

Assessment and Optimisation of 3D Optical Topography for Brain Imaging

Teresa M. M. Correia

A thesis submitted to
University College London
for the degree of
DOCTOR OF PHILOSOPHY

Supervised by
Professor Jeremy Hebden
Dr. Adam Gibson

Department of Medical Physics and Bioengineering
University College London
2010

Declaration

I, Teresa Correia confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Abstract

Optical topography has recently evolved into a widespread research tool for non-invasively mapping blood flow and oxygenation changes in the adult and infant cortex. The work described in this thesis has focused on assessing the potential and limitations of this imaging technique, and developing means of obtaining images which are less artefactual and more quantitatively accurate.

Due to the diffusive nature of biological tissue, the image reconstruction is an ill-posed problem, and typically under-determined, due to the limited number of optodes (sources and detectors). The problem must be regularised in order to provide meaningful solutions, and requires a regularisation parameter (λ), which has a large influence on the image quality. This work has focused on three-dimensional (3D) linear reconstruction using zero-order Tikhonov regularisation and analysis of different methods to select the regularisation parameter. The methods are summarised and applied to simulated data (deblurring problem) and experimental data obtained with the University College London (UCL) optical topography system.

This thesis explores means of optimising the reconstruction algorithm to increase imaging performance by using spatially variant regularisation. The sensitivity and quantitative accuracy of the method is investigated using measurements on tissue-equivalent phantoms.

Our optical topography system is based on continuous-wave (CW) measurements, and conventional image reconstruction methods cannot provide unique solutions, i.e., cannot separate tissue absorption and scattering simultaneously. Improved separation between absorption and scattering and between the contributions of different chromophores can be obtained by using multispectral image reconstruction. A method is proposed to select the optimal wavelength for optical topography based on the multispectral method that involves determining which wavelengths have overlapping sensitivities.

Finally, we assess and validate the new three-dimensional imaging tools using *in vivo* measurements of evoked response in the infant brain.

Acknowledgements

I could not have asked for a better supervisor than Professor Jeremy Hebden. I would like to thank him for giving me the opportunity to carry out this PhD. I am truly grateful for his continuous support, encouragement and guidance throughout my research and writing of this thesis. I would like to thank my second supervisor, Dr. Adam Gibson, for his guidance, advice, critical comments and friendship.

I would like to thank Professor Simon Arridge for his advice and suggestions. He was always very helpful and patient with my questions. I would also like to thank Dr. Martin Schweiger for his help and suggestions.

I would like to thank all my colleagues at the Biomedical Optics Research Laboratory for sharing their knowledge and providing an enjoyable research environment. A special thanks to my research colleagues and friends, Dr. Mauren Abreu de Souza, Dr. Christos Panagiotou, Dr. Caroline Reid, Dr. Jan Laufer, Dr. Thomas Allen, Dr. Louise Enfield, Dr. Nick Everdell, Dr. Chris Colley, Paul Burke, Rob Cooper, Dr. Bradley Treeby and Dr. Salavat Magazov. Other members of the team I had the pleasure to work and exchange ideas with include Dr. Gilberto Branco, Dr. Peck Hui Koh, Dr. Terence Leung, Dr. Anna Blasi and Professor Clare Elwell. I would like to thank Sarah Fox for providing the infant optical topography data. I would also like to thank Anil Banga, Imran Khakoo and Joanna Bruncker for making the thermochromic solid phantoms.

I am thankful to Fundação para a Ciência e a Tecnologia for the financial support, without which this project would not have been possible.

Finally, I would like to thank my parents for their patience, support and encouragement over the past years.

List of Publications

Correia, T, Gibson, A, Schweiger M and Hebden, J. 2009. Selection of regularization parameter for optical topography. *Journal of Biomedical Optics*, **14**, 034044. (Paper selected for the July 1, 2009 issue of Virtual Journal of Biological Physics Research).

Correia, T, Banga, A, Everdell, NL, Gibson, AP and Hebden, JC. 2009. A quantitative assessment of the depth sensitivity of an optical topography system using a solid dynamic tissue-phantom. *Physics in Medicine and Biology*, **54**, 6277-6286.

Hebden, JC, **Correia, T**, Khakoo, I, Gibson, AP and Everdell, NL. 2008. A dynamic optical imaging phantom based on an array of semiconductor diodes. *Physics in Medicine and Biology*, **53**, N407-N413.

Hebden, JC, Brunker, J, **Correia, T**, Price, BD, Gibson, AP and Everdell, NL. 2008. An electrically-activated dynamic tissue-equivalent phantom for assessment of diffuse optical imaging systems. *Physics in Medicine and Biology*, **53**, 329-337.

Correia, T, Gibson, A and Hebden, J. 2009. Identification of the optimal wavelengths in optical topography using photon density measurement functions. *Proceedings of SPIE*, **7187**, 718718. (Conference paper and poster).

Correia, T, Gibson, A and Hebden, J. 2008. Optimal selection of the regularization parameter for optical topography image reconstruction. In: *Biomedical Optics (BIOMED)*, OSA (*Optical Society of America*), BSuE49. (Conference paper and poster).

Contents

Declaration	2
Abstract	3
Acknowledgements	4
List of Publications	5
Table of Contents	6
List of Figures	10
List of Tables	19
List of Acronyms	21
1 Introduction	23
2 Tissue optics	25
2.1 Optical properties	25
2.1.1 Absorption	25
2.1.2 Scattering	26
2.1.3 Refractive index	27
2.2 Optical properties of the adult human head	27
2.2.1 Absorption spectra of the main chromophores	28
2.2.1.1 Water	28
2.2.1.2 Lipids	28
2.2.1.3 Haemoglobin	28
2.2.1.4 Cytochromes	29
2.2.2 Head tissues	30
2.2.2.1 Skin	30

2.2.2.2	Skull	32
2.2.2.3	Meninges	33
2.2.2.4	Brain	33
2.2.3	Brain Activity	35
2.2.3.1	Brain Functions	35
2.2.3.1.1	Social perception	36
2.2.3.2	Cerebral haemodynamics	37
2.2.3.3	Functional Activation	39
3	Current state of optical topography instrumentation and imaging	40
3.1	Types of instrumentation	40
3.2	Optical Topography	42
3.2.1	Main systems	42
3.2.1.1	UCL Optical topography system	45
3.3	Probe configuration	46
3.3.1	UCL topography probe	49
3.4	Image reconstruction	50
4	Imaging techniques in medicine	53
4.1	X-ray computed tomography	53
4.2	Ultrasound	54
4.3	Magnetic Resonance Imaging	56
4.4	Radionuclide Imaging	59
4.4.1	Single Photon Emission Computed Tomography	59
4.4.2	Positron Emission Tomography	59
4.5	Electrical Impedance Tomography	60
4.6	Comparison of brain imaging modalities	61
5	Image reconstruction	64
5.1	Forward problem	64
5.1.1	Radiative transfer equation	64
5.1.1.1	Diffusion equation	65
5.1.1.1.1	Boundary conditions	67
5.1.1.1.2	Source conditions	68
5.1.1.1.3	Analytical solutions - Green's functions	69
5.1.1.1.4	Numerical methods	71
5.1.1.1.4.1	Finite element method	71
5.1.1.1.4.2	Finite difference method	73
5.1.1.1.5	Validity of the diffusion approximation	74
5.1.1.2	Numerical solutions of the RTE	74
5.1.1.2.1	Monte Carlo method	74
5.1.1.2.2	Kubelka-Munk theory	75
5.2	Inverse problem	76

Contents	8
5.2.1 Jacobian calculation	76
5.2.2 Linear reconstruction	79
5.2.2.1 Tikhonov regularisation	80
5.2.2.2 Reconstruction methods	81
5.2.2.2.1 Conventional Method	81
5.2.2.2.2 Multispectral Method	82
5.2.3 Non-Uniqueness problem	83
5.2.4 Non-linear reconstruction	84
5.3 Summary	84
6 Selection of regularisation parameter	85
6.1 Singular Value Decomposition	86
6.2 Selection Methods	87
6.2.1 Generic methods	87
6.2.1.1 Methods that <i>do not</i> use noise statistics	87
6.2.1.1.1 Heuristic Method	87
6.2.1.1.2 L-curve	88
6.2.1.1.3 f-slope	89
6.2.1.1.4 Generalised Cross-Validation (GCV)	89
6.2.1.1.5 Quasi-Optimality Criterion (QOC)	90
6.2.1.2 Methods that use noise statistics	91
6.2.1.2.1 Unbiased Predictive Risk Estimator (UPRE)	91
6.2.1.2.2 Discrepancy Principle (DP)	92
6.2.1.2.3 Normalised cumulative periodogram (NCP)	92
6.2.2 Image based methods	93
6.2.2.1 Fixed Noise Figure (NF)	93
6.2.2.2 Blur Radius (BR)	93
6.2.2.3 Full Width Half Maximum	94
6.2.2.4 Contrast Ratios	95
6.2.2.5 Criteria	95
6.3 Regularisation parameter selection	95
6.3.1 Deblurring	95
6.3.2 Results	98
6.3.2.1 Example of methods and results - 5 % noise	98
6.3.2.2 All noise levels and realisations	103
6.3.3 Discussion	105
6.4 Image reconstruction of phantom data	107
6.4.1 Results	109
6.4.2 Discussion	115
6.5 Conclusions	117

7	Quantitative assessment of the depth sensitivity using tissue-like phantoms	120
7.1	Methods	123
7.1.1	Simulations	123
7.1.2	Optical topography phantom experiments	123
7.1.2.1	Liquid phantom	123
7.1.2.2	Solid phantoms	125
7.2	Variable regularisation parameter	127
7.3	Image analysis	130
7.4	Results and discussion	131
7.4.1	Simulations and liquid phantom experiments	131
7.4.2	Solid phantom	136
7.5	Conclusions	139
8	Identification of the optimal wavelengths using photon density measurement functions	141
8.1	Unique and simultaneous reconstruction	143
8.2	Interrogated volume	144
8.3	Methods	144
8.3.1	Similarity Measures	144
8.3.1.1	Sum of squared differences	144
8.3.1.2	Ratio-image uniformity	145
8.3.1.3	Correlation coefficient	145
8.3.1.4	Joint histogram and joint entropy	145
8.3.1.5	Mutual information	146
8.3.1.6	Test photon density measurement functions	147
8.3.2	Adult head model	147
8.3.2.1	Homogeneous model	147
8.3.2.2	Three layer model	148
8.3.3	Optimal wavelengths	148
8.3.4	Image reconstruction	149
8.4	Results and Discussion	149
8.4.1	Similarity measures test on photon density measurement functions	149
8.4.2	Optimal wavelengths	152
8.4.3	PMDF comparison	158
8.4.3.1	Volume	158
8.4.3.2	Full Width Half Maximum	158
8.4.3.3	Penetration depth	162
8.4.3.4	Sensitivity depth	163
8.4.3.5	Maximum sensitivity	164
8.4.3.6	Discussion	165
8.4.4	Image reconstruction	169
8.5	Conclusions	174

<i>Contents</i>	10
9 Imaging of functional brain activation	176
9.1 Methods	176
9.1.1 Stimuli	176
9.1.2 Data acquisition and processing	177
9.1.3 Image reconstruction	178
9.2 Results	178
9.3 Discussion	189
9.4 Conclusions	192
10 Summary and future developments	193
10.1 Use of priors in linear reconstruction	195
Appendix A Absorption and reduced scattering coefficients table	197
Appendix B Appendix	198
B.1 Spatially variant regularisation - summarised results and reconstructed images .	198
B.2 Reconstructed images	200
B.2.1 Simulations	200
B.2.2 Liquid phantom	204

