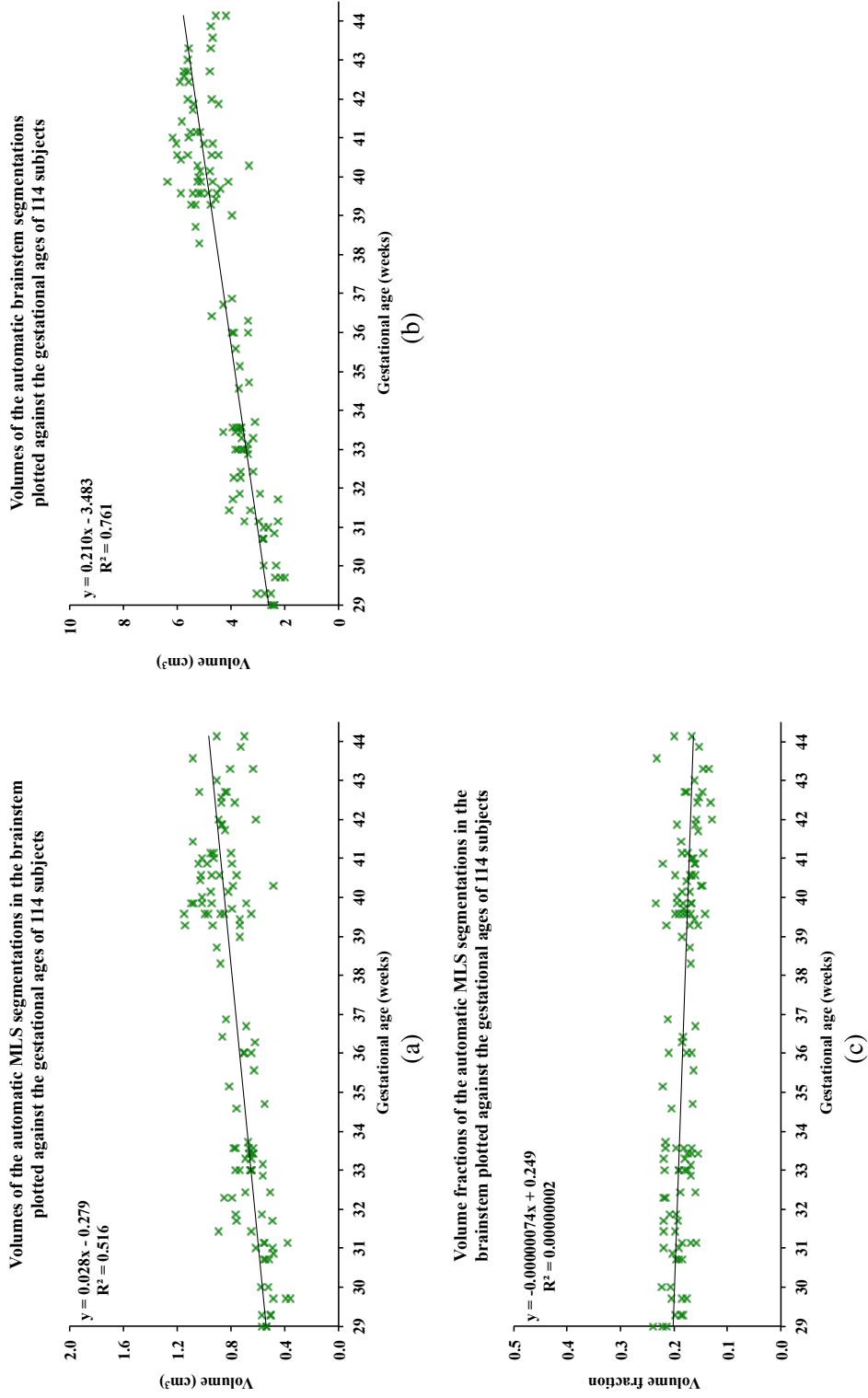


Figure 5.19: (a) Volumes of the automatic segmentations for myelin-like signals (MLS) in the brainstem obtained using the proposed method GMM-PV-MRF, (b) volumes of the automatic brainstem segmentations obtained during image pre-processing, and (c) volume fractions of MLS plotted against the gestational ages (GAs) of 114 subjects between 29 and 44 weeks GA. The MLS volume appears to grow linearly with GA, which results in an approximately constant MLS volume fraction after correcting for the different sizes of the brainstem in individual subjects.

pearance of new myelination. In Chapter 6, we will demonstrate the different myelination progressions in the ROIs by constructing spatio-temporal growth models for MLS.

5.5 Conclusion

In this chapter, we presented an EM segmentation framework for the tissue likely to contain myelin, which we referred to as myelin-like signals (MLS) throughout the paper, on T₂w brain MR images of preterm neonates.

The two major challenges of this task are the inadequate spatial prior information and the large number of PV voxels typical for the developing brain. The publicly available neonatal brain atlases (Kuklisova-Murgasova et al., 2011; Serag et al., 2012) do not include a probability map for myelin, and manual annotations are impractical in clinical use or for large-scale researches. The PV voxels shift the intensities of MLS towards BKG, which makes the EM optimisation process unstable and causes over-estimation of the MLS class.

To overcome these difficulties, we introduced an additional Gaussian to explicitly model the PV voxels. Van Leemput et al. (2003) observed that, without any spatial prior information, the extra Gaussian hampers the segmentation robustness. The previous PV estimation methods thus used prior knowledge of the spatial locations of the composing pure tissues to guide the search for PV voxels (Shattuck et al., 2001; Van Leemput et al., 2003; Tohka et al., 2004; Manjon et al., 2010; Cardoso et al., 2011; Ledig et al., 2012). Because there are no probabilistic or manual atlases available for MLS, we identified the PV locations entirely based on tissue connectivities among the triplet of classes in the local neighbourhood via second-order MRFs. This approach performed robustly and accurately with respect to manually annotated MLS in different ROIs for subjects between 29 and 44 weeks GA, regardless of the volume fraction used to obtain the initial MLS segmentations. This is crucial in practice as the choice of the initial threshold is difficult without manual annotations. The independence of the automatic MLS segmentations against the initial threshold is also important when assessing deviation from the average MLS growth model in order to estimate the brain maturation status.

One limitation is that we have not utilised all the multi-spectral imaging information that is available. Myelinated GM and WM structures are shown differently using T_1w and T_2w sequences (Barkovich, 1998; Counsell et al., 2002). In general, myelinated GM structures, such as the thalamic and brainstem nuclei, are more conspicuous on T_2w images whereas myelinated WM structures, particularly the axonal tracts, are better visualised on T_1w images (Barkovich, 1998). Here we only assessed MLS based on T_2w acquisitions, because the PLIC tracts were shown clearly on the T_2w images in this particular dataset. In Chapter 4, we did include T_1w images for segmentation of myelinated white matter in the NeoBrainS12 challenge (Wang et al., 2012), but the purely intensity-based multivariate GMM resulted in considerable over-estimation without either explicit PV modelling or spatial regularisation. In the future, we will incorporate both T_1w and T_2w information in the current EM framework with explicit PV modelling via second-order MRFs.

We will demonstrate in Chapter 6 that our method of MLS segmentation provides the basis for the spatial-temporal analysis that can track brain maturation in fragile preterm neonates using routinely acquired MR images. In particular, a growth model constructed based on the automatic MLS segmentations of the 114 infants can be used to estimate the deviation from the normal myelination trajectory for individual subjects. This can be useful in clinical practice for preterm neonates with subtle brain maturational abnormalities, who would benefit from early medical interventions soon after birth.

Chapter 6

Spatio-temporal analysis for assessing neonatal brain maturation

6.1 Introduction

Preterm neonates born before 37 gestational weeks are commonly assessed on the basis of birth weight. However, Rorke and Riggs (1969) observed histologically that, in some neonates, the brain appears more mature than might be anticipated from body weight, and neonates who are well developed according to the measurement of body weight may also experience delayed myelination. As explained in Chapter 3, Dittrich et al. (2014) proposed an iterative algorithm that integrates fetal brain segmentation, atlas construction as well as age estimation. The morphological age of a fetus, rather than the nominal age, was estimated by maximising the similarity between its brain segmentation and the spatio-temporal atlas. The experiments on fetuses with abnormal cortical development demonstrated that the algorithm is able to detect significant differences compared to a control group.

At the end of Chapter 5, we showed the evidence that myelin-like signals (MLS) on T₂w neonatal brain MR images are closely associated with ongoing myelination, even though the signal properties reflect several maturational processes that occur simultane-

The work described in this chapter has been published in the *Proceedings of the 13th International Symposium on Biomedical Imaging (ISBI)* (Wang et al., 2016).

ously (Korogi et al., 1996). More importantly, these conventional MR images are routinely acquired during clinical imaging of fragile preterm neonates (Rutherford et al., 2006; Srinivasan and Rutherford, 2008). Therefore, the aim of this chapter is to correlate MLS on T_2w images with gestational age (GA), so that brain maturation in preterm neonates can be assessed directly from clinical MRI data.

The chapter is organised as follows. Section 6.3 describes a pipeline where we construct a spatio-temporal growth model for MLS in a predefined region of interest (ROI) using logistic regression, and use the obtained model for age estimation. In Section 6.4, we apply the proposed pipeline to characterise the MLS development in two ROIs, namely the thalami and brainstem, based on 114 preterm infants aged between 29 and 44 weeks GA. The logistic MLS growth models in both ROIs are compared with those constructed by a commonly used kernel regression method (Kuklisova-Murgasova et al., 2011). We evaluate both regression models by estimating the GAs of the 114 subjects using a leave-one-out procedure.

The spatio-temporal trajectory of MLS is consistent with the known myelination patterns from both histological studies (Chi et al., 1977; Kinney et al., 1988) and clinical MRI observations (Counsell et al., 2002; Rutherford, 2002). In particular, the logistic model successfully captures the onset of myelination in the posterior limbs of the internal capsule (PLIC) at approximately 40 weeks GA, which is an important landmark for evaluating postnatal brain development (Bird et al., 1989). Compared to the kernel-based approach, we demonstrate that the logistic regression model is less affected by the distribution of subjects at different ages, and hence predicts the GAs with a smaller root mean squared error (RMSE).

6.2 Materials

We use the same dataset as in Chapter 5, which includes 114 preterm infants scanned between 29 and 44 weeks GA. The parameters of the T_2w fast spin-echo acquisition can be found in Section 5.2.1.

6.3 Analysis pipeline

In Fig. 6.1, we present a pipeline for assessment of neonatal brain maturation from T₂w images. During the pre-processing step, we correct the image bias field and segment MLS in a predefined ROI for the 114 subjects in the dataset (see explanations in Chapter 5). A spatio-temporal growth model for MLS in the ROI is then constructed using logistic regression based on the automatic segmentations. We can assess the brain maturation status in a given subject by estimating its GA using the MLS growth model. In the next section, this pipeline will be applied separately in the ROIs of the thalami and brainstem where the majority of myelination occurs in the perinatal period.

6.3.1 Construction of the spatio-temporal growth model

6.3.1.1 Transformation to the common reference space

The MLS segmentations obtained from Chapter 5 consist of continuous fractions of the MLS class calculated using Eq. 5.18 in the spaces of individual subjects. The first step is to transform all the segmentations to a common reference space.

Using free-form deformation (FFD) non-rigid registration (Rueckert et al., 1999), we register the bias corrected T₂w image of each subject with the dilated ROI of a single reference subject at 36 weeks GA chosen from the dataset. We use normalised mutual information (NMI) as the similarity measure and 10 mm control point spacing to achieve the optimal alignment of the myelinated brain structures in the ROI. The MLS segmentations are then transformed accordingly from each subject's space to the common reference space. The calculation of an average space is omitted in this study because it does not affect our analysis. The aim is to co-align the images in the common reference space and to identify the anatomical correspondence of the brain structures in the thalami and brainstem.

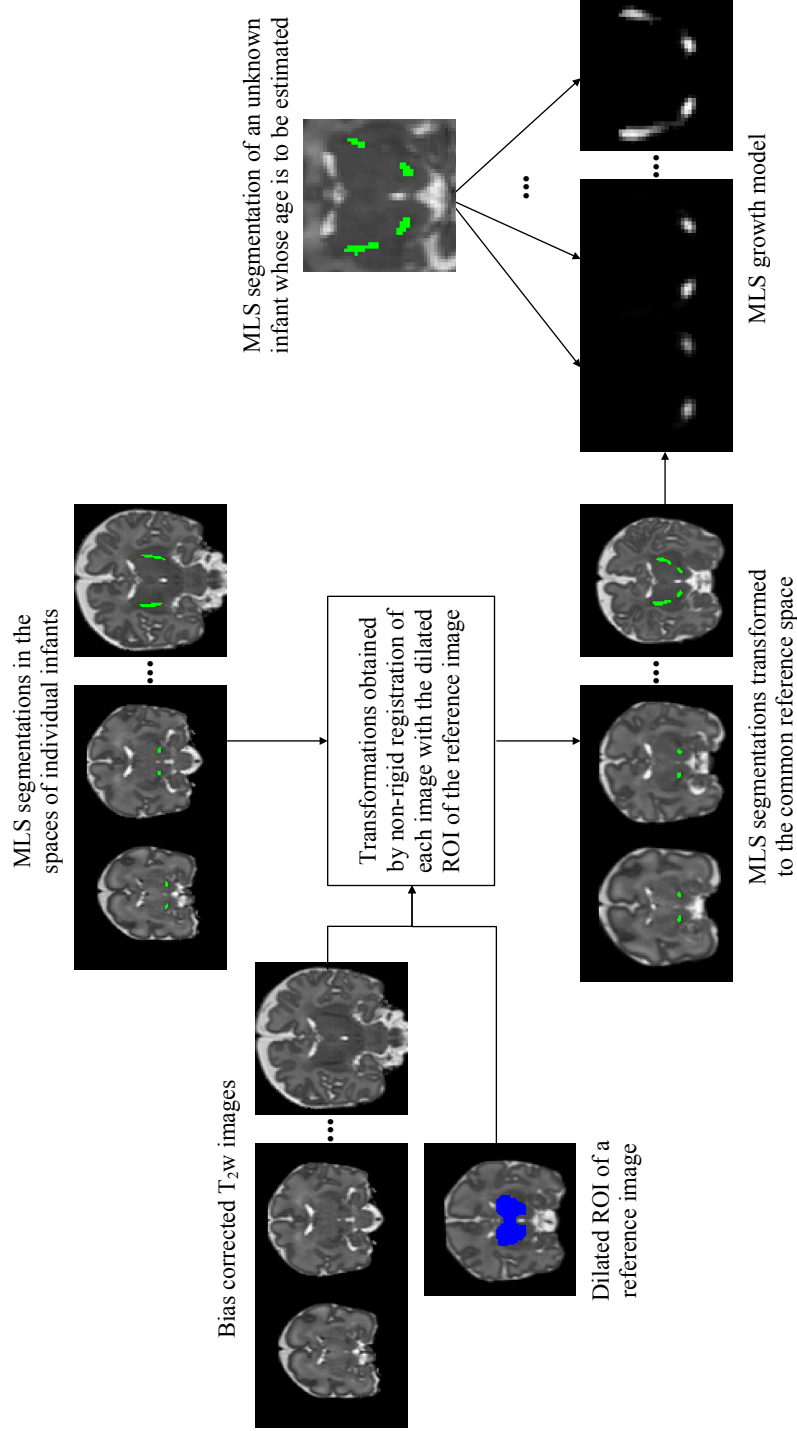


Figure 6.1: Schematic flow of the proposed spatio-temporal analysis. Firstly, myelin-like signals (MLS) are segmented in a predefined region of interest (ROI) for 114 preterm infants in the dataset. These segmentations are transformed to a common reference space by non-rigid registration (Rueckert et al., 1999) of each bias corrected T_{2w} image with the dilated ROI of a reference image. An MLS growth model is then constructed using logistic regression based on the transformed segmentations. Lastly, the gestational age of a given subject can be estimated by assessing the deviation between its MLS segmentation and the average growth model. In this study, we choose the thalami and brainstem as separate ROIs for assessing myelination in the perinatal period.