List of Figures

2.1	Absorption spectrum of water.	28
2.2	Absorption spectrum of lipid (pig fat).	28
2.3	Absorption spectra of oxy- and deoxy-haemoglobin.	29
2.4	Absorption spectra of cytochrome c oxidase (Cyt aa3), cytochrome b (Cyt b) and cytochrome c (Cyt c).	30
2.5	Skull [Schünke <i>et al.</i> , 2007].	32
2.6	Meninges: dura mater, arachnoid mater and pia mater [Siegel & Sapru, 2005].	33
2.7	Cross-section of the brain showing the grey and white matter [Schünke <i>et al.</i> , 2007].	34
2.8	Absorption spectra of total haemoglobin (HbT).	35
2.9	Absorption coefficient μ_a spectra for skin, skull and intracerebral tissue.	36
2.10	Reduced scattering coefficient μ'_s spectra for skin, skull and intracerebral tissue.	36
2.11	Brain [Chiras, 2005].	37
2.12	Regions of the brain involved in social cognition.	38
3.1	Types of measurements of light through tissue.	41
3.2	UCL topography system with 32 laser sources and 16 detectors.	45
3.3	Block diagram of the UCL topography system	46
3.4	Hitachi probe configurations [Kawaguchi <i>et al.</i> , 2001].	47
3.5	Array used by Franceschini <i>et al</i> [2000] to probe the motor cortex during a hand tapping protocol.	47
3.6	Three optode arrangements: LA - lattice arrangement, DA - double-density arrangement and QA - quadruple-density arrangement [Yamamoto <i>et al.</i> , 2002].	48
3.7	High-density array with 24 sources (red) and 28 detectors (blue). The measurement channels are represented by the green lines [Zeff <i>et al.</i> , 2007].	49
3.8	Array showing the 16 sources (8 at each wavelength), 8 detectors and 3 source-detector distances. D1-D3 are the source-detector separations.	50

3.9	Array showing 64 channels. All the source-detector pairs collect data. The red and blue lines represent short and large separations respectively. D1-D15 are all the possible source-detector separations in cm.	50
3.10	Images of activation in the visual cortex obtained using a high-density array with 24 sources and 28 detectors.	52
4.1	Fourth generation CT scanner geometry [Michael, 2001].	54
4.2	Fifth generation CT scanner [Wolbarst & Hendee, 2006].	54
4.3	Computer tomography scan of an adult brain [Zimmermana <i>et al.</i> , 2006].	55
4.4	Harmonic imaging ultrasound and cranial computed tomography (CT) for comparison [Seidel <i>et al.</i> , 2004].	56
4.5	Ultrasound scan of a cerebral tumour located near the lateral ventricle [Harrer <i>et al.</i> , 2003].	56
4.6	External magnetic field \vec{B}_0 produces the net magnetisation \vec{M}_0 in the direction of the magnetic field [McRobbie & E A Moore, 2007].	57
4.7	T1 and T2 relaxation times after a 90° RF pulse.	57
4.8	After the 90° RF pulse, T2 relaxation causes rapid decrease in the transverse magnetisation and at the same time, T1 causes a slower increase in the longitudinal magnetisation \vec{M}_0 [McRobbie & E A Moore, 2007].	58
4.9	fMRI activation maps showing the regions of greatest response to social stimuli: eye, mouth and hand movements [Pelphrey <i>et al.</i> , 2005].	59
4.10	PET statistical maps showing regions involved in the assessment of facial emotion [Nakamura <i>et al.</i> , 1999].	60
4.11	EIT adjacent drive pattern [Graham & Adler, 2007].	61
4.12	EIT transverse, coronal and sagittal images of the response to visual and motor stimuli [Bagshaw <i>et al.</i> , 2003].	61
5.1	Geometry used to derive the Green's function solution to the diffusion equation for a semi-infinite homogeneous medium.	70
5.2	Geometry used to derive the Green's function solution to the diffusion equation for a infinite homogeneous slab.	70
5.3	A slab finite element mesh.	73
5.4	Absorption photon measurement density function for an optode separation of 35 mm.	79
5.5	Simplified diagrams illustrating the determination of the sensitivity function.	80
6.1	Tikhonov filter factors f_i	86
6.2	Example of an L-curve.	88
6.3	Original test image.	96
6.4	Blurred image with 5% added Gaussian noise.	96
6.5	Picard condition for the deblurring problem.	97
6.6	Relative error for the regularised solution for a single realisation of 5% noise.	97
6.7	Reconstructed image with $\lambda_{opt} = 0.038$	97

6.8	(a) L-Curve method and (b) curvature for different values of λ for the deblurring example.	99
6.9	f-slope as a function of $\ln(1/\lambda)$	99
6.10	Fixed NF as a function of λ	99
6.11	Blur radius method.	100
6.12	Generalised-cross validation function.	100
6.13	Quasi-optimality criterion.	100
6.14	FWHM, maximum intensity of ROI and weighted FWHM of the deblurred images.	101
6.15	Contrast to noise ratio curve.	102
6.16	DP method.	102
6.17	UPRE function.	102
6.18	Normalised cumulative periodograms for different regularisation parameters. The red straight line is the NCP for the λ selected using the NCP method.	103
6.19	Mean relative error of λ_{opt} , L-curve, f-slope, NCP, $\text{CNR} \cdot \Psi^{-1}$, GCV, QOC, DP and UPRE methods for different noise levels.	104
6.20	Phantom used to acquire optical topography data.	108
6.21	DPC for 3 different absorbing targets.	110
6.22	Images reconstructed of target C1.	111
6.23	Position of the target on the x axis of the reconstructed image for different λ values.	112
6.24	Position of the target on the y axis of the reconstructed image for different λ values.	112
6.25	Maximum absorption coefficient for each image reconstructed.	112
6.26	FVHM of the target in the images.	112
6.27	L-curve method for the optical topography problem.	113
6.28	The f-slope method fails to find a regularisation parameter λ for experimental data.	113
6.29	The GCV method and its minima for the phantoms in study.	113
6.30	Plots of the CNR method as a function of the regularisation parameter	114
6.31	The QOC fails to exhibit a minimum for C1 and C3.	115
6.32	DP method and the respective zeros for each function.	115
6.33	UPRE functions and its minima.	115
6.34	Normalised cumulative periodograms for C1 represented by the blue lines, C2 by the red and C3 the black lines.	115
7.1	The experimental setup.	124
7.2	Phantom experiments.	125
7.3	Schematic of the solid phantom containing seven targets impregnated with thermochromic dye located at different depths.	126
7.4	Solid phantom where each diode is embedded within a cylindrical target containing thermochromic pigment [Hebden <i>et al.</i> , 2008a].	127

7.5	Exponential and sigmoidal fit to the normalised regularisation parameters found using the L-curve method.	128
7.6	Graph showing how the apparent depth z_m of target in the image (reconstructed from simulated data) varies with true depth.	132
7.7	Graph showing how the apparent depth z_m of target in the image (reconstructed from experimental data) varies with true depth.	132
7.8	Each bar represents the depth root squared error (z_{RSE}) of the target in the images obtained using each regularisation type.	133
7.9	PSF width root squared error (PSF_{RSE}) of the reconstructed images using different regularisation types.	133
7.10	FVHM root squared error ($FVHM_{RSE}$) of the target in the images obtained using each regularisation type.	133
7.11	Contrast root squared error (C_{RSE}) of the reconstructed images using different regularisation types.	133
7.12	Lateral spatial accuracy of the target in the images obtained using constant regularisation (blue) and NOSER type regularisation (red).	134
7.13	PSF width of images obtained using constant regularisation (blue) and NOSER type regularisation (red).	134
7.14	FVHM of the target in the images obtained using constant regularisation (blue) and NOSER type regularisation (red).	134
7.15	Contrast accuracy of the reconstructed images using constant regularisation (blue) and NOSER type regularisation (red).	134
7.16	Image reconstructed using NOSER regularisation of targets T3 (left) and T2 (right).	135
7.17	Image reconstructed using NOSER regularisation of targets T3 (left) and T4 (right).	135
7.18	Optical topography images of each target in the xz plane.	136
7.19	Graph showing how resolution varies with depth.	137
7.20	Graph showing how contrast varies with depth.	137
7.21	Graph showing the difference between the real target position and apparent position in the image.	137
7.22	Graph showing how the depth of the target in the image varies with true depth.	137
7.23	Images of four perturbations embedded in a solid phantom, reconstructed using a spatially variant regularisation parameter, NOSER.	138
7.24	Images of two targets embedded in a solid phantom, reconstructed using NOSER types regularisation.	139
8.1	Three layer adult head model and probe placement.	148
8.2	Cross-section xy at $z=11.7$ mm of the medium used in the simulation.	149
8.3	2D photon density measurement functions (PMDFs) calculated for different optical properties in the plane xz at $y=42.5$ mm. PMDFs were compared against the reference PMDF ($PMDF_r$) to test the similarity measures.	151

8.4	2D absolute differences between PMDF_7 and each PMDF.	152
8.5	Plot of the sum of squared differences SSD, condition number \mathbf{k} and residual norm R , where each point represents a set of three wavelengths.	153
8.6	Histogram of the optimal wavelength distributions that satisfy the three wavelength selection conditions, for the homogeneous model, three wavelengths and two chromophores (HbO_2 and HHb).	154
8.7	Histogram of the optimal wavelength distributions that satisfy the high residual and low condition number criteria, for the homogeneous model, three wavelengths and two chromophores (HbO_2 and HHb).	154
8.8	Histogram of the optimal wavelength distributions that satisfy the three wavelength selection conditions, for four wavelengths and two chromophores (HbO_2 and HHb).	154
8.9	Histogram of the optimal wavelength distributions that satisfy the high residual and low condition number criteria, for four wavelengths and two chromophores (HbO_2 and HHb).	154
8.10	Histogram of the optimal wavelength distributions that satisfy the three wavelength selection conditions, for the three layer model, three wavelengths and two chromophores (HbO_2 and HHb).	155
8.11	Histogram of the optimal wavelength distributions that satisfy the three wavelength selection conditions, for the three layer model, four wavelengths and two chromophores (HbO_2 and HHb).	155
8.12	Sum of squared differences SSD, condition number \mathbf{k} and residual norm R for different wavelength pairs.	156
8.13	Histogram for the wavelength pairs obtained for condition $\text{SSD} \leq 0.005$, $\mathbf{k} \leq 8$ and $R \geq 1.5 \times 10^{-15}$	156
8.14	Wavelengths pairs for 6 conditions.	156
8.15	Histogram of the wavelengths that satisfy the low condition number criteria.	156
8.16	Histogram for wavelengths with the shortest distance to the point $(\text{SSD}, \mathbf{k}) = (0, 0)$	157
8.17	Sum of squared differences SSD and condition number \mathbf{k} for different wavelength pairs.	157
8.18	PMDFs Sum of squared differences SSD for wavelength pairs in the range 650-990 nm for absorption and diffusion.	157
8.19	PMDF profiles for 4 different wavelengths, for the homogeneous model.	159
8.20	PMDF profiles for 4 different wavelengths, for the 3 layer model.	160
8.21	Relative volume of the PMDF_{μ_a} (left) and PMDF_{κ} (right) for the homogeneous head model (top) and layered model (bottom).	161
8.22	FWHM for the depth direction of the PMDF_{μ_a} (left) and PMDF_{κ} (right) profiles for the homogeneous head model (top) and layered model (bottom).	161
8.23	Lateral FWHM of the PMDF_{μ_a} (left) and PMDF_{κ} (right) profiles for the homogeneous head model (top) and layered model (bottom).	162

8.24	Penetration depth of the PMDF_{μ_a} (left) and PMDF_{κ} (right) profiles for the homogeneous head model (top) and layered model (bottom).	163
8.25	Maximum and mean sensitivity depths of the PMDF_{μ_a} (left) and PMDF_{κ} (right) profiles for the homogeneous head model (top) and layered model (bottom). . .	164
8.26	Maximum sensitivity of the PMDF_{μ_a} (left) and PMDF_{κ} (right) profiles for the homogeneous head model (top) and layered model (bottom).	165
8.27	Reconstructed images of oxy-haemoglobin (left), deoxy-haemoglobin (centre) and scattering amplitude (right) from measurements at wavelengths 650 nm, 720 nm and 910 nm.	169
8.28	Reconstructed images of oxy-haemoglobin (left), deoxy-haemoglobin (centre) and scattering amplitude (right) from measurements at wavelengths 680 nm, 720 nm and 830 nm.	169
8.29	Reconstructed images of oxy-haemoglobin (left), deoxy-haemoglobin (centre) and scattering amplitude (right) from measurements at wavelengths 650 nm, 710 nm, 730 nm and 910 nm.	170
8.30	Reconstructed images of oxy-haemoglobin (left), deoxy-haemoglobin (centre) and scattering amplitude (right) from measurements at wavelengths 680 nm, 710 nm, 730 nm and 830 nm.	170
8.31	Reconstructed images of oxy-haemoglobin (left) and deoxy-haemoglobin (right) from measurements at wavelengths 690 nm and 880 nm.	171
8.32	Reconstructed images of oxy-haemoglobin (left) and deoxy-haemoglobin (right) from measurements at wavelengths 710 nm and 970 nm.	171
8.33	The xz cross section of the medium used in the simulation. Image reconstructed using the correct sensitivity matrix. Image reconstructed using the erroneous sensitivity matrix.	172
8.34	PMDF_r and PMDF_{720} profiles. The marks on PMDF_{720} represent a unit perturbation in the medium and the arrows indicate where it would occur if PMDF_r was used to locate the perturbation.	173
8.35	Depth difference between equal sensitivities of PMDF_r and PMDFs at wavelengths selected using our method (680 nm, 720 nm and 830 nm) and wavelengths selected using the two criteria method (650 nm, 720 nm and 910 nm).	174
8.36	Depth difference between equal sensitivities of PMDF_{720} and PMDFs at the other wavelengths selected using our method (680 nm and 830 nm) and wavelengths selected using the two criteria method (650 nm and 910 nm). . . .	174
9.1	Social stimuli presented to 13 healthy five-month old infants, which consisted of hand, eye and mouth movements for 10 seconds per trial [Lloyd-Fox <i>et al.</i> , 2009b].	177
9.2	Lateral and frontal array, showing the approximate position of the array of the infant head.	177

9.3	Reconstructed images of changes in [HbO ₂] occurring in the left hemisphere over time, showing the response to hand movements.	180
9.4	Reconstructed images of changes in [HHb] occurring in the left hemisphere over time, showing the response to hand movements.	180
9.5	Reconstructed images of changes in [HbO ₂] occurring in the left hemisphere over time, showing the response to mouth movements.	181
9.6	Reconstructed images of changes in [HHb] occurring in the left hemisphere over time, showing the response to mouth movements.	181
9.7	Reconstructed images of changes in [HbO ₂] occurring in the left hemisphere over time, showing the response to eye movements.	182
9.8	Reconstructed images of changes in [HHb] occurring in the left hemisphere over time, showing the response to eye movements.	182
9.9	Reconstructed images of changes in [HbO ₂] occurring in the right hemisphere over time, showing the response to hand movements.	183
9.10	Reconstructed images of changes in [HHb] occurring in the right hemisphere over time, showing the response to hand movements.	183
9.11	Reconstructed images of changes in [HbO ₂] occurring in the right hemisphere over time, showing the response to mouth movements.	184
9.12	Reconstructed images of changes in [HHb] occurring in the right hemisphere over time, showing the response to mouth movements.	184
9.13	Reconstructed images of changes in [HbO ₂] occurring in the right hemisphere over time, showing the response to eye movements.	185
9.14	Reconstructed images of changes in [HHb] occurring in the right hemisphere over time, showing the response to eye movements.	185
9.15	Reconstructed images of changes in [HbO ₂] occurring in the frontal lobe over time, showing the response to hand movements.	186
9.16	Reconstructed images of changes in [HHb] occurring in the frontal lobe over time, showing the response to hand movements.	186
9.17	Reconstructed images of changes in [HbO ₂] occurring in the frontal lobe over time, showing the response to mouth movements.	187
9.18	Reconstructed images of changes in [HHb] occurring in the frontal lobe over time, showing the response to mouth movements.	187
9.19	Reconstructed images of changes in [HbO ₂] occurring in the frontal lobe over time, showing the response to eye movements.	188
9.20	Reconstructed images of changes in [HHb] occurring in the frontal lobe over time, showing the response to eye movements.	188
9.21	<i>xz</i> cross-sectional images of changes in [HbO ₂] in the right hemisphere evoked by eye movements.	189
9.22	<i>xz</i> cross-sectional images of changes in [HbO ₂] in the right hemisphere evoked by mouth movements.	189

9.23	Schematic illustration of the regions of maximum haemodynamic activity for the three experimental conditions, for the left pad (left), frontal pad (centre) and right pad (right). For both hemispheres the posterior temporal area is on the left and the anterior area is on the right.	189
B.1	Images reconstructed from simulated data using a constant regularisation parameter λ_c , for absorption perturbations located at (a) 10 mm, (b) 20 mm and (c) 30 mm below the surface.	200
B.2	Images reconstructed from simulated data using an exponential decay function (E) as regularisation parameter, for absorption perturbations located at (a) 10 mm, (b) 20 mm and (c) 30 mm below the surface.	200
B.3	Images reconstructed from simulated data using a sigmoid decay function (S) as regularisation parameter, for absorption perturbations located at (a) 10 mm, (b) 20 mm and (c) 30 mm below the surface.	201
B.4	Images reconstructed from simulated data using the NOSER type regularisation parameter, for absorption perturbations located at (a) 10 mm, (b) 20 mm and (c) 30 mm below the surface.	201
B.5	Images reconstructed from simulated data using the NOSER type regularisation parameter and data covariance matrix C , for absorption perturbations located at (a) 10 mm, (b) 20 mm and (c) 30 mm below the surface.	202
B.6	Images reconstructed from simulated data using the variance uniformisation (VU) regularisation parameter, for absorption perturbations located at (a) 10 mm, (b) 20 mm and (c) 30 mm below the surface.	202
B.7	Images reconstructed from simulated data using the covariance matrix R as regularisation parameter, for absorption perturbations located at (a) 10 mm, (b) 20 mm and (c) 30 mm below the surface.	203
B.8	Images reconstructed from simulated data using the covariance matrices R and C as regularisation parameter, for absorption perturbations located at (a) 10 mm, (b) 20 mm and (c) 30 mm below the surface.	203
B.9	Images reconstructed from phantom data using a constant regularisation parameter λ_c , for absorption perturbations located at (a) 12 mm, (b) 22 mm and (c) 32 mm below the surface.	204
B.10	Images reconstructed from phantom data using an exponential decay function (E) as regularisation parameter, for absorption perturbations located at (a) 12 mm, (b) 22 mm and (c) 32 mm below the surface.	204
B.11	Images reconstructed from phantom data using a sigmoid decay function (S) as regularisation parameter, for absorption perturbations located at (a) 12 mm, (b) 22 mm and (c) 32 mm below the surface.	205
B.12	Images reconstructed from phantom data using the NOSER type regularisation parameter, for absorption perturbations located at (a) 12 mm, (b) 22 mm and (c) 32 mm below the surface.	205

- B.13 Images reconstructed from phantom data using the NOSER type regularisation parameter and data covariance matrix C , for absorption perturbations located at (a) 12 mm, (b) 22 mm and (c) 32 mm below the surface. 206
- B.14 Images reconstructed from phantom data using the variance uniformisation (VU) regularisation parameter, for absorption perturbations located at (a) 12 mm, (b) 22 mm and (c) 32 mm below the surface. 206
- B.15 Images reconstructed from phantom data using the covariance matrix R as regularisation parameter, for absorption perturbations located at (a) 12 mm, (b) 22 mm and (c) 32 mm below the surface. 207
- B.16 Images reconstructed from phantom data using the covariance matrices R and C as regularisation parameter, for absorption perturbations located at (a) 12 mm, (b) 22 mm and (c) 32 mm below the surface. 207

List of Tables

2.1	Optical properties found in the literature of brain, skull and skin.	31
4.1	Main characteristics of medical imaging modalities. Cost: \$ - < £300k, \$\$ - £300k – 1 million, \$\$\$ - > £1million.	62
6.1	Regularisation parameter λ and relative error ε obtained using the selection methods for the deblurring problem for a single noise realisation of 5% noise. .	98
6.2	Regularisation parameters λ obtained using the selection methods for the deblurring problem with different noise levels.	105
6.3	Relative error for the deblurring problem with different noise levels.	106
6.4	Regularisation parameters for the experimental data using different selection methods	116
7.1	Targets used in the experiment.	124
8.1	Optical properties used to generate PMDFs which were compared against the reference PMDF _r to find the SSD, CC, RIU, H, MI, NMI and $z_{centroid}$	150
8.2	PMDFs with large and small sensitivity volumes to find the SSD, CC, RIU, H, MI, NMI and $z_{centroid}$	152
A.1	Absorption and reduced scattering coefficients for skin, skull and brain tissue. .	197
B.1	Weighted average depth for the simulation and liquid phantom. The uncertainty in depth corresponds to the pixel size, which is 3.3 mm. The RSE values are the global depth accuracy.	198
B.2	PSF width obtained from the reconstructed images of the simulation and liquid phantom. The uncertainty is given by the pixel size, which has a mean value of 3.4 mm. The RSE values represent the global PSF width.	199
B.3	Normalised FVHM. The uncertainty is 0.0001. The RSE values represent the FVHM for all depths.	199

B.4	Contrast of the target in the reconstructed images from simulation and phantom data. The RSE values are the global contrast.	199
B.5	Apparent position (x and y) of the target in the images, spatial accuracy SA_{xy} , PSF width, FVHM and contrast.	199

List of Acronyms

BR	Blur Radius
CC	Correlation coefficient
CNR	Contrast to Noise Ratio
CSF	Cerebrospinal Fluid
CT	Computed tomography
CW	Continuous Wave
DPC	Discrete Picard Condition
DP	Discrepancy Principle
EIT	Electrical Impedance Tomography
FD	Frequency Domain
FEM	Finite Element Method
fMRI	functional Magnetic Resonance Imaging
FFT	Fast Fourier Transform
FVHM	Full Volume Half Maximum
FWHM	Full Width Half Maximum
GCV	Generalized Cross-Validation
IFG	Inferior Frontal Gyrus
MI	Mutual information
MRI	Magnetic Resonance Imaging
MBLL	Modified Beer-Lambert Law
mPFC	medial Prefrontal Cortex
NCP	Normalised Cumulative Periodogram
NF	Noise Figure
NIRS	Near Infrared Spectroscopy
NMI	Normalised Mutual Information
NOSE	Newton's One Step Reconstructor
PD	Penetration Depth
PET	Positron Emission Tomography
PMDF	Photon Measurement Density function
PSF	Point Spread Function
pSTS	posterior Superior Temporal Sulcus
QOC	Quasi-Optimality Criterion
RIU	Ratio-image uniformity
ROI	Region Of Interest
RSE	Root Squared Error
RTE	Radiative Transfer Equation
SPECT	Single Positron Emission Computed Tomography
SSD	Sum of squared differences
SVD	Singular Value Decomposition
TD	Time Domain
TOAST	Temporal Optical Absorption and Scattering Tomography
UCL	University College London
UPRE	Unbiased Predictive Risk Estimator
US	Ultrasound
VU	Variance Uniformisation

CHAPTER 1

Introduction

This PhD project is focused on establishing the optimum performance and inherent limitations of three dimensional optical imaging of evoked response in the brain using optical topography systems.