6.3.1.2 Logistic regression

We fit a voxelwise logistic function to the transformed segmentations to simulate the average change of the MLS fraction over time at each location:

$$f_i(t) = \frac{\gamma_i}{1 + e^{-(\alpha_i + \beta_i t)}} \quad (6.1)$$

Here $f_i(t)$ denotes the average fraction of the MLS class at voxel i in the reference space, modelled as a function of GA t . The parameter γ_i corresponds to the final MLS fraction when myelination is complete in voxel i , as shown in Fig. 6.2. Therefore, the voxels with γ_i values less than one are likely to contain PV, whereas the estimated values of γ_i at non-PV voxels are equal to one. The time point t_i^A in Fig. 6.2 is determined by the negative ratio between the parameters α_i and β_i :

$$t_i^A = -\frac{\alpha_i}{\beta_i} \quad (6.2)$$

We consider t_i^A a critical time point because it is associated with the maximum growth rate at voxel i , and marks the GA when the anatomical location becomes myelinated on average. Thus, at time point t_i^A :

$$\left(\frac{df_i}{dt} \right)_{max} = \frac{\gamma_i \beta_i}{4} \quad (6.3)$$

We optimise the parameters α_i , β_i and γ_i using a Levenberg-Marquardt non-linear least squares algorithm (Levenberg, 1944).

When constructing the growth model, we only use the voxels containing sufficient MLS where the rapid exponential growth of the fitted logistic function occurs largely in the age range between 29 and 44 weeks GA, and discard the others as outliers. Therefore, the parameter estimates at the selected voxels need to fulfil the following conditions:

- (i) $0.5 \leq \gamma_i \leq 1$: The value of γ_i needs to be greater than or equal to 0.5 to ensure that voxel i will have contained sufficient MLS by the end of the myelination process.

We reject the PV voxels that cannot reach a final MLS fraction of 0.5.

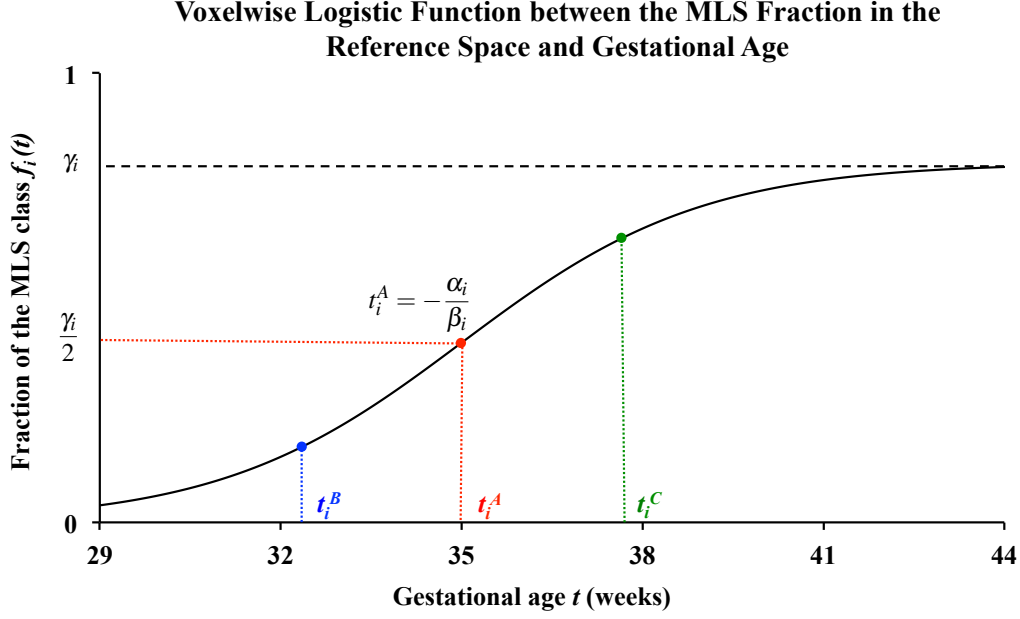


Figure 6.2: Voxelwise logistic function between the fraction of myelin-like signals (MLS) in the reference space and gestational age (GA). The parameter γ_i implies the ultimate MLS fraction contained in voxel i . The time point t_i^A is associated with the maximum growth rate at voxel i , and marks the GA when the anatomical location becomes myelinated on average. t_i^B and t_i^C correspond to the maximum and minimum curvatures respectively.

- (ii) $\beta_i > 0$: We further require β_i and hence the maximum slope in Eq. 6.3 to be positive, so that the fitted logistic function increases monotonically.
- (iii) $24 \leq t_i^A \leq 43$: According to the myelination sequence described in Counsell et al. (2002), myelin is visible on T₂w images as early as 25 weeks GA in the posterior brainstem and the ventrolateral nuclei (VLN) of the thalami. Myelination progresses towards the superior and anterior regions of the brain, and becomes evident in the PLIC tracts at approximately 40 weeks GA, followed by the corticospinal tracts at 42 weeks GA. Thus based on this prior knowledge, we only include the voxels where the t_i^A estimates are between 24 and 43 weeks GA.

6.3.2 Estimation of gestational ages

We can estimate the GA of a given subject, as compared to the nominal GA, based on the underlying brain maturation status by evaluating the deviation between its MLS seg-

mentation and the average growth model. We calculate the sum of squared differences (SSD) between the growth model and the MLS segmentation, and minimise the SSD with respect to the GA t using gradient descent:

$$SSD(t) = \frac{1}{I} \sum_{i=1}^I (f_i(t) - s_i^T)^2 \quad (6.4)$$

$$\hat{t} = \arg \min_t SSD(t) \quad (6.5)$$

The MLS fraction $f_i(t)$ is evaluated at time point t and voxel i from the fitted logistic function using Eq. 6.1. s_i^T is the MLS fraction in the subject's transformed segmentation in the reference space, and I the number of voxels.

To improve the accuracy of age estimation, we only include the voxels that experience the full range of change in the MLS fraction from zero to γ_i between 29 and 44 weeks GA when calculating the SSD. Thus we introduce the time points t_i^B and t_i^C in Fig. 6.2, associated with the maximum and minimum curvatures respectively:

$$t_i^{B,C} = -\frac{\ln(2 \pm \sqrt{3}) + \alpha_i}{\beta_i} \quad (6.6)$$

We require both t_i^B and t_i^C to be between 29 and 44 weeks GA, and replace Condition (iii) in Section 6.3.1 with $29 \leq t_i^{B,C} \leq 44$ while the other constraints remain unchanged.

6.4 Results and discussion

We applied the proposed pipeline in the thalami and brainstem separately based on the 114 preterm infants in the dataset, and repeated the analysis using a common kernel-based approach (Kuklisova-Murgasova et al., 2011). In this section, we first confirmed the transformations obtained from FFD non-rigid registration (Rueckert et al., 1999) were diffeomorphic. We then examined the logistic simulation of MLS progression against the known myelination patterns in preterm infants (Counsell et al., 2002; Kinney et al., 1988). We visually compared the growth models constructed using the logistic and kernel-based methods respectively, which were further evaluated by predicting the GAs of the 114

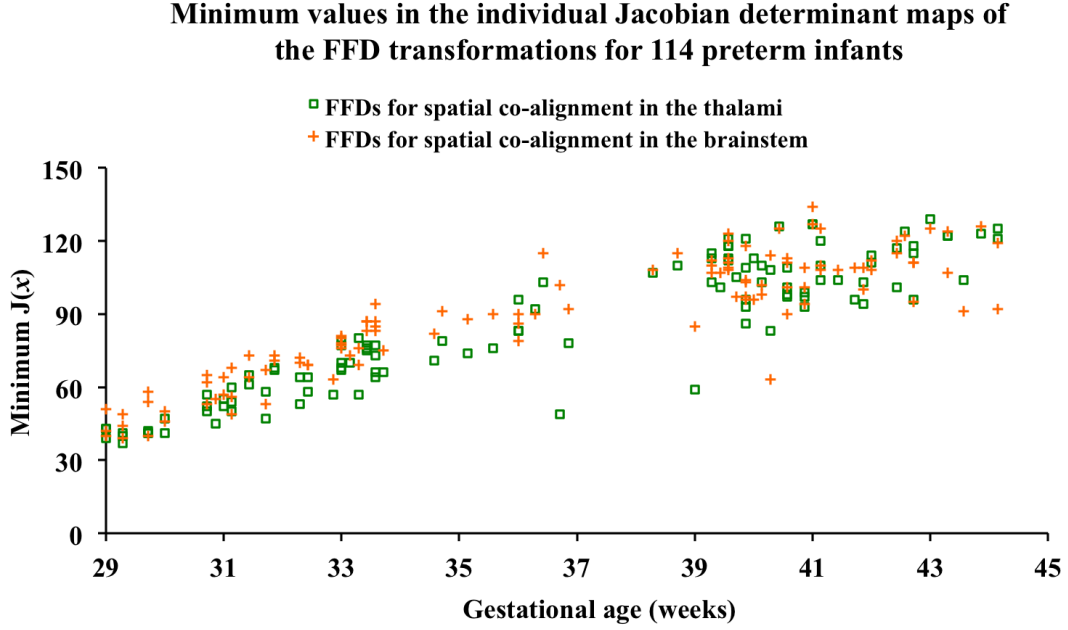


Figure 6.3: Minimum values in the individual Jacobian determinant maps of the free-form deformation (FFD) transformations computed for co-aligning the regions of thalami and brainstem, respectively, for 114 preterm infants. Since the minimum Jacobian determinants are all positive, the transformations are confirmed to be diffeomorphic and preserve topology.

subjects using a leave-one-out procedure.

6.4.1 Jacobian determinants of the FFD transformations

As mentioned in Section 3.2.4.4, diffeomorphic registration algorithms preserve topology. We also explained in Section 3.5.2 that the Jacobian determinant can be used as a measure of the volume change induced by the transformation, and is greater than zero for topology preserving transformations.

Here we generated maps of the Jacobian determinants for the total FFD transformations, including the global affine transformations and the local B-spline displacement fields, that were computed for achieving spatial co-alignment in the thalami and brainstem respectively. In Fig. 6.3, we plotted the minimum values in the individual Jacobian determinant maps against the GAs of the 114 infants for each ROI. We confirmed that the transformations later used to construct the MLS growth model in the thalami and brainstem were diffeomorphic. This result was consistent with our expectation because we had

applied large control point spacing to restrict the transformations to the global shape of the central brain structures without being affected by the local contrast variations.

6.4.2 Logistic simulation results compared to the known myelination patterns

6.4.2.1 Spatio-temporal trajectories of MLS in the thalami and brainstem

We chose the thalami and brainstem as the ROIs for assessing neonatal brain maturation. The spatio-temporal trajectories of MLS in both ROIs are shown in Figs. 6.4 and 6.5 respectively for preterm infants aged between 29 and 44 weeks GA.

We found previously in Section 5.4.5 that the volume fraction of MLS in the thalami appeared to increase exponentially with GA between 29 and 44 weeks GA, whereas the brainstem maintained a nearly constant MLS volume fraction. The spatio-temporal modelling results further explained the volume calculations, and confirmed that there was a marked difference in the onset and rate of myelination in the two ROIs.

It can be seen from Fig. 6.4 that the VLN and subthalamic nuclei (STN) in the thalami were likely to be myelinated before 29 weeks GA. Moreover, the logistic model successfully captured the arrival of MLS in both of the PLIC tracts at approximately 40 weeks GA. This is an important landmark for evaluating neonatal brain development (Bird et al., 1989). The brainstem in Fig. 6.5, however, appeared to be well developed by the time of 29 weeks GA, and there was relatively little new myelination in the age range of our investigation.

Therefore, MLS progressed from inferior to superior, and from posterior to anterior, which was consistent with the general myelination sequence in the central nervous system during human infancy (Kinney et al., 1988). Note that because we corrected for the different sizes of each ROI in individual subjects through non-rigid registration (Rueckert et al., 1999), the simulated MLS progression was entirely attributed to the appearance of new myelinated brain structures without contribution from the overall brain growth.

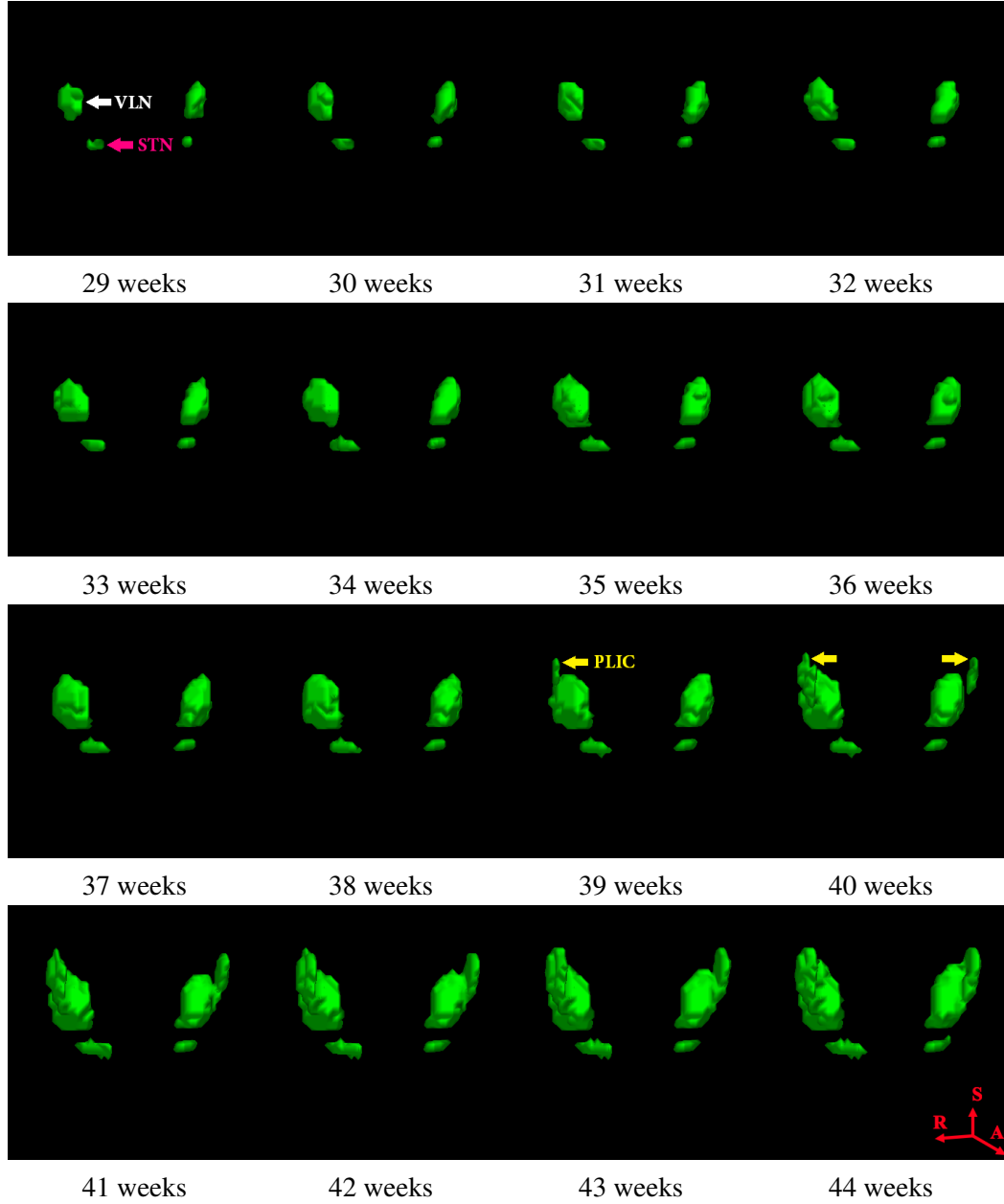


Figure 6.4: Logistic simulation of the spatio-temporal trajectory for myelin-like signals (MLS) in the thalami between 29 and 44 weeks gestational age (GA). We binarise the MLS fraction at 0.5 for presentation in 3D surface rendering. The ventrolateral nuclei (VLN, white arrow) and subthalamic nuclei (STN, pink arrow) appear to be myelinated before 29 weeks GA. MLS becomes evident in both tracts of the posterior limbs of the internal capsule (PLIC, yellow arrow) at approximately 40 weeks GA, which is an important landmark for evaluating neonatal brain development. As the different thalamus sizes in individual subjects have been corrected through non-rigid registration, the progression of MLS is completely due to the appearance of new myelination without contribution from the overall brain growth.

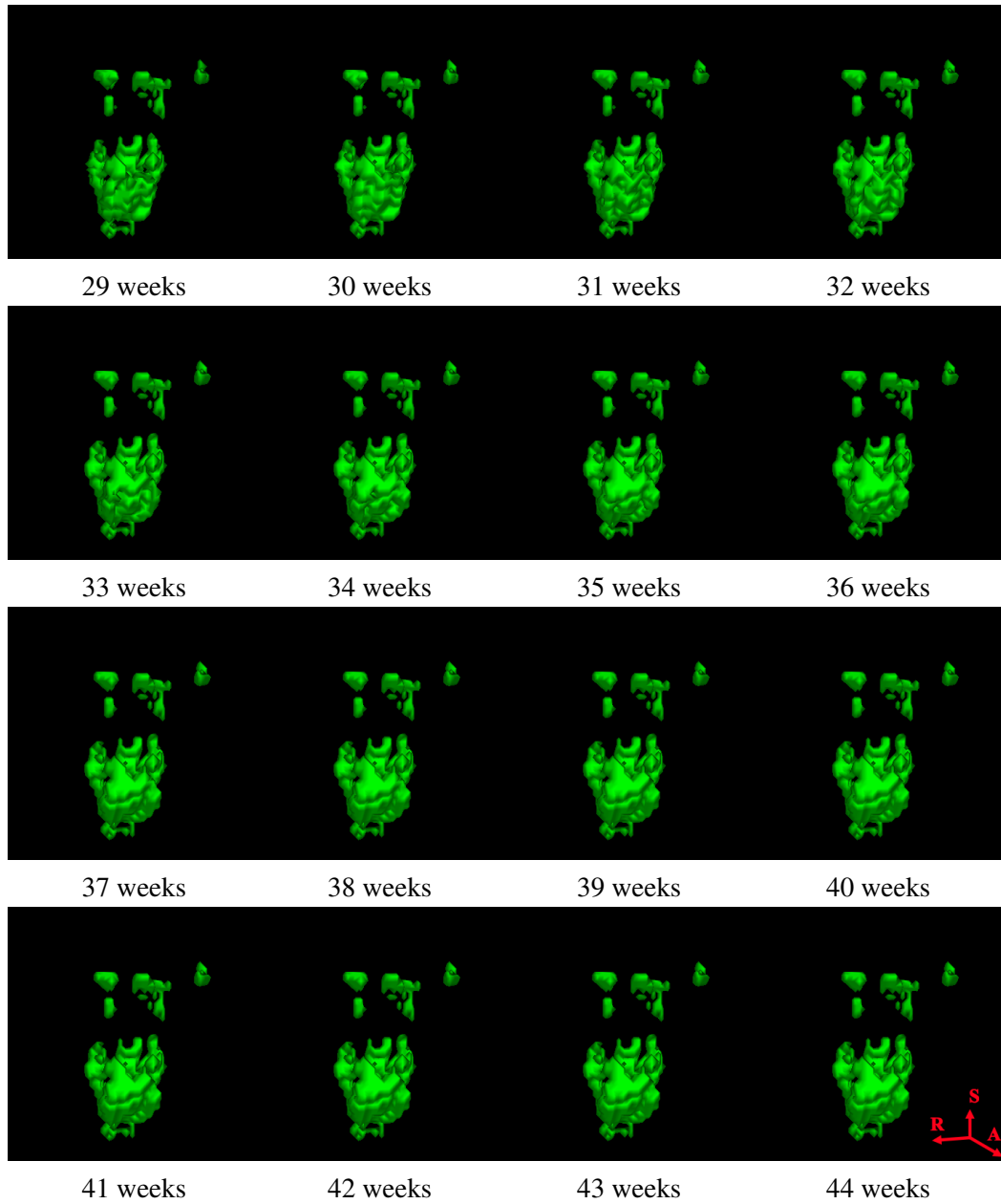


Figure 6.5: Logistic simulation of the spatio-temporal trajectory for myelin-like signals (MLS) in the brainstem between 29 and 44 weeks gestational age (GA). We binarise the MLS fraction at 0.5 for presentation in 3D surface rendering. The brainstem appears to be well developed by the time of 29 weeks GA with relatively little MLS progression compared to the thalamus region. Again the overall increase in the brainstem size has been corrected through non-rigid registration.

6.4.2.2 3D parameter maps

In order to investigate the regional differences in the myelination process, we extracted the optimised logistic parameters at each voxel i , and generated 3D maps in both ROIs for:

- 1) The average myelination age t_i^A (Fig. 6.6);
- 2) The time interval Δt_i between $0.25\gamma_i$ and $0.75\gamma_i$ (Fig. 6.7);
- 3) The ultimate MLS fraction for full growth γ_i (Fig. 6.8).

As explained in Section 6.3.1, t_i^A marks the GA when voxel i becomes myelinated on average. Therefore, the only myelination site that emerged between 29 and 44 weeks GA was the characteristic PLIC. The spatial relation and myelination sequence between the VLN and PLIC further suggested that myelinated WM fibres arose from the VLN, and proceeded radially until reaching the PLIC tracts at approximately 40 weeks GA. In addition, we confirmed that the brainstem matured earlier than the thalami with all the t_i^A values close to or below 30.

We compared the t_i^A map with the myelination timetable described by Counsell et al. (2002) for preterm infants based on clinical observations from T₂w MRI. The expected myelination GAs for the labelled brain structures in Fig. 6.6 are listed in Table 6.1, which highly agreed with the t_i^A values at the corresponding locations.

Region of interest	Brain structure	Myelination age (gestational weeks)
Thalami	Ventrolateral nuclei	25
	Subthalamic nuclei	28
	Posterior limbs of the internal capsule	40
Brainstem	Vestibular nuclei	25
	Medial longitudinal fasciculi	29
	Medial lemnisci	30

Table 6.1: Myelination timetable for the labelled brain structures in Fig. 6.6, described by Counsell et al. (2002) for preterm infants based on clinical observations from T₂w MRI.

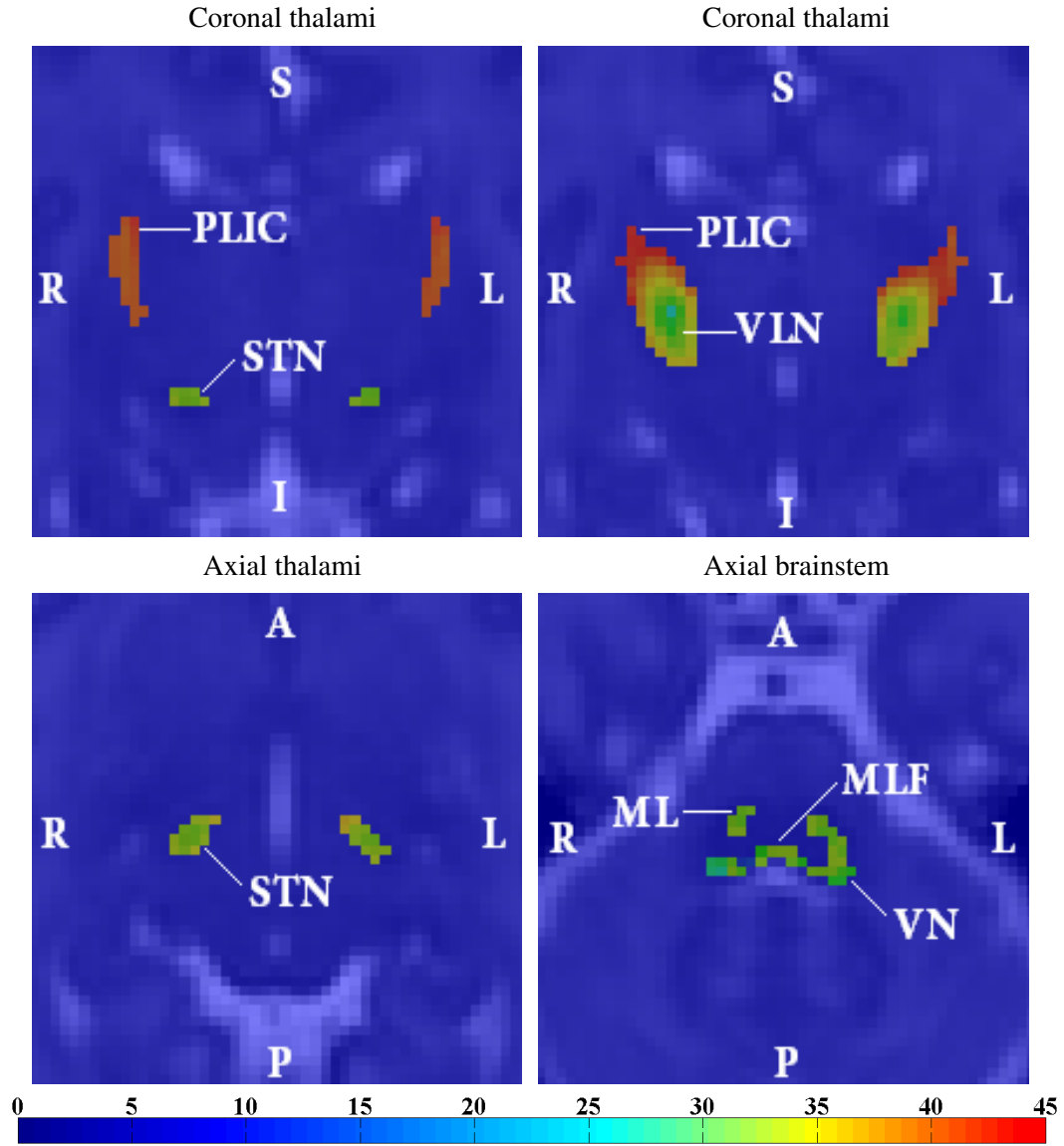


Figure 6.6: 3D map for the average myelination age t_i^A in the thalami and brainstem. The only myelination site that emerges between 29 and 44 weeks GA is the characteristic PLIC. Myelinated white matter fibres appear to arise from the VLN, and proceed radially until reaching the PLIC tracts at approximately 40 weeks GA. In contrast, the brainstem appears to develop earlier than the thalami with all the t_i^A values close to or below 30. Abbreviations: PLIC–posterior limb of the internal capsule, VLN–ventrolateral nucleus, STN–subthalamic nucleus, ML–medial lemniscus, MLF–medial longitudinal fasciculus, VN–vestibular nucleus.