Optical topography monitors changes in the optical properties (absorption μ_a and reduced scattering μ'_s coefficients) and haemodynamic activity of the brain, from changes in near infrared (NIR) light reflectance measurements. The technique exploits the different absorption spectra of tissue chromophores (in particular, the spectra of oxy-haemoglobin and deoxy-haemoglobin) to measure their concentration levels (refer to §2). Optical topography has been used to measure functional activity noninvasively in studies of cognitive processes, to examine brain maturation in infants, and also to identify cerebral pathological conditions [Koizumi *et al.*, 2003; Strangman *et al.*, 2002] (refer to §3 for a review of the state-of-the-art of optical topography). The advantages of optical topography imaging over other existing imaging techniques (e.g. functional Magnetic Resonance Imaging (fMRI), Positron Emission Tomography (PET) and Electrical Impedance Tomography (EIT)) are its fast acquisition rate, the use of relatively inexpensive technology, its portability (enabling images to be acquired at bedside), and its use as a continuous monitor of brain activity (for more detail refer to §4). The source and detector fibers are attached to the head which makes it less sensitive to motion artifacts. The technique can distinguish between changes in oxy-haemoglobin concentration $\Delta[\text{HbO}_2]$ and changes in deoxy-haemoglobin concentration $\Delta[\text{HHb}]$, whereas fMRI can only measure blood flow and deoxy-haemoglobin concentration [HHb]. The penetration depth is limited by the diffuse nature of light in tissue to approximately half the source-detector distance [Strangman *et al.*, 2002]. Hence, optical topography can only image superficial cortical areas. Another limitation, which is also a consequence of the highly scattering nature of tissue, is the lack of spatial resolution, but localisation can be improved by optimising probe design (e.g. use of overlapping measurements) (refer to §3).

Multiple measurements using different combinations of sources and detectors enable

images to be generated, representing the optical properties of the tissues directly beneath the probe. However, the image reconstruction problem is ill-posed and under-determined, and therefore the direct solution of the least squares problem is not possible (refer to §5 for further detail). Regularisation methods are essential to solve ill-posed problems by introducing some constraints on the solution. This is achieved by the regularisation parameter, which controls the tradeoff between the fidelity to the measurements and the smoothness of the solution. Several methods have been used to select the regularisation parameter, which are summarised in this thesis and tested for simulated and experimental data (refer to §6).

In almost all the published literature images reconstructed from optical topography measurements are represented as two dimensional (2D) maps, whereas here we focus on three dimensional (3D) imaging. The reconstruction of 2D images is a less under-determined problem than the reconstruction of 3D images, given that the number of unknowns increases for volumetric reconstructions.

One of the main disadvantages of optical topography is its low penetration depth, leading to limited depth information. Perturbations in the medium can be more accurately located, in particular in the depth direction, by using a spatially variant regularisation. Optical topography simulations are used to analyse the influence of different spatially variant regularisations on the reconstructed images (refer to §7). Thereafter, the performance of the UCL system and image reconstruction algorithm is quantified, first using a liquid phantom and then using a dynamic solid phantom based on a thermochromic pigment.

Our optical topography system is based on continuous-wave (CW) measurements. However, it has been theoretically demonstrated that conventional image reconstruction methods cannot provide unique solutions from CW measurements [Arridge & Lionheart, 1998]. Multispectral reconstruction methods can help overcome this non-uniqueness problem and minimise cross-talk between chromophores [Corlu *et al.*, 2005]. This thesis presents a method of determining the optimum wavelengths based on the multispectral method that minimises the differences between interrogated volumes at different wavelengths (refer to §8). This method guarantees that the interrogated volumes at different wavelengths are similar and, hence, the cerebral activation measurements contain information from the same region. The main characteristics of the optically interrogated volumes are quantified using sensitivity profiles generated at different wavelengths using a homogeneous and three-layer adult head model. We also determine the influence of the selected optimal wavelengths on the accuracy of derived oxy- and deoxy-haemoglobin concentration changes.

Finally, the new three dimensional imaging tools are used to reconstruct maps of $\Delta[\text{HbO}_2]$ and $\Delta[\text{HHb}]$ from measurements of evoked response in the infant brain (refer to §9).

Near infrared light within the range 600-900 nm can travel relatively deeply into biological tissue due to the existence of a "spectral window", where absorption by the main chromophores, water and haemoglobin, is low. At higher wavelengths absorption is dominated by water and at lower wavelengths the absorption by haemoglobin is very high.

This chapter introduces the basic interaction of light with tissue. The main chromophores and anatomy of the adult human head are also described. Finally, the origin of optical contrast during functional brain activity is described.

2.1 Optical properties

Light travelling through tissue can be either absorbed or scattered. Tissue is highly scattering because photons describe a path whose direction is randomly changed by the chaotic variation in refractive index at cellular and sub-cellular level. Light is also absorbed in tissue by chromophores. This phenomenon occurs when a light wave strikes the particle within the medium (e.g. a haemoglobin molecule), whose electrons have the same natural frequency than the frequency of the light wave. The electrons will absorb the energy of the electromagnetic wave and oscillate with the frequency of the incident wave. Finally, this vibrational energy is converted into other forms of energy such as thermal energy.

2.1.1 Absorption

For a single chromophore, the absorption efficiency Q_a is given by the ratio of its absorption cross section σ_a to its actual geometric cross section A . That is, for a chromophore with an area A the effective cross-sectional area caused by absorption, where no photons propagate, is σ_a . The probability that a photon is absorbed per unit length in a medium containing chromophores with a volume density ρ , is the absorption coefficient μ_a , which is expressed as

$$\mu_a = \rho\sigma_a. \quad (2.1)$$

The absorption mean free path between absorption events is $1/\mu_a$. For an homogeneous

absorbing medium without scattering and a collimated beam with intensity I_0 , the intensity of light after travelling a distance l in the medium is

$$I = I_0 \exp(-\mu_a l), \quad (2.2)$$

which is known as the Beer-Lambert law. If the medium contains n different chromophores contributing to the overall absorption, the absorption coefficient is expressed as

$$\mu_a(\lambda) = \sum_n \varepsilon_n(\lambda) c_n, \quad (2.3)$$

where ε_n is the extinction coefficient, which is wavelength dependent, and c_n is the concentration of each chromophore.

2.1.2 Scattering

Most scattering interactions in tissue are elastic, so that the energy is conserved, and only the direction of propagation is changed. For a scattering particle, the scattering efficiency Q_s is given by the ratio of the scattering cross section σ_s and the geometric cross section S . The probability that a photon is scattered in a medium containing many scattering particles is called the scattering coefficient μ_s , expressed as

$$\mu_s = \rho_s \sigma_s, \quad (2.4)$$

where ρ_s is the volume density of scattering particles. The mean free path between scattering events is $1/\mu_s$. An equation similar to equation (2.2) can be written for the case of a homogeneous scattering medium by replacing μ_a with μ_s .

The probability of a photon travelling in direction \hat{s} being scattered into direction \hat{s}' is given by the normalised scattering phase function $p(\hat{s}, \hat{s}')$. Therefore, the integral over all angles is

$$\int_{4\pi} p(\hat{s}, \hat{s}') d\hat{s}' = 1. \quad (2.5)$$

If scattering is symmetric around the direction of travel, then for a random scattering medium the average phase function is a function of the angle θ between the incident and scattered directions: $p(\hat{s}, \hat{s}') = p(\cos(\theta))$. The anisotropy factor g is the mean cosine of the scattering angle

$$g = \int_{-1}^1 p(\cos(\theta)) \cos(\theta) d(\cos(\theta)), \quad (2.6)$$

which is a measure of the anisotropy of the phase function. When $g = 1$ the scattering vanishes and light is forward directed, for $g = 0$ scattering is isotropic and for $g = -1$ light is backscattered.

In biological tissue multiple scattering events occur, and light in tissue is strongly forward directed, with an anisotropy factor in the range $0.69 \leq g \leq 0.99$ [Cheong *et al.*, 1990]. The reduced (transport) scattering coefficient $\mu'_s = \mu_s(1 - g)$ is the effective number of isotropic scattering events per unit length, which is equivalent to many anisotropic scattering events. This

parameter is fundamental in the diffusion theory of light propagation through random media.

According to Mie scattering theory, the wavelength dependence of the reduced scattering coefficient μ'_s is given by

$$\mu'_s(\lambda) = a\lambda^{-b}, \quad (2.7)$$

where a is the scatter amplitude and b the scatter power, and both depend on the size, density of the scatterers and refractive index variations in the medium.

The total attenuation coefficient is defined as

$$\mu_t = \mu_a + \mu_s. \quad (2.8)$$

Similarly, the total transport attenuation coefficient is

$$\mu_{tr} = \mu_a + \mu'_s. \quad (2.9)$$

2.1.3 Refractive index

The speed of light in a medium with refractive index n is

$$v = \frac{c}{n}, \quad (2.10)$$

where c is the speed of light in a vacuum. Light travelling across media with different refractive indices causes changes in the direction of propagation. The average refractive index of biological tissue is generally assumed to be 1.4, which is between that of lipid and water [Bolin *et al.*, 1989]. Consider a photon with an angle of incidence θ_1 passing through a medium with refractive index n_1 enters a medium with refractive index n_2 at an angle of refraction θ_2 , Snell's law relates these quantities as follows

$$n_1 \sin(\theta_1) = n_2 \sin(\theta_2). \quad (2.11)$$

2.2 Optical properties of the adult human head

Although absorption of near infrared light is low in tissue, the absorption spectra of the main chromophores are different. In particular the spectra of oxy-haemoglobin (HbO₂) and deoxy-haemoglobin (HHb) are distinct (refer to §2.2.1.3), hence measurements performed at two different wavelengths, usually chosen on each side of the isosbestic point (wavelength at which the absorption is the same, approximately 805 nm), can provide information about relative concentrations of HbO₂ and HHb, and therefore about blood oxygenation. Concentrations of cytochrome oxidase, which is an indication of tissue oxygenation (found in the cell mitochondrial membrane), lipid and water can also be determined using appropriate wavelengths [Elwell, 1995; Cope & Delpy, 1988]. Typically the concentration of cytochrome c oxidase is less than 10% of that of haemoglobin [Cooper & Springett, 1997]. While concentrations of water and lipid are usually constant over time, haemoglobin and cytochromes depend on oxygenation levels and metabolism. The absorption spectra for the main tissue chromophores are described below.

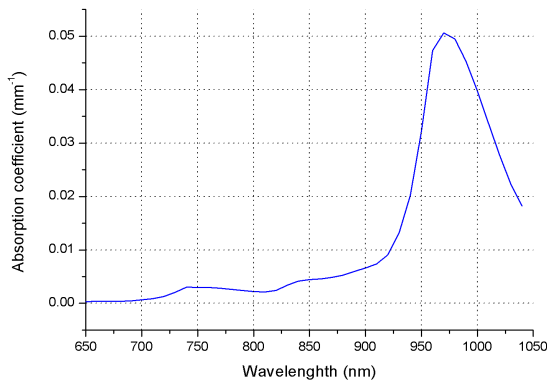


Figure 2.1: Absorption spectrum of water.