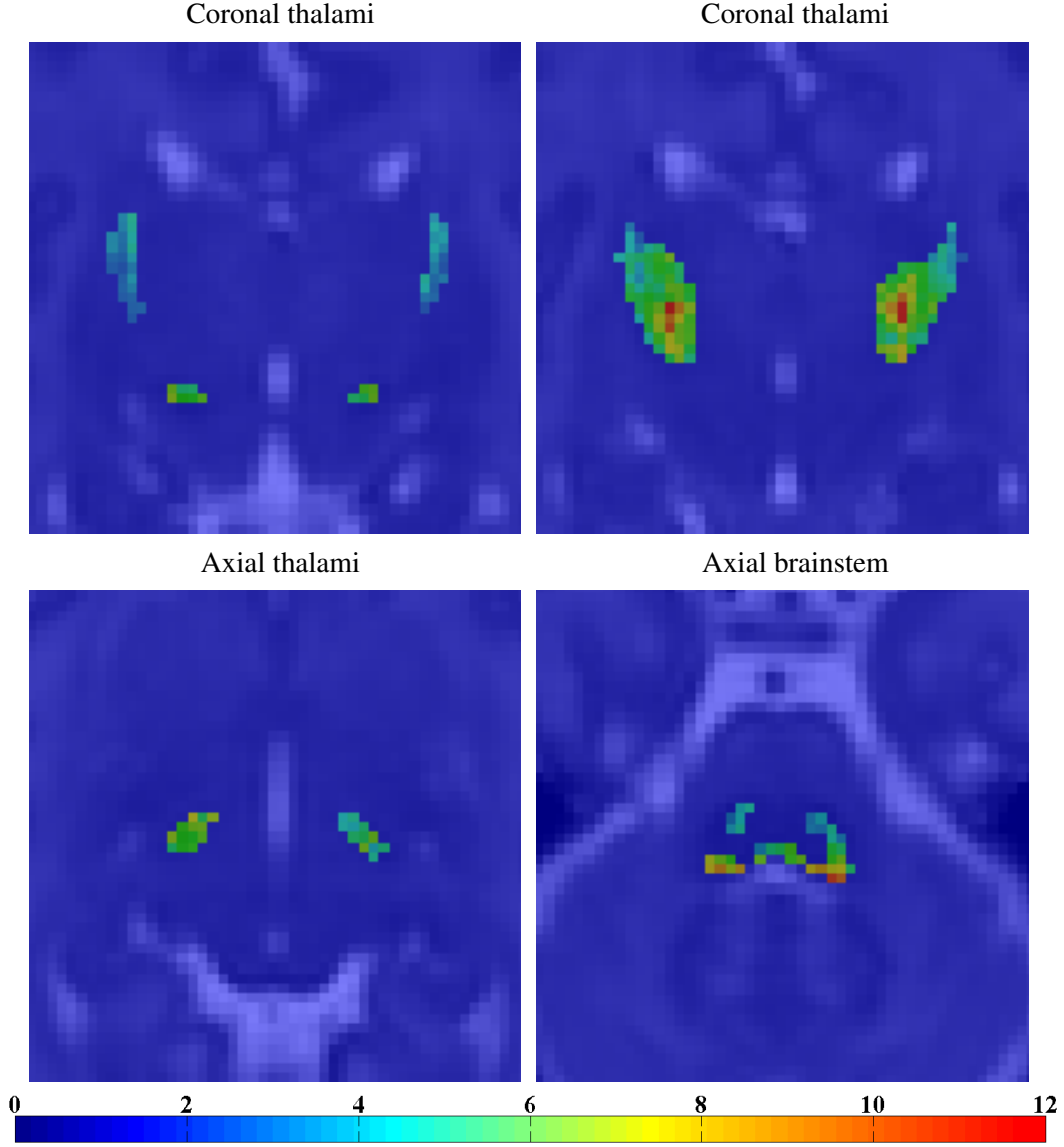


Figure 6.7: 3D map for the time interval Δt_i between $0.25\gamma_i$ and $0.75\gamma_i$ in the thalami and brainstem. The centre of the VLN displays the slowest maturation rate, and the fraction of myelin-like signals (MLS) increases from $0.25\gamma_i$ to $0.75\gamma_i$ in nearly 12 weeks, despite the fact that the same location becomes myelinated the earliest at approximately 25 weeks GA. The fastest MLS growth with Δt_i values below 4 weeks occurs in the PLIC and ML, which confirms the general myelination patterns towards the superior and anterior of the brain. The brain structures have been labelled and the abbreviations explained in Fig. 6.6.

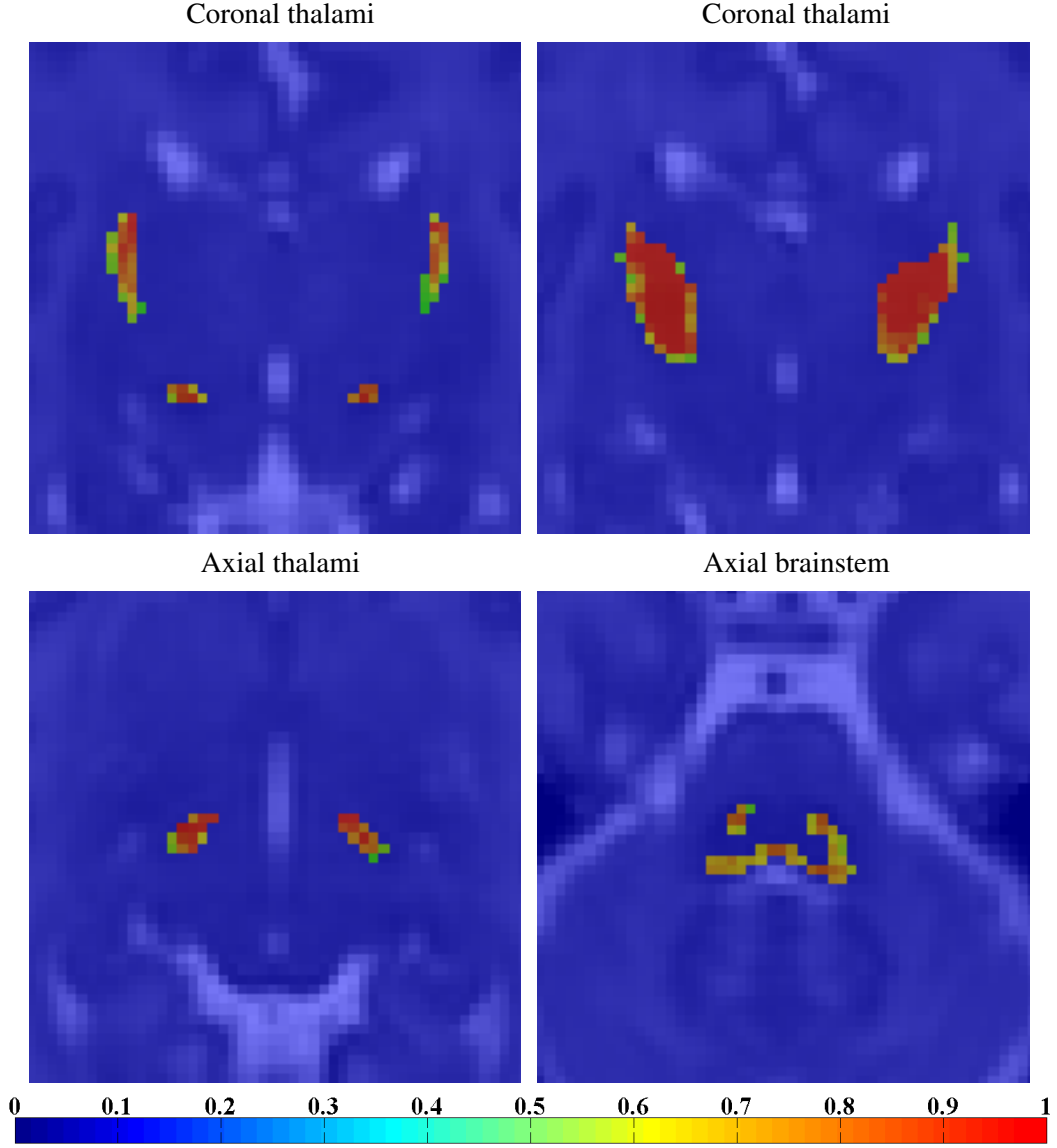


Figure 6.8: 3D map for the ultimate fraction of myelin-like signals γ_i (see Fig. 6.2) in the thalami and brainstem. Non-partial volume (PV) voxels with γ_i values equal to one are commonly found in large and well defined nuclei, such as the VLN, whereas small and irregularly shaped structures, such as the PLIC and those in the brainstem, contain more PV voxels with γ_i values less than one. The PV voxels whose γ_i values are below 0.5 have been discarded during model construction. The brain structures have been labelled and the abbreviations explained in Fig. 6.6.

The logistic model not only estimated the average myelination age at every location, but also informed about the rate of myelination through the time interval between key points. The centre of the VLN displayed the slowest maturation rate, and the increase of MLS fraction from $0.25\gamma_i$ to $0.75\gamma_i$ lasted nearly 12 weeks, despite the fact that the same location became myelinated the earliest at approximately 25 weeks GA. The fastest MLS growth with Δt_i values below 4 weeks occurred in the PLIC and medial lemnisci (ML). This finding confirmed the general myelination patterns towards the superior and anterior of the brain.

Fig. 6.8 verified our proposition for the PV locations in Figs. 5.12 and 5.13 in Chapter 5. Non-PV voxels with γ_i values equal to one were commonly found in large and well-defined nuclei, such as the VLN, while small and irregularly shaped structures, such as the PLIC and those in the brainstem, contained more PV voxels with γ_i values less than one. Due to the insufficient resolution of the clinical MR images, the sizes of the small myelination sites were comparable to the voxel sizes, and hence were more prone to the PV effect. Note that we already excluded the PV voxels whose γ_i values were below 0.5 when constructing the growth model, most of which appeared at the boundary of the VLN.

6.4.3 Comparing the growth models constructed using logistic and kernel-based methods

Kuklisova-Murgasova et al. (2011) created a probabilistic neonatal brain atlas using a kernel-based regression method (Davis et al., 2007; Ericsson et al., 2008). Similarly, we constructed MLS growth models in both ROIs between 29 and 44 weeks GA using a Gaussian kernel of one-week standard deviation (SD), which we then visually compared with the spatio-temporal trajectories in Figs. 6.4 and 6.5 obtained from logistic modelling. The differences were mainly found in the thalami close to 39 weeks GA, as illustrated in Fig. 6.9.

The logistic model first showed MLS in the right PLIC tract at approximately 39 weeks GA, and myelination developed in both tracts by the time of 40 weeks GA. In

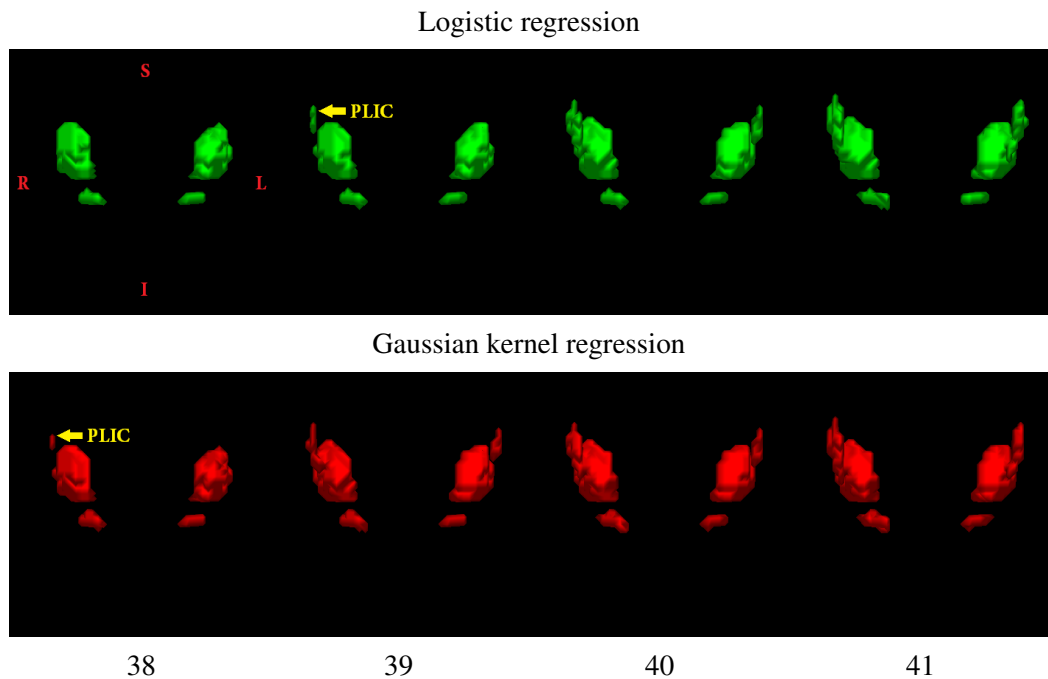


Figure 6.9: Spatio-temporal trajectories (coronal view) for myelin-like signals (MLS) in the thalami between 38 and 41 weeks gestational age (GA), simulated using logistic (green) and kernel (red) regression. The MLS fraction is binarised at 0.5 for presentation in 3D surface rendering. In the logistic model, MLS develops in both tracts of the posterior limbs of the internal capsule (PLIC) at approximately 40 weeks GA, which occurs one week earlier in the kernel regression model.

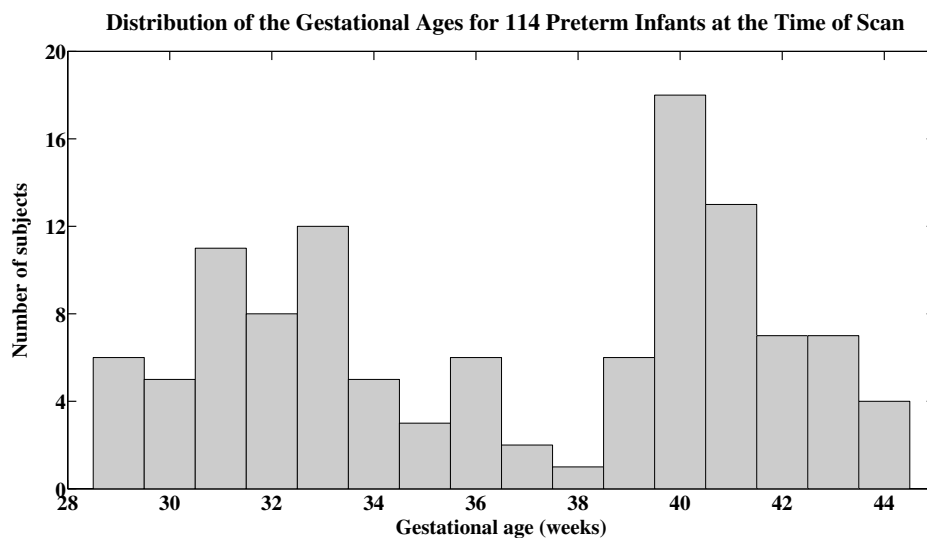


Figure 6.10: Distribution of the gestational ages for 114 preterm infants at the time of scan. Many were scanned at approximately 40 gestational weeks.

the kernel regression model, the first appearance of MLS in the right PLIC tract and myelination of both tracts were one week earlier than in the logistic model.

The inconsistency was due to the fact that the kernel-based approach was affected by the distribution of subjects at different ages. We obtained the non-parametric kernel regression model at GA t by computing the Gaussian weighted average of the transformed MLS segmentations for the 114 preterm infants. Those infants with GAs close to the time point t were assigned higher weights. It can be seen from Fig. 6.10 that only one infant was scanned in the age group of 38 weeks GA, and 6 infants in the group of 39 weeks GA, as compared to 18 in the group of 40 weeks GA. As a result, the appearance of the kernel regression model became falsely more mature at 38 and 39 weeks GA, and the landmark of the myelinated PLIC arrived earlier than in the logistic model. The kernel-based method also encountered this issue at the ends of the age range. In fact, the choice of one-week SD for the Gaussian kernel was a compromise in order to achieve a smooth growth model while eliminating the influence of the non-uniform subject distribution.

6.4.4 Evaluation of the logistic and kernel regression methods through leave-one-out age estimation

We evaluated the logistic and kernel regression methods through leave-one-out age estimation for the 114 preterm infants. The Gaussian kernel of one-week SD used in Section 6.4.3 remained unchanged. We predicted the GA of a test subject by minimising the SSD between the transformed MLS segmentation of the test subject and the growth model constructed based on the remaining 113 subjects. In Figs. 6.11 and 6.12, we presented the age estimate results calculated using the logistic and kernel-based growth models for MLS in the thalami and brainstem respectively.

We calculated the RMSE of the predicted GAs as follows:

$$RMSE = \sqrt{\frac{1}{N} \sum_{n=1}^N (\hat{t}_n - T_n)^2} \quad (6.7)$$

where $\hat{t}_n$ is the predicted GA of subject n , T_n the nominal GA from clinical records, and

Region of interest	Regression model	RMSE (weeks)
Thalami	Logistic	1.409
	Gaussian kernel	1.811
Brainstem	Logistic	2.559
	Gaussian kernel	2.744

Table 6.2: Root mean squared errors (RMSEs) of the predicted gestational ages (GAs) compared to the nominal GAs for 114 preterm infants. The GAs are estimated using the logistic and kernel-based growth models for myelin-like signals (MLS) in the thalami and brainstem respectively. Logistic regression in the thalami achieves the smallest RMSE. This is because most of the MLS spatio-temporal changes between 29 and 44 weeks GA are located in the thalami, and the logistic model is less affected by the variations in the distribution of subjects at different ages.

N the number of subjects. The resultant RMSEs are summarised in Table 6.2.

The logistic model for MLS in the thalami achieved the smallest RMSE in predicting the GAs for the 114 preterm infants. It may be seen from Fig. 6.11(b) that the kernel-based method yielded large estimate errors particularly at the ends of the age range (see the green arrows). As explained in Section 6.4.3, the non-parametric kernel regression model obtained by Gaussian weighted averaging of the MLS segmentations was affected by the variations in the distribution of infants at different ages. The early GAs were under-estimated because the kernel-based model was drawn towards the older infants and appeared more mature than might be expected. The situation was reversed at the late GAs, and caused over-estimation. The same systematic errors incurred close to 39 weeks GA were less evident due to the limited number of infants at these ages.

Table 6.2 also showed that the predictions given by both regression methods were more accurate in the thalami than in the brainstem. The known neurodevelopment (Chi et al., 1977; Counsell et al., 2002) as well as our results in Section 6.4.2 suggested that most of the MLS spatio-temporal changes related to brain maturation between 29 and 44 weeks GA were located in the thalami, which was thus more suitable for age estimation than the brainstem. In addition, both regression models produced outlier predictions in the brainstem, as pointed by the orange arrows in Fig. 6.12(b). These outliers were not seen in the thalami, and were caused by failed MLS segmentation in the presence of substantial

PV voxels. As shown in Fig. 6.8, the small myelination sites in the brainstem contained more PV voxels as a result of their comparable sizes to the image resolution.

Lastly, we tested the sensitivity of the age estimate results to the selection of the reference subject. We constructed the logistic and kernel-based MLS growth models in the thalami and brainstem in the space of another reference subject at 36 weeks GA, and calculated the GAs for the 114 preterm infants using the same leave-one-out procedure. The new RMSEs in Table 6.3 were close to those in Table 6.2 obtained using the previous reference subject. Pairwise Student’s t-tests at the 5% significance level showed no significant differences in the GAs predicted using the two reference subjects, regardless of the ROI and the choice of the regression model.

Region of interest	Regression model	RMSE (weeks)
Thalami	Logistic	1.508
	Gaussian kernel	1.846
Brainstem	Logistic	2.552
	Gaussian kernel	2.857

Table 6.3: Root mean squared errors (RMSEs) of the gestational ages (GAs) predicted for 114 preterm infants using a different reference subject at 36 weeks GA. The new RMSEs are close to those in Table 6.2 obtained using the previous reference subject. Pairwise Student’s t-tests at the 5% significance level show no significant differences in the GAs predicted using the two reference subjects.

6.5 Conclusion

In this chapter, we developed a method of constructing a spatio-temporal model for MLS in a predefined ROI using voxelwise logistic regression based on the automatic segmentations obtained from T₂w images. The model was then applied to estimate the GAs of 114 preterm infants in a leave-one-out procedure. We investigated the differences in the myelination process between the thalami and brainstem from 29 to 44 weeks GA. Additionally, we compared the logistic function with a non-parametric kernel-based model (Kuklisova-Murgasova et al., 2011), which involved voxelwise Gaussian weighted averaging of individual MLS segmentations.

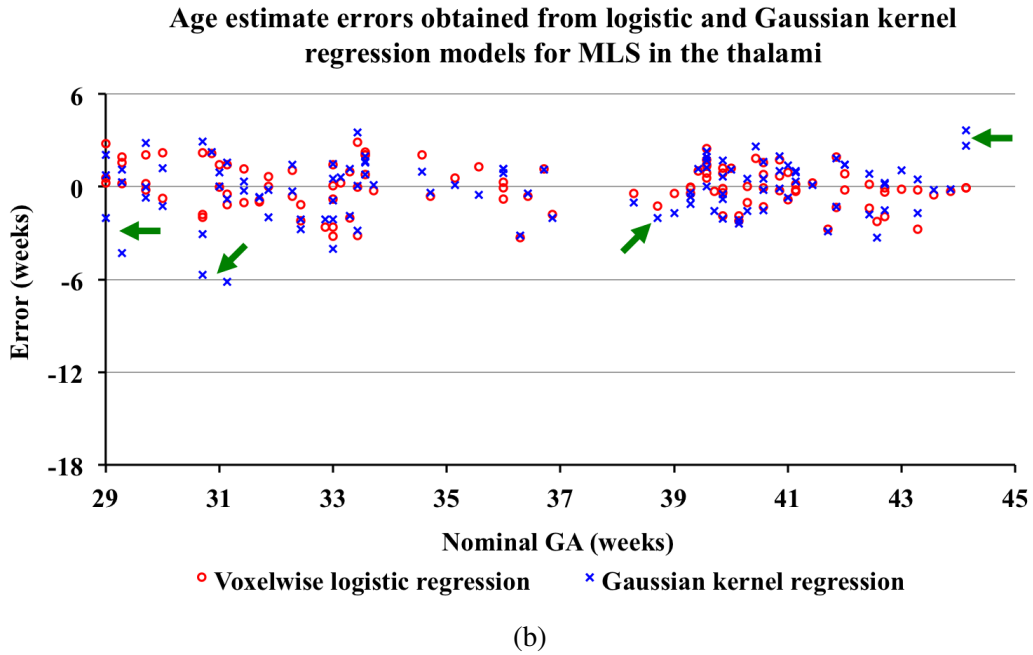
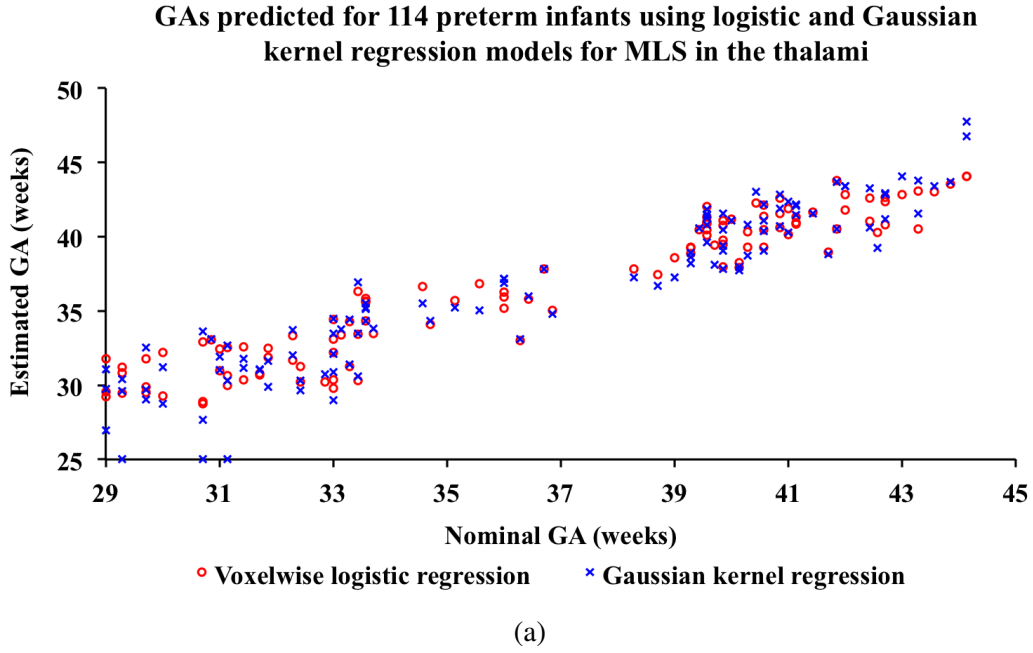
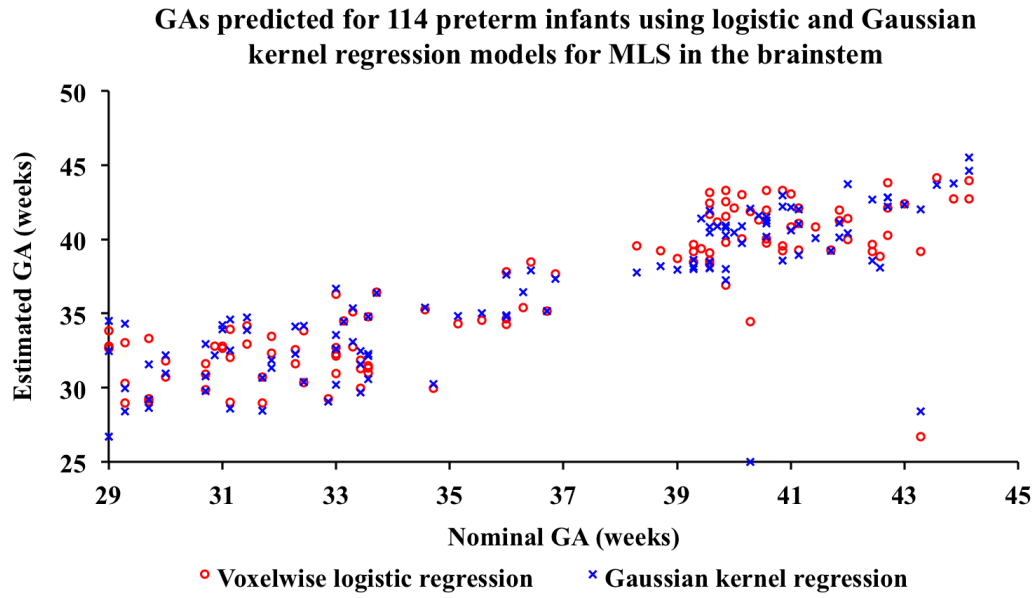
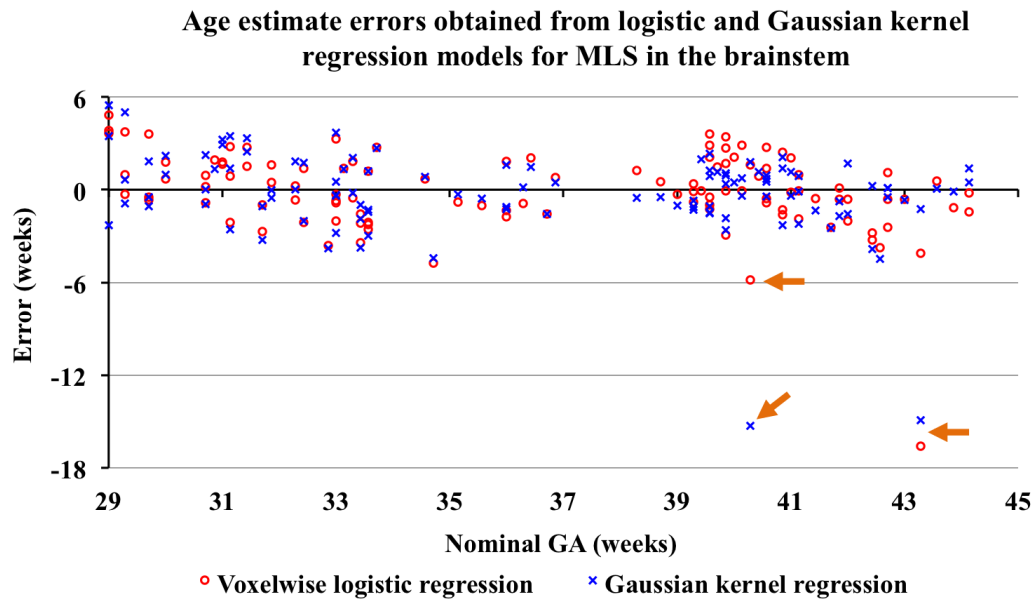


Figure 6.11: (a) Gestational ages (GAs) predicted for 114 preterm infants and (b) age estimate errors compared to the nominal GAs using the logistic and kernel-based growth models for myelin-like signals in the thalami. The kernel-based method yields large estimate errors at the ends of the age range, as pointed by the green arrows, due to the non-uniform distribution of infants at different ages. The same systematic errors incurred close to 39 weeks GA are less evident due to the limited number of infants at these ages.



(a)



(b)

Figure 6.12: (a) Gestational ages (GAs) predicted for 114 preterm infants and (b) age estimate errors compared to the nominal GAs using the logistic and kernel-based growth models for myelin-like signals (MLS) in the brainstem. The predictions given by both regression methods display larger deviations from the nominal GAs compared to Fig. 6.11. This is because most of the MLS spatio-temporal changes between 29 and 44 weeks GA are located in the thalami, whereas the brainstem provides relatively little information on brain maturation during this particular postnatal period. The outlier predictions pointed by the orange arrows are caused by failed MLS segmentation in the presence of substantial partial volume voxels.

We found that the brainstem matured earlier than the thalami, and showed very limited age-dependent MLS progression from 29 weeks GA onward. Consequently, the GAs of the 114 preterm infants predicted based on MLS in the brainstem yielded a larger RMSE compared to the thalami, irrespective of the modelling method used. The differences between the logistic and kernel-based regression models were best seen in the thalami. The logistic model indicated that MLS arrived in the landmark site of PLIC (Bird et al., 1989) at approximately 40 weeks GA in great consistency with the clinical MRI observations (Counsell et al., 2002; Rutherford, 2002), whereas myelination of the PLIC tracts occurred one week earlier in the kernel-based model. This is because the weighted averaging using a Gaussian kernel of fixed width was affected by the variations in the distribution of infants at different ages, which hence lowered the accuracies of the predicted GAs.

It is interesting that the spatio-temporal analysis results are consistent with the known myelination patterns, but our method still has some limitations. Because the T_2w sequence is not designed specifically for imaging myelin (Korogi et al., 1996), we classified each voxel automatically as one of the MLS, PV and background (BKG) classes using the segmentation method proposed in Chapter 5, and calculated the voxelwise fraction of the MLS class for construction of the growth model. The additional segmentation step introduced errors, as compared to Dean III et al. (2014) in which myelin development is modelled from the myelin water fraction (MWF) maps measured directly using myelin water imaging (MWI) (Alonso-Ortiz et al., 2015). However, it must be taken into account that MWI is clinically impractical when imaging fragile preterm neonates (Rutherford, 2002; Dubois et al., 2014). We instead presented an alternative approach that is easily applicable in clinical practice, and that shows great promise in age estimation and brain maturation assessment. Comparisons of the logistic model with adaptive kernel regression (Serag et al., 2012) and other parametric functions, such as Gompertz (Gompertz, 1825) and LogOdds (Pohl et al., 2006b), remains to be investigated in the future.