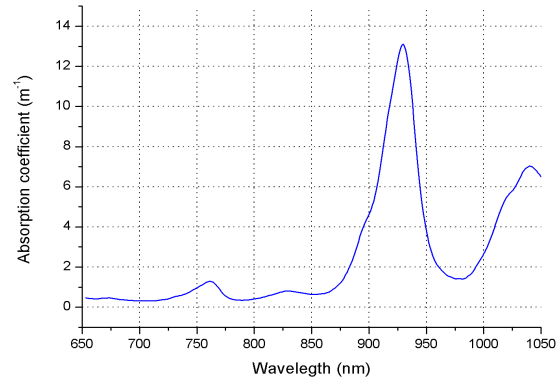


Figure 2.2: Absorption spectrum of lipid (pig fat).

2.2.1 Absorption spectra of the main chromophores

2.2.1.1 Water

The adult brain contains on average 80% water [Leung *et al.*, 2005]. The absorption spectrum of water is shown in figure 2.1 over the wavelength range 650-1040 nm [Matcher *et al.*, 1994]. Absorption is relatively low for shorter wavelengths, allowing light to travel several centimetres into the tissue. It increases rapidly at about 900 nm showing a strong absorption peak at 970 nm, and drops beyond this wavelength.

2.2.1.2 Lipids

The absorption spectrum of 100% lipid, obtained from pig lard, is shown in figure 2.2 for the range 650-1050 nm [van Veen *et al.*, 2004]. A strong absorption peak can be seen at 930 nm. An adult brain contains on average 11.6% lipid [White *et al.*, 1991]. Although the spectra for water and lipids are very similar, the contribution of the latter to the overall absorption of brain tissue is smaller due to its lower concentration. However, the proportion of lipids in the white matter is higher than in the grey matter.

2.2.1.3 Haemoglobin

Near Infrared Spectroscopy (NIRS) is mainly used to monitor changes in oxy- and deoxy-haemoglobin concentrations ($\Delta[\text{HbO}_2]$, $\Delta[\text{HHb}]$). From these measurements one is able to retrieve changes in cerebral blood volume (CBV), cerebral blood flow (CBF), oxygen saturation (SO_2) and total haemoglobin concentration ($[\text{HbT}]$). The specific absorption spectra of HbO_2 and HHb for wavelengths from 650 nm to 1050 nm can be seen in figure 2.3 [Cope, 1991]. The spectra show that oxy-haemoglobin has higher absorption for wavelengths larger than the isosbestic point at 805 nm, whereas deoxy-haemoglobin dominates for wavelengths below this point. The high absorption below 600 nm sets the lower limit for wavelengths used in NIRS and the upper limit at 900 nm is conditioned by the high absorption of water. The haemoglobin concentration in adult brain tissue is approximately 80 μM and oxygen saturation (SO_2) is 70% [Cope, 1991; Choi *et al.*, 2004].

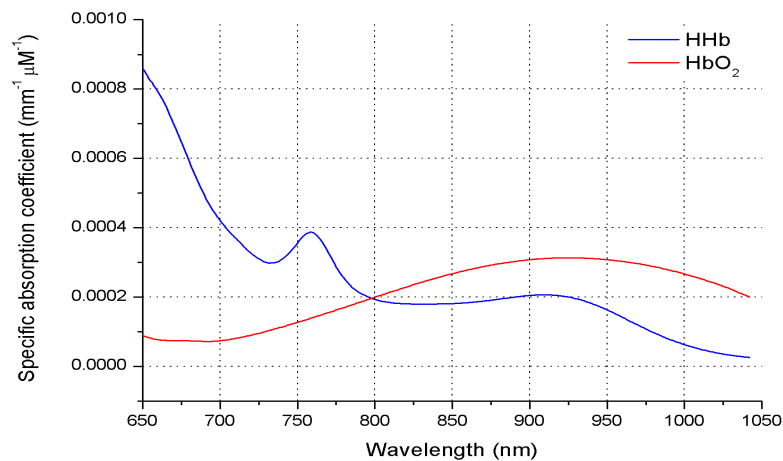


Figure 2.3: Absorption spectra of oxy- and deoxy-haemoglobin.

2.2.1.4 Cytochromes

Cytochromes are enzymes in the mitochondrial respiratory electron transport chain. Cytochromes can be found in the reduced or oxidised state and at the beginning of the electron chain the components tend to be more reduced while at the end they are likely to be almost fully oxidised [Cope, 1991]. The first component accepts electrons from a donor and passes them to the subsequent acceptor and so on until Cytochrome b (Cyt b) passes electrons to cytochrome c (Cyt c), which passes them to cytochrome c oxidase (Cyt aa3), the terminal electron acceptor of the chain, where oxygen is reduced to water.

The absorption spectra of cytochromes depend on their redox state. The absorption spectra for Cyt aa3, Cyt b and Cyt c are shown in figure 2.4 for the oxidised and reduced state [UCL, 2005]. Cytochrome c oxidase is more absorbing in the NIR region than any other cytochrome, and its spectrum has a broad absorption peak around 830 nm. For wavelengths below 700 nm the combined absorption of cytochromes is stronger. Although the absorption magnitudes of haemoglobin and cytochrome c oxidase are similar, the concentration of Cyt aa3 is much lower than that of haemoglobin. As mentioned previously the concentration of Cyt aa3 is less than 10 % of that of haemoglobin, and the typical total concentration of Cyt aa3 is considered to be 8 μM . Since about 80% of Cyt aa3 is oxidised it follows that the concentration of oxidised Cyt aa3 (oxCyt aa3) is 6.4 μM and the remaining 1.6 μM is in the reduced state (rCyt aa3). The ratios for reduced rCyt c : rCyt b : rCyt aa3 using the values found by van der Zee [1992] for adult grey matter are 1 : 1.32 : 2.37. The same ratios can be assumed for rCyt b and rCyt c concentrations which are found to be 0.89 μM and 0.68 μM respectively. Since near the electron source cytochromes are nearly fully reduced and vice-versa at the end of the chain, different concentration percentages can be attributed to the oxidised and reduced states of each cytochrome. Cyt c can be considered to have an equivalent proportion of oxidised (50 %) and reduced (50 %) Cyt c, whereas for Cyt b which is closer to the beginning of the chain, it can be considered that 60 % is in the reduced state and 40 % in the oxidised state. Consequently the concentrations of Cyt b and Cyt c are 2.37 μM and 1.36 μM respectively, and it follows that the

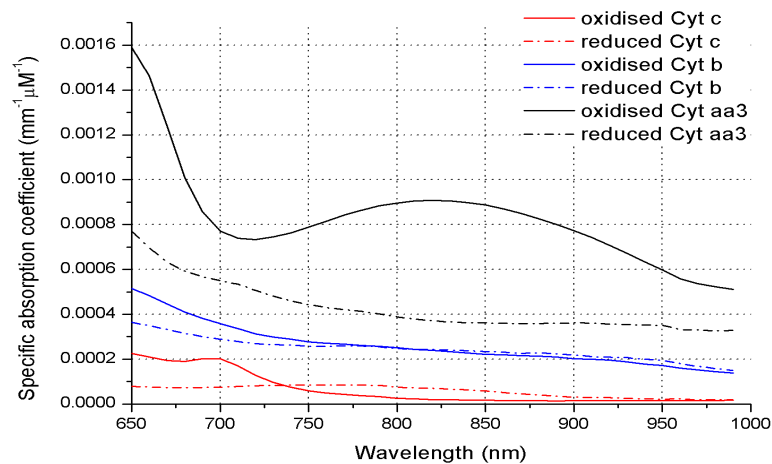


Figure 2.4: Absorption spectra of cytochrome c oxidase (Cyt aa3), cytochrome b (Cyt b) and cytochrome c (Cyt c).

total cytochrome concentration is $11.73 \mu\text{M}$.

2.2.2 Head tissues

The brain is surrounded and protected by the skin, skull, meninges (dura mater, arachnoid mater and pia mater), and cerebrospinal fluid (CSF). The central part of the brain is the white matter and the thin outer layer of grey matter, the cerebral cortex. Table 2.1 summarises the optical properties of the main head tissues found in the literature. These values correspond to measurements both *in vivo* and *in vitro*, but do not necessarily represent the true optical properties. *In vivo* methods to measure the optical properties require an indirect measurement, such as NIRS, where the optical properties may contain contributions from several types of tissue. The exception is when measurements are performed during open brain surgery. Meanwhile the water and residual blood are removed from *in vitro* samples and dyes are usually applied to increase contrast of specific tissues. Therefore, the optical properties of tissue samples may be altered.

The typical refractive index of soft tissue is about 1.4. Notice that scattering is much lower for neonatal brain than for the adult brain, in particular for white matter [van der Zee, 1992]. For neonatal brain the differences between grey and white matter are less obvious. This is a consequence of the smaller degree of myelination of the white matter and number of folds present in the neonatal brain than in the adult brain. Due to the soft and thin skull of neonates, combined with a small head size and thin CSF layer, it is more likely that light will penetrate into deeper regions of the brain than for the adult brain.

2.2.2.1 Skin

The skin varies in thickness from approximately 2-4 mm. The skin has two main layers, the epidermis and dermis. Underneath the skin there is the hypodermis, which mainly consists of adipose tissue. The skin protects against physical damage, external hazardous substances and helps regulate the body temperature.

Tissue	sample	λ (nm)	μ_a (mm ⁻¹)	μ'_s (mm ⁻¹)
Infant grey matter [Bevilacqua <i>et al.</i> , 1999]	<i>in vivo</i>	674	0.017-0.018	0.99-1.12
	<i>in vivo</i>	811	0.018-0.019	0.48-0.74
	<i>in vivo</i>	849	0.018-0.019	0.45-0.74
	<i>in vivo</i>	956	0.021-0.022	0.42-0.80
Infant white matter [Bevilacqua <i>et al.</i> , 1999]	<i>in vivo</i>	674	0.017	1.34
	<i>in vivo</i>	849	0.013	0.98
	<i>in vivo</i>	956	0.030	0.84
Infant skull [Bevilacqua <i>et al.</i> , 1999]	<i>in vivo</i>	674	0.021	1.19
	<i>in vivo</i>	849	0.022	0.91
	<i>in vivo</i>	956	0.036	0.77
Neonate brain [Zhao <i>et al.</i> , 2005]	<i>in vivo</i>	788	0.006-0.009	0.79-1.04
	<i>in vivo</i>	832	0.007-0.010	0.72-0.97
Adult white matter [Yaroslavsky <i>et al.</i> , 2002]	<i>in vitro</i>	670	0.07	6.02
	<i>in vitro</i>	850	0.10	4.1
Adult grey matter [Yaroslavsky <i>et al.</i> , 2002]	<i>in vitro</i>	670	0.02	0.84
Adult grey matter [Gebhart, 2006]	<i>in vitro</i>	650-900	0.04-0.09	0.7-1.1
Adult white matter [Gebhart, 2006]	<i>in vitro</i>	650-900	0.07-0.08	3.6-5.2
Adult grey matter [van der Zee, 1992]	<i>in vitro</i>	650-900	0.04-0.06	1.9-2.2
Adult white matter [van der Zee, 1992]	<i>in vitro</i>	650-900	0.02-0.03	8-10
Neonate grey matter [van der Zee, 1992]	<i>in vitro</i>	650-900	0.04-0.08	0.4-0.9
Neonate white matter [van der Zee, 1992]	<i>in vitro</i>	650-900	0.04-0.07	0.5-1.2
Pig skull [Firbank <i>et al.</i> , 1993]	<i>in vitro</i>	650-900	0.04-0.05	2.63-1.32
Dermis + epidermis [Simpson <i>et al.</i> , 1998]	<i>in vitro</i>	650-900	0.013-0.029	1.63-2.62

Table 2.1: Optical properties found in the literature of brain, skull and skin.