Chapter 7

Conclusion

This thesis has been devoted to the development of an imaging marker based on myelination visualised using the conventional T₂w MRI for automatic and prognostic evaluation of neonatal brain maturation. In this final chapter, we summarise the contributions, discuss the limitations, and outline the directions for future research.

7.1 Summary of contributions

The contributions of this thesis are threefold:

- 1) **Development of an atlas-based neonatal brain MRI segmentation method in the NeoBrainS12 challenge.** We adapted the SPM segmentation tool (Ashburner and Friston, 2005) to a dedicated neonatal brain atlas (Kuklisova-Murgasova et al., 2011). Partial volume (PV) misclassifications were corrected by connected component labelling. Myelinated white matter (MWM), which is not included in the atlas, was segmented by fitting a multivariate Gaussian mixture model (GMM) to both T₁w and T₂w intensities of the aligned images in each of the two regions: i) basal ganglia and thalami (BGT); ii) brainstem.

The proposed method achieved Dice coefficients (DCs) above 0.7 with respect to the manual annotations for all the tissue types except MWM. In fact, MWM segmentation was significantly underperformed by all the algorithms in the challenge due to the

absence of atlas spatial prior (Isgum et al., 2015). The purely intensity-based GMM in our approach considerably over-estimated MWM despite the incorporation of multi-spectral MRI data, which we then improved in the next contribution.

- 2) **Development of an EM framework with explicit PV modelling via second-order MRFs for segmenting myelin-like signals (MLS) on T₂w neonatal brain MR images.** We introduced the term MLS because T₂w signals reflect several maturational processes simultaneously including myelination (Korogi et al., 1996). We defined an MLS segmentation protocol for delineating manual annotations in two regions of interest (ROIs): the thalami and brainstem. The ROIs were obtained using the previously proposed atlas-based method, but the subsequent MLS segmentation in the each ROI was atlas-free. To tackle the absence of an MLS atlas and the PV effect, we fitted three Gaussians to model the intensities of MLS, PV voxels and the background (BKG) tissue, respectively. The spatial priors were generated entirely based on tissue connectivities among the triplet of classes in the local neighbourhood.

This approach performed robustly and accurately with respect to the manual annotations in both ROIs for 16 test subjects aged at one-week intervals between 29 and 44 weeks GA, regardless of the volume fraction used to obtain the initial MLS segmentation estimates. The independence of MLS segmentations from the initial volume fraction is crucial for the construction of an unbiased spatio-temporal model in the next contribution.

- 3) **Construction of separate MLS growth models in the thalami and brainstem of the preterm brain from 29 to 44 weeks GA, and age estimation using the growth model.** We non-rigidly registered (Rueckert et al., 1999) T₂w images of 114 subjects with a single reference subject at 36 weeks GA. The MLS segmentations in each ROI were transformed accordingly to the common reference space where we fitted a voxel-wise logistic function to construct the spatio-temporal model. We estimated the GA of a preterm infant by minimising the deviation between its MLS segmentation and the normal growth model.

The simulated MLS trajectories in the thalami and brainstem were consistent with the known myelination patterns, particularly the arrival of MLS in the posterior limbs of the internal capsule (PLIC) at approximately 40 weeks GA (Counsell et al., 2002). We demonstrated that the logistic model predicted the GAs of the 114 subjects in a leave-one-out fashion with a smaller root mean squared error, as compared to a commonly used kernel regression method (Davis et al., 2007; Kuklisova-Murgasova et al., 2011) which is affected by the uneven distribution of subjects at different ages.

7.2 Limitations and future work

7.2.1 Short-term goals

Below we describe the short-term goals that would address the limitations of the proposed methods.

- 1) **Multivariate MLS segmentation.** In Chapter 4, we segmented MWM in the Neo-BrainS12 challenge by fitting a two-class GMM to the T_1w and T_2w intensities in each ROI of the aligned images. However, this approach neither considered the PV effect, nor did it provide any spatial regularisation. We thus took a step back in Chapter 5, and used T_2w images only to segment MLS with explicit PV modelling via second-order MRFs, which can be naturally extended to handle multi-spectral MRI data. As discussed at the end of Chapter 5, myelinated grey matter (GM) nuclei are more conspicuous on T_2w image, whereas myelinated white matter (WM) tracts are better visualised in the T_1w contrast (Barkovich, 1998). Using both T_1w and T_2w images might improve the specificity of the conventional MRI techniques in terms of indicating myelination.
- 2) **Assessing MLS in other brain regions.** The current analysis focused on the thalami and brainstem because these are the regions where the majority of myelination occurs during the perinatal period. However, myelination is also present in the cerebellum from approximately 25 weeks GA, and progresses to the WM in the cerebral

hemispheres in late infancy (Counsell et al., 2002). Each ROI has a different predictive power that best assesses a particular period of brain development. For example, we demonstrated in Chapter 6 that the thalami with the most prominent MLS growth between 29 and 44 weeks GA produced the most accurate age estimates for preterm infants in this age range. It is therefore desirable to include the cerebellum and cerebral hemispheres in the analysis, while taking into consideration the challenges of MLS segmentation in these regions. For example, PV voxels in the cerebellum might contain all three of the GM, WM and cerebrospinal fluid (CSF), and the cerebral hemispheres frequently show substantial WM intensity inhomogeneities in preterm infants.

- 3) **Multiple PV Gaussians.** In Fig. 1.3 in Chapter 1, we illustrated that the intensity distribution of MLS is substantially mixed up with that of BKG in the histogram due to the intermediate PV intensities. We introduced a single spatially constrained Gaussian to approximate the tail region between the MLS and BKG intensity distributions. Adding more PV Gaussians could improve the fit to the flat tail region, provided that there are valid spatial priors.
- 4) **Spatial co-alignment.** In Chapter 6, we non-rigidly registered the T_2w images into a common reference space using free-form deformations (FFD) (Rueckert et al., 1999). However, spatial co-alignment and hence anatomical correspondence can be disrupted due to the contrast change over time caused by the continuous myelination and decreasing water content. Therefore, we had chosen large control point spacing to restrict the FFD transformations to only the global shape of the central brain structures, so that local variations in the tissue contrast do not affect the registration results. It will be interesting to develop a parametric intensity model to characterise the temporal contrast change, which would allow for finer FFD resolutions.
- 5) **Evaluation of the logistic MLS growth model.** Currently, we evaluated the logistic regression model as compared to fixed-width kernel smoothing by visually inspecting the appearance of MLS in the landmark site of PLIC (Bird et al., 1989), and by leave-one-out age estimation for 114 preterm infants. Comparisons with adaptive kernel

regression (Serag et al., 2012) that addresses uneven sampling distributions as well as other parametric functions, such as Gompertz (Gompertz, 1825) and LogOdds (Pohl et al., 2006b), remains to be investigated in the future.

7.2.2 Long-term outlooks

It is widely believed that, among the premature infants without brain lesions, some might develop long-term impairments while the majority generally “catch up” with other children born full term (Rorke and Riggs, 1969). This thesis has been focusing on the detection of early signs of neurodevelopmental delays. It would also be of great interest to investigate the “catching up” cohort. For instance, do the preterm infants indeed reach the same brain maturation level as full term babies at some later time? If they do, when does this “catching up” occurs, and are there any regional differences? Answers to these questions are yet not available. Investigations on the “catching up” process might provide some insights into the effects of preterm birth.

Also, the MLS-based image analysis tools developed in this thesis could be applied potentially to fetal brain MRI for assessment of prenatal brain development. The information on myelination extracted from clinical MR images might be fused with cranial ultrasonography which is unable to depict myelination (Van Wezel-Meijler et al., 2010).

7.3 Conclusion

Most of the babies born prematurely adjust well during growth, but some experience learning difficulties at school age. In clinical practice, birth weights are used primarily as a measure of brain prematurity for early screening, which sometimes cannot fully reflect the neurodevelopmental conditions. Routinely acquired T_2w brain MRI of preterm neonates carry signatures that are thought to be associated with an important maturational process known as myelination. In this thesis, we therefore developed an image analysis pipeline that achieves automatic segmentation and spatio-temporal modelling of the characteristic myelin-like signals (MLS) using T_2w neonatal brain MRI.

The proposed segmentation method via second-order MRFs is robust against PV corruption, different initialisations and variations in the size and shape of the ROI. The average MLS growth model enables us to re-estimate the GAs of preterm infants according to the underlying brain maturation status. It is hoped that patients with abnormal brain development compared to the majority of the preterm population could be identified at an early stage to prevent long-term disabilities.

List of Publications

- Isgum, I., M.J.N.L. Benders, B. Avants, M.J. Cardoso, S.J. Counsell, E.F. Gomez, L. Gui, P.S. Huppi, K.J. Kersbergen, A. Makropoulos, A. Melbourne, P. Moeskops, C.P. Mol, M. Kuklisova-Murgasova, D. Rueckert, J.A. Schnabel, V. Srhoj-Egekher, J. Wu, S. Wang, L.S. De Vries, and M.A. Miergever (2015). “Evaluation of automatic neonatal brain segmentation algorithms: The NeoBrainS12 challenge”. In: *Med. Image Anal.* 20.1, pp. 135–151.
- Wang, S., M. Kuklisova-Murgasova, C. Ledig, J.V. Hajnal, S.J. Counsell, and J.A. Schnabel. “Segmentation with partial volume modeling via second-order MRFs: Application to tracking neonatal brain maturation”. In submission.
- Wang, S., M. Kuklisova-Murgasova, and J.A. Schnabel (2012). “An atlas-based method for neonatal MR brain tissue segmentation”. In: *Proceedings of the MICCAI Grand Challenge: Neonatal Brain Segmentation*, pp. 28–35.
- Wang, S., M. Kuklisova-Murgasova, J.V. Hajnal, C. Ledig, and J.A. Schnabel (2016). “Regression analysis for assessment of myelination status in preterm brains with magnetic resonance imaging”. In: *Proceedings of the 13th International Symposium on Biomedical Imaging (ISBI)*, pp. 278–281.

Bibliography

- Adams, R. and L. Bischof (1994). “Seeded region growing”. In: *IEEE Trans. Pattern Anal. Mach. Intell.* 16.6, pp. 641–647.
- Aisenson, M.R. (1950). “Closing of the anterior fontanelle”. In: *Pediatr.* 6.2, pp. 223–226.
- Alexa, M. (2002). “Linear combination of transformations”. In: *ACM Trans. Graph.* 21.3, pp. 380–387.
- Aljabar, P., K.K. Bhatia, M. Kuklisova-Murgasova, J.V. Hajnal, J.P. Boardman, L. Srinivasan, M.A. Rutherford, L.E. Dyet, A.D. Edwards, and D. Rueckert (2008). “Assessment of brain growth in early childhood using deformation-based morphometry”. In: *NeuroImage* 39, pp. 348–358.
- Aljabar, P., R.A. Heckemann, J.V. Hajnal, and D. Rueckert (2009). “Multi-atlas based segmentation of brain images: Atlas selection and its effect on accuracy”. In: *NeuroImage* 46.726-738.
- Alonso-Ortiz, E., I.R. Levesque, and G.B. Pike (2015). “MRI-based myelin water imaging: A technical review”. In: *Magn. Reson. Med.* 73, pp. 70–81.
- Anbeek, P., L.L. Vincken, F. Groenendaal, A. Koeman, M.J.P. Van Osch, and J. Van der Grond (2008). “Probabilistic brain tissue segmentation in neonatal magnetic resonance imaging”. In: *Pediatr. Res.* 63.2, pp. 158–163.
- Arsigny, V., P. Fillard, X. Pennec, and N. Ayache (2007). “Geometric means in a novel vector space structure on symmetric positive-definite matrices”. In: *SIAM J. Matrix Anal. Appl.* 29.1, pp. 328–347.
- Arteachevarria, X., A. Munoz-Barrutia, and C. Ortiz-de Solorzano (2009). “Combination strategies in multi-atlas image segmentation: Application to brain MR data”. In: *IEEE Trans. Med. Imag.* 28.8, pp. 1266–1277.
- Ashburner, J. and K.J. Friston (2005). “Unified segmentation”. In: *NeuroImage* 26, pp. 839–851.

- Ashburner, J., C. Hutton, R.S.J. Frackowiak, I. Johnsrude, C. Price, and K.J. Friston (1998). "Identifying global anatomical differences: Deformation-based morphometry". In: *Hum. Brain Mapp.* 6, pp. 348–357.
- Avants, B. and J.C. Gee (2004). "Geodesic estimation for large deformation anatomical shape averaging and interpolation". In: *NeuroImage* 23, S139–S150.
- Avants, B.B., C.L. Epstein, M. Grossman, and J.C. Gee (2008). "Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain". In: *Med. Image Anal.* 12.1, pp. 26–41.
- Ball, G., J.P. Boardman, D. Rueckert, P. Aljabar, T. Arichi, N. Merchant, I.S. Gousias, A.D. Edwards, and S.J. Counsell (2012). "The effect of preterm birth on thalamic and cortical development". In: *Cereb. cortex* 22, pp. 1016–1024.
- Barkovich, A.J. (1998). "MR of the normal neonatal brain: Assessment of deep structures". In: *Am. J. Neuroradiol.* 19, pp. 1397–1403.
- Basser, P.J., J. Mattiello, and D. Le Bihan (1994). "Estimation of the effective self-diffusion tensor from the NMR spin echo". In: *J. Magn. Reson.* 103, pp. 247–254.
- Besag, J. (1974). "Spatial interaction and the statistical analysis of lattice systems". In: *J. R. Stat. Soc. Ser. B Stat. Methodol.* 36.2, pp. 192–236.
- Besag, J. (1986). "On the statistical analysis of dirty pictures". In: *J. R. Stat. Soc. Ser. B Stat. Methodol.* 48.3, pp. 259–302.
- Bhatia, K.K., J.V. Hajnal, B.K. Puri, A.D. Edwards, and D. Rueckert (2004). "Consistent groupwise non-rigid registration for atlas construction". In: *Proceedings of the 2nd International Symposium on Biomedical Imaging (ISBI)*, pp. 908–911.
- Bird, C.R., M. Hedberg, B.P. Drayer, P.J. Keller, R.A. Flom, and J.A. Hodak (1989). "MR assessment of myelination in infants and children: Usefulness of marker sites". In: *Neuroradiology* 10, pp. 731–740.
- Bleasdale, J. and J. Nelder (1960). "Plant population and crop yield". In: *Nature* 188, p. 342.
- Blencowe, H., S. Cousens, M.Z. Oestergaard, D. Chou, A. Moller, R. Narwal, A. Adler, C.V. Garcia, S. Rohde, L. Say, and J.E. Lawn (2012). "National, regional and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications". In: *Lancet* 379, pp. 2162–2172.