The epidermis is the outermost layer and is mostly composed of dead cells containing keratin, a protein that helps skin hydration. The deepest layer of the epidermis contains melanocytes that produce melanin. Melanin is a pigment that gives skin its natural colour and it absorbs light in the ultraviolet (UV) and visible range. The amount of melanin varies between individual and so does the attenuation coefficient.

Total Absorption Spectrum The absorption spectrum of skin has contributions from three main chromophores: haemoglobin, water and melanin. Haemoglobin concentration in skin is approximately 65 μM and oxygen saturation SO_2 is 71% [Zonios *et al.*, 2001]. The volume fraction of water content is approximately 60% [Meglinski & Matcher, 2002]. The wavelength dependence of the melanin absorption coefficient can be approximated by the expression [Jacques, 1998]

$$\mu_{a,melanin} = (6.6 \times 10^{10})\lambda^{-3.33} \quad (\text{mm}^{-1}), \quad (2.12)$$

where λ is the wavelength represented in units of nm. For moderately pigmented Caucasian skin the melanin content is approximately 13% [Jacques, 1998]. Considering the contributions from these chromophores and using equation 2.3, the total absorption coefficient of the skin layer can be represented as

$$\mu_{a,skin}(\lambda) = \varepsilon_{HbO_2} [HbT] SO_2 + \varepsilon_{HHb} [HbT] (1 - SO_2) + \mu_{a,H_2O} W + \mu_{a,melanin} M, \quad (2.13)$$

where M is the fraction of melanin content, W is the fraction of water content, μ_{a,H_2O} is the absorption coefficient of 100% water (mm^{-1}), ε_{HHb} is the specific absorption coefficient of HHb ($\mu\text{M}^{-1} \text{mm}^{-1}$) and ε_{HbO_2} is the specific absorption coefficient of HbO₂ ($\mu\text{M}^{-1} \text{mm}^{-1}$).

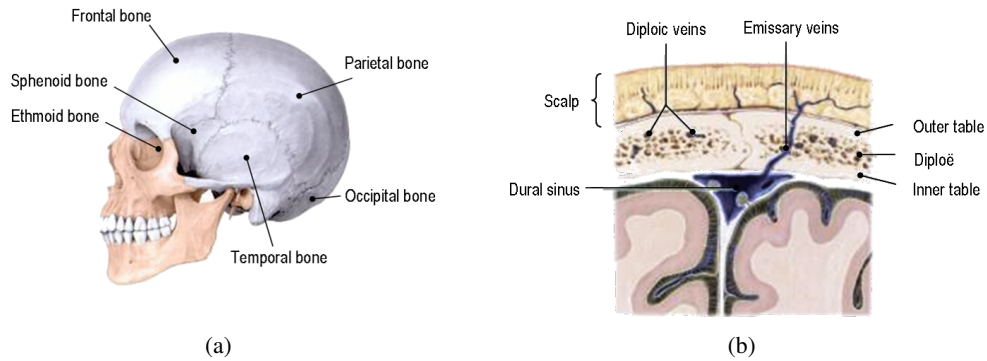


Figure 2.5: Skull [Schünke *et al.*, 2007]. (a) Cranial bones. (b) Magnified view of the skull, showing the bone structure, emissary veins, diploic veins and venous sinus.

The absorption coefficient values show a good agreement with the values presented by Simpson *et al* [1998].

Scattering spectrum The wavelength dependence of the reduced scattering coefficient of skin has been modelled as a combination of Mie and Rayleigh scattering given by [Bashkatov *et al.*, 2005]

$$\mu'_{s,skin} = 7.37\lambda^{-0.22} + 1.1 \times 10^{11}\lambda^{-4} \text{ (mm}^{-1}\text{)}, \quad (2.14)$$

where λ is the wavelength in nm. The first term corresponds to Mie scattering due to large collagen fibers and the second to Rayleigh scattering mainly due to small collagen fibrils, present in the dermis.

2.2.2.2 Skull

The cranium protects the brain against external injuries. There are 8 cranial bones, connected by a immobile fibrous connective tissue called sutures. It consists of a frontal, a sphenoid, a ethmoid, a occipital bone, two parietal and two frontal bones. Figure 2.5(a) shows the left lateral view of the cranial bones. Most cranial bones are flat bones, which are typically thin and slightly curved. They consist of internal and external surfaces of compact bone with a layer of spongy bone in between (see figure 2.5(b)). The spongy layer, the diploë, contains red bone marrow, where blood cell production occurs. The compact layer is responsible for the strength and rigidity of the skull. The emissary veins connect the intracranial venous sinuses with the veins of the scalp (see figure 2.5(b)). The diploic veins are located in the diploë and are connected to the emissary veins (see figure 2.5(b)). The skull does not have a homogeneous surface, the thickness of skull varies between 4 mm and 10 mm. It is thickest in the parieto-occipital regions and thinnest in the temporal region. The bone plates of the neonate skull are not fused and are connected by loose fibrous connective tissue called the fontanelles. The large gaps between bones are reduced by ossification of the joints by 18 months of age. Between the age of 7 and 13 years the skull approximates the adult size.

Firbank *et al* [1993] measured the optical properties of samples of bone from pig skull in the NIR. Absorption and scattering spectra of skull used in this thesis are described in [Firbank *et al.*, 1993].

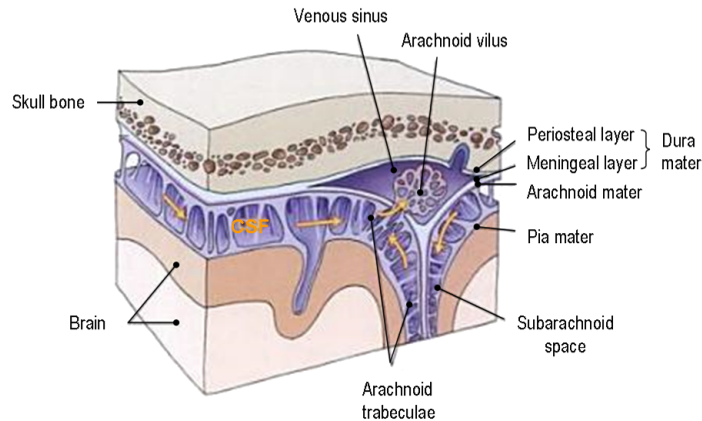


Figure 2.6: Meninges: dura mater, arachnoid mater and pia mater [Siegel & Sapru, 2005].

2.2.2.3 Meninges

The meninges are below the skull and above the brain. From the outermost layer inward they are the dura mater, the arachnoid mater, and the pia mater (see figure 2.6). The dura mater is a two-layer membrane separated in some regions to form the venous sinuses, which collect venous blood from the brain. It restricts the displacement of the brain when the head is moved by forming partitions in the brain. The space between the arachnoid and pia membranes is the subarachnoid space, which is approximately 1-2 mm thick and filled with CSF (see figure 2.6). The CSF is produced within the ventricles and acts as a shock absorbing medium, provides support and nourishes the brain. CSF is a transparent non-scattering fluid with optical properties similar to water.

The arachnoid villi are small knoblike projections of the arachnoid mater into the dura sinuses, through which the CSF flows (see figure 2.6). These structures function only in one direction, i.e., blood from the sinuses cannot flow back into the subarachnoid space. The arachnoid membrane is connected to the pia mater through small strands of fibrous connective tissue, the arachnoid trabecula. The largest blood vessels transverse the subarachnoid space, which are transformed into tiny capillaries in the pia mater that penetrate deep into the brain.

2.2.2.4 Brain

The largest part of the brain is the cerebrum, which is divided into two halves, the right and left hemispheres. The surface of the cerebrum is a 2-4 mm thick layer made of grey matter, known as cerebral cortex (see figure 2.7). This layer is highly folded, with many ridges and dips, the gyri and sulci. The folding increases the brain surface allowing the existence of a larger number of neurons. The neurons are nerve cells that consist of a body, the soma, and a long myelinated fiber, the axon. The cerebral cortex consists mainly of neuronal cell bodies. Beneath this layer is the white matter (see figure 2.7) formed by a large number of axons. The neurons, from different parts of the cortex, connect and communicate through the axons in the white matter. Information is ultimately transmitted to the spinal cord. White matter is more scattering than grey matter due to the presence of myelin sheaths around the axons.

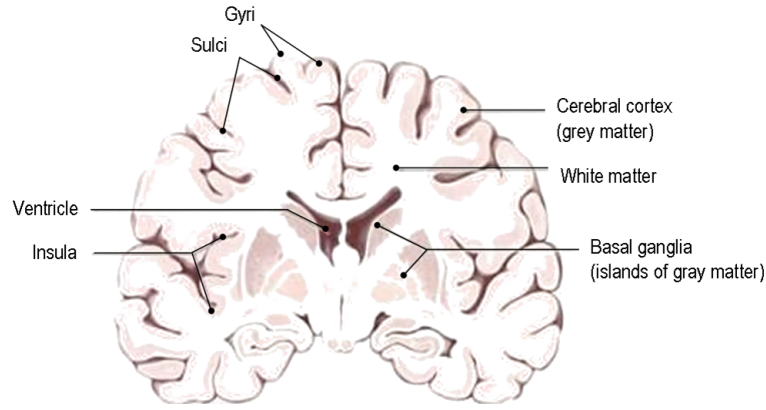


Figure 2.7: Cross-section of the brain showing the grey and white matter [Schünke *et al.*, 2007].

Total Absorption Spectrum Using the typical concentrations in the brain outlined previously for each chromophore, the total absorption coefficient is calculated from equation (2.3) as:

$$\begin{aligned} \mu_{a,brain} = & \varepsilon_{HbO_2} [HbT] SO_2 + \varepsilon_{HHb} [HbT] (1 - SO_2) + \mu_{a,H_2O} W + \mu_{a,Lipid} L + \\ & \sum_{aa3,b,c} \varepsilon_{oxCyt\ x} [oxCyt\ x] Ox + \varepsilon_{rCyt\ x} [rCyt\ x] (1 - Ox) + B \end{aligned} \quad (2.15)$$

where L is the fraction of lipid content, $\mu_{a,Lipid}$ is the absorption coefficient of 100% fat (mm^{-1}), $\varepsilon_{rCyt\ x}$ is the specific absorption coefficient of reduced cytochrome x ($\mu\text{M}^{-1} \text{mm}^{-1}$), $\varepsilon_{oxCyt\ x}$ is the specific absorption coefficient of oxidised cytochrome x ($\mu\text{M}^{-1} \text{mm}^{-1}$), Ox is the fraction of oxidised $Cyt\ x$ and B is a background absorption. The spectrum can be seen in figure 2.8, together with the total haemoglobin spectrum and the HbT spectrum with added chromophores. Adding the contribution of water to the HbT absorption coefficient results in a slight increase in absorption beyond 700 nm and becomes more significant for wavelengths larger than 900 nm. Lipids do not have a large contribution to the total absorption, there is only a slight increase in absorption around the absorption peak of fat at 930 nm. Cytochromes act approximately as a constant absorber beyond 750 nm with a slightly higher contribution from Cyt aa3 around its broad peak at 830 nm. Leung *et al* [2005] calculated the intracerebral absorption coefficient from the absorption contributions of HbO_2 , HHb, for a total haemoglobin concentration of 61 μM , 71% water and a wavelength independent background absorption B , which they set to a constant value so that the total absorption matches the value $\mu_a = 0.036 \text{ mm}^{-1}$ used by Okada and Delpy [Okada & Delpy, 2003a]. This spectrum is also shown in figure 2.8. The background absorption in equation (2.15) was set to $B=0.012 \text{ mm}^{-1}$ (figure 2.8) and it is clear that the contribution from the cytochromes has some influence on the total value of absorption considering that the assumptions made are approximately correct. The difference between the HbT + H_2O + Fat + Cyt +B spectrum and the one obtained by Leung *et al* [2005] is more evident for wavelengths below 750 nm and above 830 nm.

Scattering Spectrum Matcher *et al* [1997] measured μ'_s of the forehead of different human adults over the wavelength range 760 nm to 900 nm and observed that the scattering coefficient

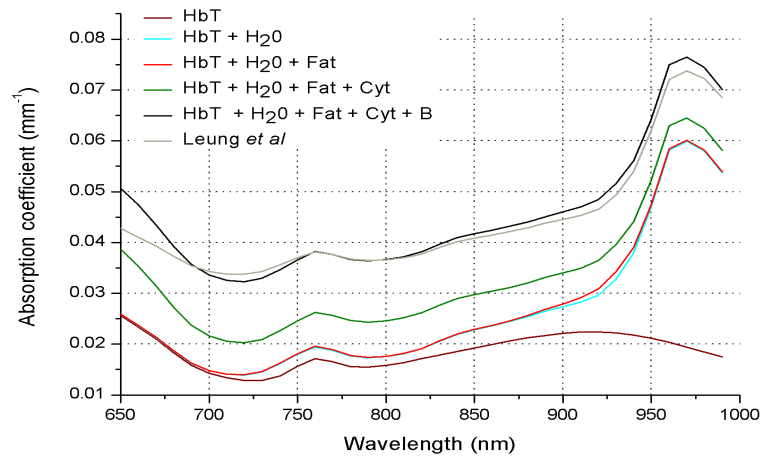


Figure 2.8: Absorption spectra of total haemoglobin (HbT), total haemoglobin with water (H_2O) added, fat added to the previous spectrum, cytochrome (Cyt) absorption added, background (B) added and spectrum by Leung *et al* [2007].

is approximately linear and represented by the relation $\mu'_s(\lambda) = a\lambda + b$ where $a = -6.5 \times 10^{-4}$ ($\text{mm}^{-1} \text{nm}^{-1}$) and $b = 1.45$ (mm^{-1}). They also fitted the experimental spectra to Eq.(2.7) and obtained $a = 32.08$ and $b = 0.53$. van der Zee [1992] measured μ'_s for grey matter *in vitro* and found $\mu'_s(800 \text{ nm})=2.2 \text{ mm}^{-1}$, which can be used as a reference number and included in the exponential relation found by Matcher *et al* [Matcher *et al.*, 1997], which then becomes

$$\mu_{s',brain}(\lambda) = \frac{2.2 (a\lambda^{-b})}{(a800^{-b})} \quad (2.16)$$

where a and b are the same as those in the previous exponential relation. Over the wavelength range 650 nm to 990 nm the maximum value at 650 nm corresponds to $\mu'_s=2.43 \text{ mm}^{-1}$ and it decays to the minimum value $\mu'_s=1.83 \text{ mm}^{-1}$ at 990 nm.

Figure 2.9 shows the absorption coefficients μ_a for skin, skull and brain tissue, for the wavelength range 650 - 990 nm. The reduced scattering coefficient spectra for the main head layers is shown in figure 2.10. A table with the absorption and reduced scattering coefficients for skin, skull and brain tissue can be found in appendix A. The optical properties for skin were obtained using equations (2.13) and (2.14), for brain tissue using equations (2.15) and (2.16) and for skull these values are described in Firbank *et al* [1993].

2.2.3 Brain Activity

Optical topography is commonly used to monitor functional responses to a stimulus. Different areas of the cerebral cortex can be activated since they respond to different stimuli. The main regions of the brain and their functions are briefly described. It is also important to understand cerebral oxygenation and blood flow.

2.2.3.1 Brain Functions

Large sulci in the cerebral cortex divide each hemisphere of the brain into four lobes: the frontal, temporal, parietal, and occipital (see figure 2.11(a)). The lobes have specific functions and can

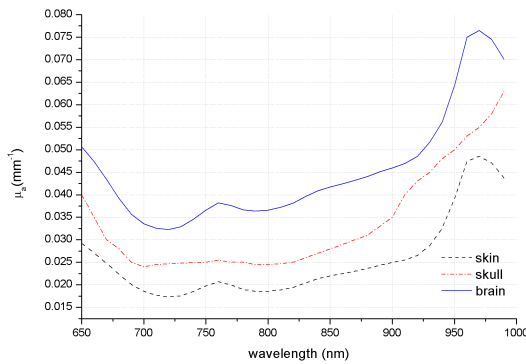


Figure 2.9: Absorption coefficient μ_a spectra for skin, skull and intracerebral tissue.

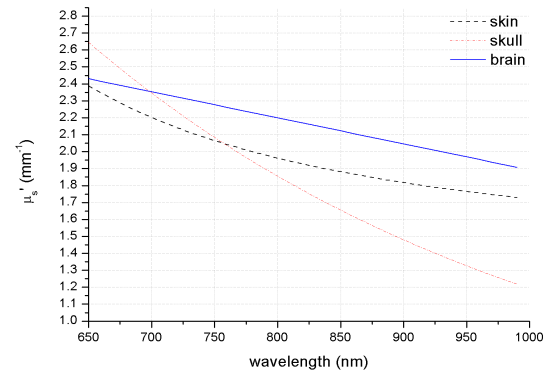


Figure 2.10: Reduced scattering coefficient μ_s' spectra for skin, skull and intracerebral tissue.

be subdivided into sensory, motor and association areas. Information processed in the right hemisphere is associated with signals received from the left side of the body and vice-versa. The motor areas are responsible for muscle activity, the sensory areas interpret sensory impulses and the association areas process and store higher level information. The motor cortex is located in front of the somatosensory cortex, in the frontal lobe. The motor cortex controls voluntary motor activity. Other parts of the frontal lobe are related to speech ability, emotions, problem solving, planning, among other functions. The somatosensory cortex is located in the parietal lobe. It contains receptors that process diverse stimuli, such as touch, pressure, temperature and pain, and identify its spatial origin. The parietal lobe also controls the ability to understand language. The visual cortex in the occipital lobe is associated with visual processing. The visual cortex is organised retinotopically, since each activated area of the retina corresponds to a specific area of the visual cortex. Different visual stimuli activate different specific neuronal regions in the visual cortex, since specific regions decompose the stimulation signal into shape, movement and colour. The auditory cortex is located in the temporal lobe and is responsible for the perception of the sound stimulus. It has a tonotopic organisation, which means that different areas respond to different frequencies. The main functional regions of the brain are shown in figure 2.11(b).

2.2.3.1.1 Social perception

Several areas of the brain are involved in social perception, which is the capacity to interact and understand others. It involves understanding the intentions of other individuals based on their words, facial expressions, body gestures, posture, behaviour, etc, and from these clues, predict their actions, make decisions and communicate with them. The brain areas involved in social cognition include the medial prefrontal cortex (mPFC), the posterior superior temporal sulcus (pSTS), the temporoparietal junction (TPJ), the inferior frontal gyrus (IFG), the interparietal sulcus (IPS), the anterior cingulate cortex (ACC), the anterior insula (AI) and the amygdala (see figure 2.12) [Blakemore, 2008].

The amygdala is located deep within the temporal lobe and is involved in the recognition of expressions such as fear. It is also activated in the presence of untrustworthy faces. The ACC is located within the medial frontal lobe and is anatomically connected to the amygdala, and it also reacts to fearful expressions. The AI is a cerebral structure that lies deep within

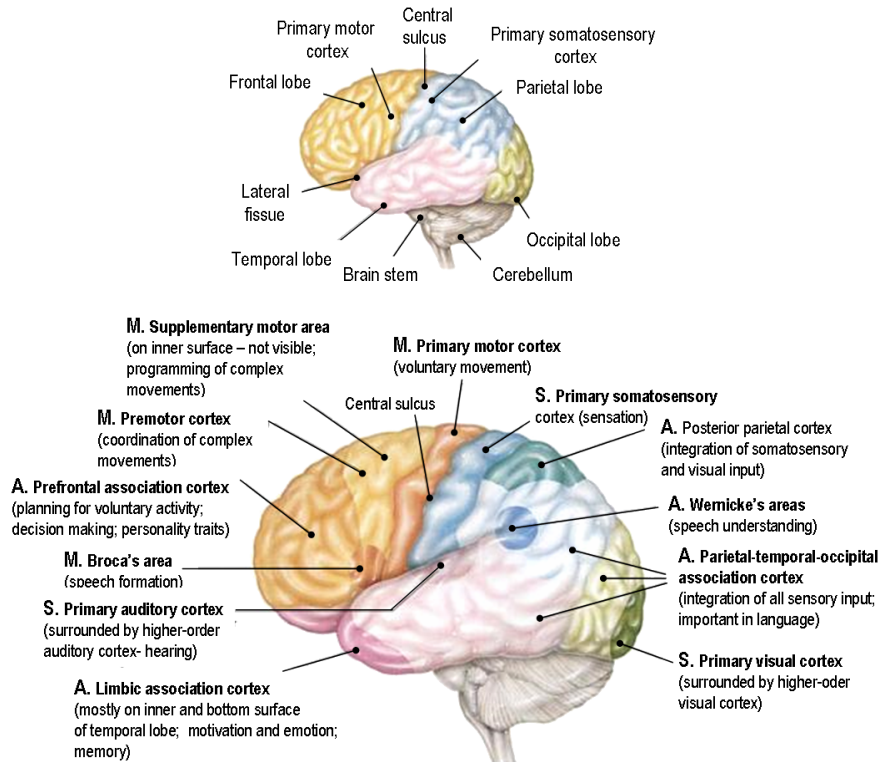


Figure 2.11: Brain [Chiras, 2005]. (a) Cerebral lobes. (b) Functional regions. The cerebral cortex has three principal functional areas: motor (M), sensory (S) and association (A) areas.

the lateral sulcus and processes expressions of disgust. The mPFC is a part of the prefrontal cortex (anterior part of the frontal lobe) involved in processing emotions, such as guilt and embarrassment. Mentalising tasks, i.e., understanding the mental state of other individuals and of ourselves, also activates this regions. The activity of the mPFC increases between childhood and adolescence and then decreases from adolescence to adulthood. The temporal lobe has two main sulci, which divide it into three gyri: the superior, middle and inferior temporal gyri. The pSTS is the sulcus separating the superior and middle temporal gyri. It activates with facial movement, for example, when the subject observes someone moving their eyes. It also activates with other biological motions which are socially relevant, such as hand gestures. The TPJ is the area where the temporal and parietal lobes meet and is involved in the process of mentalising. The IFG is located in the frontal lobe and is involved in emotional empathy, face and speech perception. The IPS separates the superior parietal lobule from the inferior parietal lobule in the posterior parietal cortex (located above the occipital lobe). It is involved in the analysis of shifts in eye gaze direction, which indicates if the attention is focused on the subject or not.