- Boardman, J.P., S.J. Counsell, D. Rueckert, O. Kapellou, K.K. Bhatia, P. Aljabar, J.V. Hajnal, J.M. Allsop, M.A. Rutherford, and A.D. Edwards (2006). “Abnormal deep grey matter development following preterm birth detected using deformation-based morphometry”. In: *NeuroImage* 32, pp. 70–78.
- Broit, C. (1981). “Optimal registration of deformed images”. PhD thesis. University of Pennsylvania.
- Brown, M.A. and R.C. Semelka (2003). *MRI: basic principles and applications*. Hoboken, New Jersey: Wiley.
- Buades, A., B. Coll, and J.M. Morel (2005). “A review of image denoising algorithms, with a new one”. In: *Multiscale Model. Simulat.* 4.2, pp. 490–530.
- Cardoso, M.J., M.J. Clarkson, G.R. Ridgway, M. Modat, N.C. Fox, S. Ourselin, and the Alzheimer’s Disease Neuroimaging Initiative (2011). “LoAd: A locally adaptive cortical segmentation algorithm”. In: *NeuroImage* 56, pp. 1386–1397.
- Chan, T.F., B.Y. Sandberg, and L.A. Vese (2000). “Active contours without edges for vector-valued images”. In: *J. Vis. Commun. Image Represent.* 11.2, pp. 130–141.
- Chi, J.G., E.C. Dooling, and F.H. Gilles (1977). “Gyral development of the human brain”. In: *Ann. Neurol.* 1, pp. 86–93.
- Christensen, G.E., R.D. Rabbitt, and M.I. Miller (1996). “Deformable templates using large deformation kinematics”. In: *IEEE Trans. Image Process.* 5.10, pp. 1435–1447.
- Cocosco, C.A., A.P. Zijdenbos, and A.C. Evans (2003). “A fully automatic and robust brain MRI tissue classification method”. In: *Med. Image Anal.* 7, pp. 513–527.
- Counsell, S.J., E.F. Maalouf, A.M. Fletcher, P. Duggan, M. Battin, H.J. Lewis, A.H. Herlihy, A.D. Edwards, G.M. Bydder, and M.A. Rutherford (2002). “MR imaging assessment of myelination in the very preterm brain”. In: *Am. J. Neuroradiol.* 23, pp. 872–881.
- Cover, T.M. and P.E. Hart (1967). “Nearest neighbor pattern classification”. In: *IEEE Trans. Inf. Theory* 13.1, pp. 21–27.
- Davis, B.C., P.T. Fletcher, E. Bullitt, and S. Joshi (2007). “Population shape regression from random design data”. In: *Proceedings of the 11th International Conference on Computer Vision (ICCV)*, pp. 1–7.

- De Vries, L.S., K.J. Rademaker, F. Groenendaal, P. Eken, I.C. Van Haastert, W.P. Vandertop, R. Gooskens, and L.C. Meiners (1998). “Correlation between neonatal cranial ultrasound, MRI in infancy and neurodevelopmental outcome in infants with a large intraventricular haemorrhage with or without unilateral parenchymal involvement”. In: *Neuropediatrics* 29, pp. 180–188.
- Dean III, D.C., J. O’Muircheartaigh, H. Dirks, N. Waskiewicz, K. Lehman, Walker L., M. Han, and S.C.L. Deoni (2014). “Modeling healthy male white matter and myelin development: 3 through 60 months of age”. In: *NeuroImage* 84, pp. 742–752.
- Dempster, A.P., N.M. Laird, and D.B. Rubin (1977). “Maximum likelihood from incomplete data via the EM algorithm”. In: *J. R. Stat. Soc. Ser. B Stat. Methodol.* 39.1, pp. 1–38.
- Deoni, S.C.L., D.C. Dean III, J. Remer, H. Dirks, and J. O’Muircheartaigh (2015). “Cortical maturation and myelination in healthy toddlers and young children”. In: *NeuroImage* 115, pp. 147–161.
- Dice, L.R. (1945). “Measures of the amount of ecologic association between species”. In: *Ecology* 23.3, pp. 297–302.
- Dittrich, E., T.R. Raviv, G. Kasprian, R. Donner, P.C. Brugger, D. Prayer, and G. Langs (2014). “A spatio-temporal latent atlas for semi-supervised learning of fetal brain segmentations and morphological age estimation”. In: *NeuroImage* 18, pp. 9–21.
- Dubois, J., G. Dehaene-Lambertz, S. Kulikova, C. Poupon, P.S. Huppi, and L. Hertz-Pannier (2014). “The early development of brain white matter: A review of imaging studies in fetuses, newborns and infants”. In: *Neuroscience* 276, pp. 48–71.
- Dunn, J.C. (1974). “A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters”. In: *Journal of Cybernetics* 3, pp. 32–57.
- Dupuis, P., Grenander U., and M.I. Miller (1998). “Variational problems on flows of diffeomorphisms for image matching”. In: *Q. Appl. Math.* 56.3, pp. 587–600.
- Durrleman, S., T. Fletcher, G. Gerig, M. Niethammer, and Pennec X., eds. (2014). “Spatio-temporal image analysis for longitudinal and time-series image data”. In: *Lecture Notes in Computer Science*. Vol. 8682. Springer.
- Ericsson, A., P. Aljabar, and D. Rueckert (2008). “Construction of a patient-specific atlas of the brain: Application to normal aging”. In: *Proceedings of the 5th International Symposium on Biomedical Imaging (ISBI)*, pp. 480–483.

- Evans, A.C., A.L. Janke, D.L. Collins, and S. Baillet (2012). “Brain templates and atlases”. In: *NeuroImage* 62, pp. 911–922.
- Fitzpatrick, J.M., D.L.G. Hill, and Maurer, Jr. C.R. (2000). “Image Registration”. In: *Handbook of Medical Imaging, Volume 2. Medical Image Processing and Analysis*. Ed. by M. Sonka and J.M. Fitzpatrick. SPIE Press, pp. 447–513.
- Fonov, V., A.C. Evans, K. Botteron, C.R. Almli, R.C. McKinstry, D.L. Collins, and the Brain Development Cooperative Group (2011). “Unbiased average age-appropriate atlases for pediatric studies”. In: *NeuroImage* 54, pp. 313–327.
- Fornari, E., M.G. Knyazeva, R. Meuli, and P. Maeder (2007). “Myelination shapes functional activity in the developing brain”. In: *NeuroImage* 38, pp. 511–518.
- Gompertz, B. (1825). “On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies”. In: *Philos. Trans. R. Soc. Lond.* 115, pp. 513–583.
- Gousias, I.S., A.D. Edwards, M.A. Rutherford, S.J. Counsell, J.V. Hajnal, D. Rueckert, and A. Hammers (2012). “Magnetic resonance imaging of the newborn brain: Manual segmentation of labelled atlases in term-born and preterm infants”. In: *NeuroImage* 62, pp. 1499–1509.
- Gui, L., R. Lisowski, T. Faundez, P.S. Huppi, F. Lazeyras, and M. Kocher (2012). “Morphology-driven automatic segmentation of MR images of the neonatal brain”. In: *Med. Image Anal.* 16.8, pp. 1565–1579.
- Guillemaud, R. and M. Brady (1997). “Estimating the bias field of MR images”. In: *IEEE Trans. Med. Imag.* 16.3, pp. 238–251.
- Guimond, A., J. Meunier, and J.P. Thirion (2000). “Average brain models: A convergence study”. In: *Comput. Vis. Image Underst.* 77, pp. 192–210.
- Guimond, A., A. Roche, N. Ayache, and J. Meunier (2001). “Three-dimensional multi-modal brain warping using the demons algorithm and adaptive intensity corrections”. In: *IEEE Trans. Med. Imag.* 20.1, pp. 58–69.
- Habas, P.A., K. Kim, J.M. Corbett-Detig, F. Rousseau, O.A. Glenn, A.J. Barkovich, and C. Studholme (2010). “A spatiotemporal atlas of MR intensity, tissue probability and shape of the fetal brain with application to segmentation”. In: *NeuroImage* 53, pp. 460–470.

- Heckemann, R.A., J.V. Hajnal, P. Aljabar, D. Rueckert, and A. Hammers (2006). “Automatic anatomical brain MRI segmentation combining label propagation and decision fusion”. In: *NeuroImage* 33, pp. 115–126.
- Henkelman, R.M., G.J. Stanisz, and S.J. Graham (2001). “Magnetization transfer in MRI: A review”. In: *NMR Biomed.* 14, pp. 57–64.
- Huppi, P.S. and T.E. Inder (2001). “Magnetic resonance techniques in the evaluation of the perinatal brain: recent advances and future directions”. In: *Semin. Neonatol.* 6, pp. 195–210.
- Huttenlocher, D.P., G.A. Klanderman, and W.J. Rucklidge (1993). “Comparing images using the Hausdorff distance”. In: *IEEE Trans. Med. Imag.* 15.9, pp. 850–863.
- Hutton, T.J., B.F. Buxton, P. Hammond, and W.W. Potts (2003). “Estimating average growth trajectories in shape-space using kernel smoothing”. In: *IEEE Trans. Med. Imag.* 22.6, pp. 747–753.
- Isgum, I., M.J.N.L. Benders, B. Avants, M.J. Cardoso, S.J. Counsell, E.F. Gomez, L. Gui, P.S. Huppi, K.J. Kersbergen, A. Makropoulos, A. Melbourne, P. Moeskops, C.P. Mol, M. Kuklisova-Murgasova, D. Rueckert, J.A. Schnabel, V. Srhoj-Egekher, J. Wu, S. Wang, L.S. De Vries, and M.A. Miergever (2015). “Evaluation of automatic neonatal brain segmentation algorithms: The NeoBrainS12 challenge”. In: *Med. Image Anal.* 20.1, pp. 135–151.
- Jenkinson, M., C.F. Beckmann, T.E.J. Behrens, M.W. Woolrich, and S.M. Smith (2012). “FSL”. In: *NeuroImage* 62, pp. 782–790.
- Kinney, H.C., B.A. Brody, A.S. Kloman, and F.H. Gilles (1988). “Sequence of central nervous system myelination in human infancy. II. Patterns of myelination in autopsied infants”. In: *J. Neuropathol. Exp. Neurol.* 47.3, pp. 217–234.
- Kittler, J., M. Hatef, R.P.W. Duin, and J. Matas (1998). “On combining classifiers”. In: *IEEE Trans. Pattern Anal. Mach. Intell.* 20.3, pp. 226–239.
- Klein, S., M. Staring, and J.P.W. Pluim (2007). “Evaluation of optimization methods for nonrigid medical image registration using mutual information and B-splines”. In: *IEEE Trans. Med. Imag.* 16.12, pp. 2879–2890.
- Kloppel, S., C.M. Stonnington, J. Barnes, F. Chen, C. Chu, C.D. Good, I. Mader, L.A. Mitchell, A.C. Patel, C.C. Roberts, N.C. Fox, C.R. Jack, J. Ashburner, and R.S.J.

- Frackowiak (2008). “Accuracy of dementia diagnosis-a direct comparison between radiologists and a computerized method”. In: *Brain* 131, pp. 2969–2974.
- Korogi, Y., M. Takahashi, M. Sumi, T. Hirai, Y. Sakamoto, I. Ikushima, and H. Miyayama (1996). “MR signal intensity of the perirolandic cortex in the neonate and infant”. In: *Neuroradiology* 38, pp. 578–584.
- Kuklisova-Murgasova, M., L. Dyet, A.D. Edwards, M.A. Rutherford, J.V. Hajnal, and D. Rueckert (2007). “Segmentation of brain MRI in young children”. In: *Acad. Radiol.* 14, pp. 1350–1366.
- Kuklisova-Murgasova, M., V. Doria, L. Srinivasan, P. Aljabar, A.D. Edwards, and D. Rueckert (2009). “A spatio-temporal atlas of the growing brain for fMRI studies”. In: *Proceedings of the MICCAI Workshop: Image analysis of the developing brain*.
- Kuklisova-Murgasova, M., P. Aljabar, L. Srinivasan, S.J. Counsell, V. Doria, A. Serag, I.S. Gousias, J.P. Boardman, M.A. Rutherford, A.D. Edwards, J.V. Hajnal, and D. Rueckert (2011). “A dynamic 4D probabilistic atlas of the developing brain”. In: *NeuroImage* 54, pp. 2750–2763.
- Laule, C., I.M. Vavasour, S.H. Kolind, D.K.B. Li, T.L. Traboulsee, G.R.W. Moore, and A.L. MacKay (2007). “Magnetic resonance imaging of myelin”. In: *Neurotherapeutics* 4, pp. 460–484.
- Le Bihan, D., J. Mangin, C. Poupon, C.A. Clark, S. Pappata, N. Molko, and H. Chabriet (2001). “Diffusion tensor imaging: Concepts and applications”. In: *J. Magn. Reson. Imaging* 13, pp. 534–546.
- Ledig, C., R. Wright, A. Serag, P. Aljabar, and D. Rueckert (2012). “Neonatal brain segmentation using second order neighborhood information”. In: *Proceedings of the MICCAI Workshop: Perinatal and Paediatric Imaging*, pp. 33–40.
- Lee, S., G. Wolberg, and S.Y. Shin (1997). “Scattered data interpolation with multilevel B-splines”. In: *IEEE Trans. Visualization Comput. Graph.* 3.3, pp. 228–244.
- Levenberg, K. (1944). “A method for the solution of certain problems in least-squares”. In: *Q. Appl. Math.* 2, pp. 164–168.
- Li, S.Z. (1995). *Markov Random Field Modeling in Computer Vision*. Berlin: Springer-Verlag.

- Liew, A.W.C. and H. Yan (2006). “Current methods in the automatic tissue segmentation of 3D magnetic resonance brain images”. In: *Curr. Med. Imaging Rev.* 2.1, pp. 91–103.
- Liu, M., A. Kitsch, S. Miller, V. Chau, K. Poskitt, F. Rousseau, D. Shaw, and C. Studholme (2016). “Patch-based augmentation of expectation-maximization for brain MRI tissue segmentation at arbitrary age after premature birth”. In: *NeuroImage* 127, pp. 387–408.
- Lotufo, R. and A. Falcao (2000). “The ordered queue and the optimality of the watershed approaches”. In: *Computational Imaging and Vision*. Ed. by J. Goutsias, L. Vincent, and D.S. Bloomberg. Vol. 18. Springer. Chap. Mathematical Morphology and its Applications to Image and Signal Processing, pp. 341–350.
- MacQueen, J.B. (1967). “Some methods for classification and analysis of multivariate observations”. In: *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*. Vol. 1. University of California Press, pp. 281–297.
- Maes, F., A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens (1997). “Multimodality image registration by maximization of mutual information”. In: *IEEE Trans. Med. Imag.* 16.2, pp. 187–198.
- Makropoulos, A., I.S. Gousias, C. Ledig, P. Aljabar, A. Serag, J.V. Hajnal, A.D. Edwards, S.J. Counsell, and D. Rueckert (2014). “Automatic whole brain MRI segmentation of the developing neonatal brain”. In: *IEEE Trans. Med. Imag.* 33.9, pp. 1818–1831.
- Makropoulos, A., P. Aljabar, R. Wright, B. Huning, N. Merchant, T. Arichi, N. Tusor, J.V. Hajnal, A.D. Edwards, S.J. Counsell, and D. Rueckert (2016). “Regional growth and atlasing of the developing human brain”. In: *NeuroImage* 125, pp. 456–478.
- Manjon, J.V., J. Tohka, and M. Robles (2010). “Improved estimates of partial volume coefficients from noisy brain MRI using spatial context”. In: *NeuroImage* 53, pp. 480–490.
- MATLAB (R2013b). *Version 8.2.0.701*. Natick, Massachusetts: The MathWorks Inc.
- Mazziotta, J., A. Toga, A. Evans, P. Fox, J. Lancaster, K. Zilles, R. Woods, T. Paus, G. Simpson, B. Pike, C. Holmes, L. Collins, P. Thompson, D. MacDonald, M. Iacoboni, T. Schormann, K. Amunts, N. Palomero-Gallagher, S. Geyer, L. Parsons, K. Narr, N. Kabani, G. Le Goualher, D. Boomsma, T. Cannon, R. Kawashima, and B. Mazoyer (2001). “A probabilistic atlas and reference system for the human brain: International

- Consortium for Brain Mapping (ICBM)". In: *Phil. Trans. R. Soc. B* 356, pp. 1293–1322.
- Melbourne, A., M.J. Cardoso, G.S. Kendall, N.J. Robertson, N. Marlow, and S. Ourselin (2012). "NeobrainS12 challenge: Adaptive neonatal MRI brain segmentation with myelinated white matter class and automated extraction of ventricles I-IV". In: *Proceedings of the MICCAI Grand Challenge: Neonatal Brain Segmentation*, pp. 16–21.
- Mikkola, K., N. Ritari, V. Tammiska, T. Salokorpi, L. Lehtonen, O. Tammela, L. Paakkonen, P. Olsen, M. Korkman, and V. Fellman (2005). "Neurodevelopmental outcome at 5 years of age of a national cohort of extremely low birth weight infants who were born in 1996-1997". In: *Pediatr.* 116.6, pp. 1391–1400.
- Miller, M.I., A. Banerjee, G.E. Christensen, S. Joshi, N. Khaneja, Grenander U., and L. Matejic (1997). "Statistical methods in computational anatomy". In: *Stat. Methods Med. Res.* 6, pp. 267–299.
- Mitchell, D.G. and M. Cohen (2004). *MRI principles*. Philadelphia: Saunders.
- Mogensen, A.O. and W.R. Vieth (1973). "Mass transfer and biochemical reaction with semipermeable microcapsules". In: *Biotechnol. Bioeng.* 15.3, pp. 467–481.
- Nossin-Manor, R., D. Card, D. Morris, S. Noormohamed, M.M. Shroff, H.E. Whyte, M.J. Taylor, and J.G. Sled (2013). "Quantitative MRI in the very preterm brain: Assessing tissue organization and myelination using magnetization transfer, diffusion tensor and T1 imaging". In: *NeuroImage* 64, pp. 505–516.
- Nossin-Manor, R., D. Card, C. Raybaud, M.J. Taylor, and J.G. Sled (2015). "Cerebral maturation in the early preterm period-A magnetization transfer and diffusion tensor imaging study using voxel-based analysis". In: *NeuroImage* 12, pp. 30–42.
- O'Shea, T.M., S.J. Counsell, D.B. Bartels, and O. Dammann (2005). "Magnetic resonance and ultrasound brain imaging in preterm infants". In: *Early Hum. Dev.* 81, pp. 263–271.
- Pham, D.L. and J.L. Prince (1999). "An adaptive fuzzy C-means algorithm for image segmentation in the presence of intensity inhomogeneities". In: *Pattern Recogn. Lett.* 20, pp. 57–68.
- Pohl, K.M., W.M. Wells, A. Guimond, K. Kasai, M.E. Shenton, R. Kikinis, W.E.L. Grimson, and S.K. Warfield (2002). "Incorporating non-rigid registration into expectation maximisation algorithm to segment MR images". In: *Proceedings of the 5th Interna-*

- tional Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*, pp. 264–571.
- Pohl, K.M., J. Fisher, W.E.L. Grimson, R. Kikinis, and W.M. Wells (2006a). “A Bayesian model for joint segmentation and registration”. In: *NeuroImage* 31, pp. 228–239.
- Pohl, K.M., J. Fisher, M.E. Shenton, R.W. McCarley, W.E.L. Grimson, R. Kikinis, and W.M. Wells (2006b). “Logarithm odds maps for shape representation”. In: *Proceedings of the 9th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*, pp. 955–963.
- Prastawa, M., J.H. Gilmore, W. Lin, and G. Gerig (2005). “Automatic segmentation of MR images of the developing newborn brain”. In: *Med. Image Anal.* 9, pp. 457–466.
- Press, W.H., S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery (1992). *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge: Cambridge University Press.
- Richards, F. (1959). “A flexible growth function for empirical use”. In: *J. Exp. Bot.* 10.2, pp. 290–301.
- Rorke, L.B. and H.E. Riggs (1969). *Myelination of the brain in the newborn*. Philadelphia: Lippincott.
- Rousseau, F., P.A. Habas, and C. Studholme (2011). “A supervised patch-based approach for human brain labeling”. In: *IEEE Trans. Med. Imag.* 30.10, pp. 1852–1862.
- Rousseeuw, P.J. and K. Van Driessen (1999). “A fast algorithm for the minimum covariance determinant estimator”. In: *Technometrics* 41.3, pp. 212–223.
- Rueckert, D., L.I. Sonoda, C. Hayes, D.L.G. Hill, M.O. Leach, and D.J. Hawkes (1999). “Nonrigid registration using free-form deformations: Application to breast MR images”. In: *IEEE Trans. Med. Imag.* 18.8, pp. 712–721.
- Rueckert, D., A.F. Frangi, and J.A. Schnabel (2003). “Automatic construction of 3-D statistical deformation models of the brain using nonrigid registration”. In: *IEEE Trans. Med. Imag.* 22.8, pp. 1014–1025.
- Rueckert, D., P. Aljabar, R.A. Heckemann, J.V. Hajnal, and A. Hammers (2006). “Diffeomorphic registration using B-splines”. In: *Proceedings of the 9th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*, pp. 702–709.

- Rumack, C.M., S.R. Wilson, J.W. Charboneau, and J. Johnson (2005). *Diagnostic ultrasound*. St. Louis: Elsevier Mosby.
- Rutherford, M.A. (2002). *MRI of the Neonatal Brain*. Edinburgh: Saunders.
- Rutherford, M.A., L. Srinivasan, L. Dyet, P. Ward, J. Allsop, S.J. Counsell, and F. Cowan (2006). “Magnetic resonance imaging in perinatal brain injury: Clinical presentation, lesions and outcome”. In: *Pediatr. Radiol.* 36, pp. 582–592.
- Saigal, S. and L.W. Doyle (2008). “An overview of mortality and sequelae of preterm birth from infancy to adulthood”. In: *Lancet* 371, pp. 261–269.
- Saigal, S., L. den Ouden, D. Wolke, L. Hoult, N. Paneth, D.L. Streiner, A. Whitaker, and J. Pinto-Martin (2003). “School-age outcomes in children who were extremely low birth weight from four international population-based cohorts”. In: *Pediatr.* 112.4, pp. 943–950.
- Santiago, P. and H.D. Gage (1993). “Quantification of MR brain images by mixture density and partial volume modeling”. In: *IEEE Trans. Med. Imag.* 12.3, pp. 566–574.
- Savitz, D.A., J.W. Terry, N. Dole, J.M. Thorp, A.M. Siega-Riz, and A.H. Herring (2002). “Comparison of pregnancy dating by last menstrual period, ultrasound scanning and their combination”. In: *Am. J. Obstet. Gynecol.* 187.6, pp. 1660–1666.
- Schwarz, G. (1978). “Estimating the dimension of a model”. In: *Ann. Stat.* 6.2, pp. 461–464.
- Serag, A., P. Aljabar, G. Ball, S.J. Counsell, J.P. Boardman, M.A. Rutherford, A.D. Edwards, J.V. Hajnal, and D. Rueckert (2012). “Construction of a consistent high-definition spatio-temporal atlas of the developing brain using adaptive kernel regression”. In: *NeuroImage* 59, pp. 2255–2265.
- Shattuck, D.W., S.R. Sandor-Leahy, Schaper K.A., D.A. Rottenberg, and R.M. Leahy (2001). “Magnetic resonance image tissue classification using a partial volume model”. In: *NeuroImage* 13, pp. 856–876.
- Shi, F., P. Yap, G. Wu, H. Jia, J.H. Gilmore, W. Lin, and D. Shen (2011). “Infant brain atlases from neonates to 1- and 2-year-olds”. In: *PLOS ONE* 6.4, e18746.
- Sotiras, A., C. Davatzikos, and N. Paragios (2013). “Deformable medical image registration: A survey”. In: *IEEE Trans. Med. Imag.* 32.7, pp. 1153–1190.

- Srinivasan, L. and M.A. Rutherford (2008). "MRI of the newborn brain". In: *Paediatr. Child Health* 18.4, pp. 183–195.
- Stannard, C., A. Williams, and P. Gibbs (1985). "Temperature/growth relationships for psychrotrophic food-spoilage bacteria". In: *Food Microbiol.* 2.2, pp. 115–122.
- Studholme, C. and V. Cardenas (2004). "A template free approach to volumetric spatial normalization of brain anatomy". In: *Pattern Recogn. Lett.* 25, pp. 1191–1202.
- Studholme, C., D.L.G. Hill, and D.J. Hawkes (1999). "An onverlap invariant entropy measure of 3D medical image alignment". In: *Pattern Recognition* 32, pp. 71–86.
- Talairach, J. and P. Tournoux (1988). *Co-planar stereotaxic atlas of the human brain: 3-dimensional proportional system: An approach to medical cerebral imaging*. Stuttgart: Thieme.
- Thirion, J.P. (1998). "Image matching as a diffusion process: An analogy with Maxwell's demons". In: *Med. Image Anal.* 2.3, pp. 243–260.
- Tohka, J., A.P. Zijdenbos, and A. Evans (2004). "Fast and robust parameter estimation for statistical partial volume models in brain MRI". In: *NeuroImage* 23, pp. 84–97.
- Unser, M. (1999). "Splines: A perfect fit for signal and image processing". In: *IEEE Signal Process. Mag.* 16.6, pp. 22–38.
- Van Buchem, M.A., S.C.A. Steens, H.A. Vrooman, A.H. Zwinderman, J.C. McGowan, M. Rassek, and V. Engelbrecht (2001). "Global estimation of myelination in the developing brain on the basis of magnetization transfer imaging: A preliminary study". In: *Am. J. Neuroradiol.* 22, pp. 762–766.
- Van Leemput, K., F. Maes, D. Vandermeulen, and P. Suetens (1999a). "Automated model-based bias field correction of MR images of the brain". In: *IEEE Trans. Med. Imag.* 18.10, pp. 885–896.
- Van Leemput, K., F. Maes, D. Vandermeulen, and P. Suetens (1999b). "Automated model-based tissue classification of MR images of the brain". In: *IEEE Trans. Med. Imag.* 18.10, pp. 897–908.
- Van Leemput, K., F. Maes, D. Vandermeulen, and P. Suetens (2003). "A unifying framework for partial volume segmentation of brain MR images". In: *IEEE Trans. Med. Imag.* 22.1, pp. 105–119.

- Van Wezel-Meijler, G., S.J. Steggerda, and L.M. Leijser (2010). “Cranial ultrasonography in neonates: Role and limitations”. In: *Semin. Perinatol.* 34.1, pp. 28–38.
- Wang, L., F. Shi, W. Lin, J.H. Gilmore, and D. Shen (2011). “Automatic segmentation of neonatal images using convex optimization and coupled level sets”. In: *NeuroImage* 58, pp. 805–817.
- Wang, L., F. Shi, G. Li, Y. Gao, W. Lin, J.H. Gilmore, and D. Shen (2014). “Segmentation of neonatal brain MR images using patch-driven level sets”. In: *NeuroImage* 84, pp. 141–158.
- Wang, S., M. Kuklisova-Murgasova, and J.A. Schnabel (2012). “An atlas-based method for neonatal MR brain tissue segmentation”. In: *Proceedings of the MICCAI Grand Challenge: Neonatal Brain Segmentation*, pp. 28–35.
- Wang, S., M. Kuklisova-Murgasova, J.V. Hajnal, C. Ledig, and J.A. Schnabel (2016). “Regression analysis for assessment of myelination status in preterm brains with magnetic resonance imaging”. In: *Proceedings of the 13th International Symposium on Biomedical Imaging (ISBI)*, pp. 278–281.
- Warfield, S.K., M. Kaus, F.A. Jolesz, and R. Kikinis (2000). “Adaptive, template moderated, spatially varying statistical classification”. In: *Med. Image Anal.* 4, pp. 43–55.
- Wells, W.M., W.E.L. Grimson, R. Kikinis, and F.A. Jolesz (1996). “Adaptive segmentation of MRI data”. In: *IEEE Trans. Med. Imag.* 15.4, pp. 429–442.
- Wozniak, J.R. and K.O. Lim (2006). “Advances in white matter imaging: A review of in vivo magnetic resonance methodologies and their applicability to the study of development and aging”. In: *Neurosci. Biobehav. R.* 30, pp. 762–774.
- Wright, R., A. Makropoulos, V. Kyriakopoulou, P.A. Patkee, L.M. Koch, M.A. Rutherford, J.V. Hajnal, D. Rueckert, and P. Aljabar (2015). “Construction of a fetal spatio-temporal cortical surface atlas from in utero MRI: Application of spectral surface matching”. In: *NeuroImage* 120, pp. 467–480.
- Wu, J. and B. Avants (2012). “Automatic registration-based segmentation for neonatal brains using ANTs and Atropos”. In: *Proceedings of the MICCAI Grand Challenge: Neonatal Brain Segmentation*, pp. 36–43.
- Xue, H., L. Srinivasan, S. Jiang, M.A. Rutherford, A.D. Edwards, D. Rueckert, and J.V. Hajnal (2007). “Automatic segmentation and reconstruction of the cortex from neonatal MRI”. In: *NeuroImage* 38, pp. 461–477.

- Yakolev, P.I. and A.R. Lecours (1967). “The myelogenetic cycles of regional maturation of the brain”. In: *Regional Development of the Brain in Early Life*. Ed. by A. Mankowski. Oxford: Blackwell, pp. 3–70.
- Young, J.M., T.L. Powell, B.R. Morgan, D. Card, W. Lee, M.L. Smith, J.G. Sled, and M.J. Taylor (2015). “Deep grey matter growth predicts neurodevelopmental outcomes in very preterm children”. In: *NeuroImage* 111, pp. 360–368.
- Zhang, J. (1992). “The mean field theory in EM procedures for Markov random fields”. In: *IEEE Trans. Med. Imag.* 40.10, pp. 2570–2583.
- Zhang, Y., M. Brady, and S. Smith (2001). “Segmentation of brain MR images through a hidden Markov random field and the expectation-maximization algorithm”. In: *IEEE Trans. Med. Imag.* 20.1, pp. 45–57.