2.2.3.2 Cerebral haemodynamics

Oxygenated blood is supplied to the brain by two paired arteries, the carotid and vertebral arteries. These are connected forming the circle of Willis, which helps protect the brain against brain damage in case one of the major arteries becomes occluded, since the blood can flow across the circle of Willis. It also helps equalising the blood pressure in the anterior and posterior regions of the brain. In the brain small capillaries deliver oxygen and nutrients to

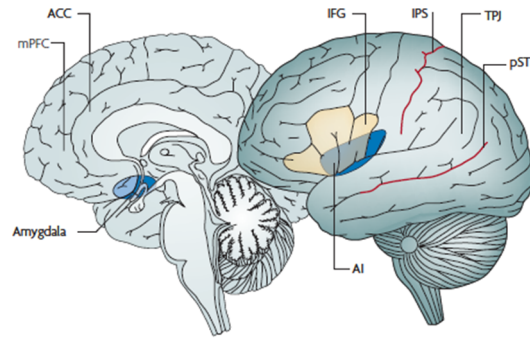


Figure 2.12: Regions of the brain involved in social cognition: medial prefrontal cortex (mPFC), posterior superior temporal sulcus (pSTS), temporoparietal junction (TPJ), inferior frontal gyrus (IFG), interparietal sulcus (IPS), anterior cingulate cortex (ACC), anterior insula (AI) and amygdala [Blakemore, 2008].

the cells. The veins collect the deoxygenated blood from the capillary system, removing the waste produced by the cells. The major fraction of venous blood is collected from the brain by the venous sinuses in the dura mater, which then transfer it to the internal jugular vein. The remaining venous blood in the brain is drained by the superficial cerebral veins and deep cerebral veins.

Optical measurements are sensitive to changes in the diameter of blood vessels. The cerebral blood vessels change their diameter in order to change the rate of blood flowing through them. The arteries are thicker and more elastic than veins, to withstand the high pressure. Approximately 70% of the blood volume is venous, because veins have a great capacity to store blood. According to the demands of the organism, veins can expand and contract controlling the amount of blood that returns to the heart. However, cerebral veins are different, since they have no valves and no muscle in their walls like other veins in the body. In the brain, the dural venous sinuses drain blood before returning to the heart.

Blood flow to the brain is approximately 700 ml/min [Buxton, 2002]. Cerebral blood flow (CBF) is the ratio between the perfusion pressure (CP) and the vascular resistance (CR):

$$CBF = \frac{CP}{CR}. \quad (2.17)$$

The perfusion pressure is the difference between the arterial pressure and the venous pressure. The blood pressure changes in the arteries when the heart pumps blood. Cerebrovascular resistance arises from friction and is given by

$$CR = V \frac{L}{D^4}, \quad (2.18)$$

where V is the blood viscosity, L is the blood vessel length and D is the vessel diameter. For the small vessels of the capillary bed the resistance is higher, hence the blood flow is slowest, allowing enough time to transfer oxygen and nutrients and remove waste products and carbon dioxide from the surrounding tissue.

Haemoglobin is a protein found in red blood cells, which has the ability to transport

oxygen and carbon dioxide. Each haemoglobin molecule is made of a globin group surrounded by four heme groups. Each heme group contains an iron atom that binds to oxygen. Therefore, the overall capacity of haemoglobin is four oxygen molecules. At high oxygen concentration the oxygen molecules bind to haemoglobin to form oxy-haemoglobin, but when the oxygen concentration is low the oxygen is released into the surrounding medium and deoxy-haemoglobin forms. In the presence of carbon dioxide the oxygen dissociates from haemoglobin easily. Deoxy-haemoglobin has a higher affinity for carbon dioxide than for oxygen. Hence, carbon dioxide is transported to the lungs in the venous blood. The oxygen saturation SO_2 , is a measure of how much haemoglobin is oxygenated relative to the total amount of haemoglobin (HbT) present in the blood vessels. Total haemoglobin is basically the sum of the concentration of oxy- and deoxy-haemoglobin. The SO_2 of arterial blood is typically 97-100 % and of venous blood is on average between 65-75% .

The brain is an extremely active organ with a high demand for oxygen and nutrients, this supply and blood flow are maintained constant by the cerebral autoregulation. Cerebral blood flow tends to maintain relatively constant despite changes in the cerebral perfusion pressure, by adjusting the vascular resistance. The resistance is altered by changing the diameter of the blood vessels. Therefore, constant blood supply to the brain is always guaranteed and the blood vessels are protected against an excessive increase in the blood pressure. The cerebral autoregulation is a very precise mechanism, and no other organ with an autoregulatory system has such sensitivity. However, the cerebral autoregulatory system is not infallible. It fails, for example, when the pressure exceeds approximately 150 mmHg, known as cerebral oedema, and when the pressure falls below about 50 mmHg, known as cerebral ischemia.

2.2.3.3 Functional Activation

When a region of the brain activates the arterial pressure remains constant, but the vessels dilate and, hence, the resistance reduces in order to increase the CBF. The CBF increases to supply additional oxygen. As a consequence the cerebral blood volume (CBV) also increases. When blood flow increases by increasing the blood volume the [HHb] remains constant. However, at the capillary bed the increase in CBF is mostly due to an increase in the blood velocity rather than capillary dilation. In this case the transit time decreases, i.e., blood spends less time in the capillaries, which reduces the capability of the capillary system to deliver oxygen to tissue. Fox and Raichle [Fox & Raichle, 1986] reported that the oxygen delivered to the brain (CBF) exceeds the oxygen consumption (cerebral oxygen metabolic rate, $CMRO_2$). As a consequence, the concentration of oxygen becomes higher in the veins than before activation, which is associated with a decrease in local deoxy-haemoglobin concentration [HHb]. Therefore, a typical functional activation causes an increase in local oxy-haemoglobin concentration [HbO_2] accompanied by a decrease in local deoxy-haemoglobin concentration [HHb]. The total haemoglobin HbT provides an estimation of CBV. Generally, the change in [HbO_2] is greater in magnitude than [HHb] and they do not necessarily reach the respective concentration peak at the same time. Optical topography can provide spatial and temporal maps representing the haemodynamic response to the cerebral activity.

Current state of optical topography instrumentation and imaging

In the 1970s Near Infrared Spectroscopy (NIRS) was first developed as a tool to monitor changes in total cerebral oxygenation [Jöbsis, 1977]. NIRS uses a single source-detector pair that provides global measurements of the concentration of tissue chromophores, without spatial information. Measurements are made in either reflection or transmission mode. Spatial maps that represent the optical properties of the object of study can be obtained by increasing the number of sources and detectors on the surface of the object. There are two imaging approaches: optical tomography and optical topography [Hebden, 2003]. Optical tomography is usually used to describe methods which involve reconstructing a 3D volume using multiple measurements of transmitted light across the volume. Meanwhile optical topography is usually applied to systems which obtain diffuse reflectance measurements from an array of source-detector pairs placed on the surface of the tissue, in order to map the changes in the optical properties of the tissue underneath the array, which can be converted into $\Delta[\text{HbO}_2]$ and $\Delta[\text{HHb}]$ [Hebden, 2003].

This chapter describes the different measurement techniques used in NIRS and imaging. Some of the most important optical topography systems and array configurations of source-detector pairs are introduced. Examples of functional activity studies performed using the different systems are also presented. Finally, an overview is given of the state-of-the-art in optical topography imaging.

3.1 Types of instrumentation

Three different measurement techniques have been implemented into NIRS and imaging systems: continuous wave (CW), frequency-domain (FD), and time-domain (TD) measurements (figure 3.1) [Gibson *et al.*, 2005]. CW systems acquire straightforward measurements of intensity attenuation (figure 3.1(a)). TD systems employ sources which emit short pulses of light and measure the times of flight of photons through tissue, providing a temporal distribution known as the Temporal Point Spread Function (TPSF), as seen in figure 3.1(b). In FD systems the light source is amplitude modulated at frequencies of few hundreds of MHz. The detected

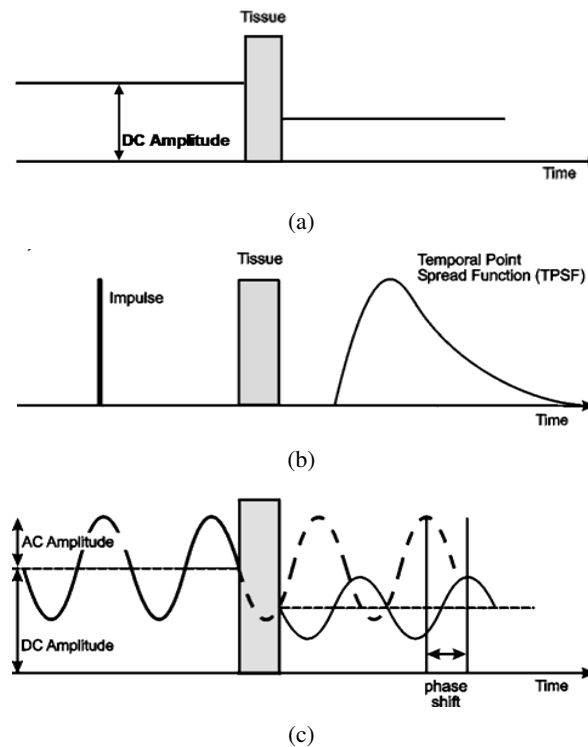


Figure 3.1: Types of measurements of light through tissue [Hebden *et al.*, 1997]. (a) Continuous wave measurement. (b) Time domain measurement. (c) Frequency domain measurement.

signal is compared to a reference signal and the phase shift and amplitude are recorded (figure 3.1(c)).

CW systems are relatively inexpensive and have a fast acquisition rate. A drawback of this type of system is the higher sensitivity to tissue immediately below the optodes, due to the so-called photon measurement density function (PMDF) which typically has a banana-like shape [Arridge, 1995]. The PMDF is a 3D map of the sensitivity of a measurement made using a given source and detector to changes in optical properties at each point within the medium [Arridge, 1995]. It represents the relative probability of detected photons having probed a certain volume. The PMDF theory and calculation method will be analysed with more detail in §5.2.1. CW measurements are highly affected by coupling changes, and for this reason the coupling coefficients have sometimes been introduced as unknowns in the reconstruction problem [Boas *et al.*, 2001c]. Another successful strategy to overcome this problem is to measure differences in intensity, where the coupling coefficients are assumed to remain constant and intensity measurements are recorded for a reference period and then during a period when changes in tissue occur [Gibson *et al.*, 2005]. However, measurements of intensity are not sufficient to separate the effects of absorption and scatter at a single wavelength [Arridge & Lionheart, 1998].

FD systems record both the amplitude and phase which allows absorption and scatter to be separated. These systems are still relatively inexpensive and have fast image acquisition (but slower than CW). TD systems use expensive electronics and their photon counting detectors are slow but highly sensitive. FD measurements are related to TD measurements via Fourier

transform. However, TD systems acquire measurements over a broad range of frequencies simultaneously, whereas FD systems typically use a single or a few discrete frequencies. TD detection is optimum when acquiring measurements at large source-detector separations, where the intensity is very low and photon counting techniques are required [Gibson *et al.*, 2005]. However, photon counting techniques inherently require lower photon fluxes and therefore need longer acquisition times [Boas *et al.*, 2001b].

The CW optical topography systems have the advantage of being fast, simple, compact and portable. In order to detect fast signals or to perform real-time imaging, fast acquisition rates are essential, hence systems which can activate all the sources simultaneously are required. Each source is modulated at a different frequency and the detected signal can be separated either by using lock-in detection or by fast Fourier transform (FFT) [Gibson *et al.*, 2005]. In lock-in detection, each lock-in amplifier is locked to the modulation frequency of the appropriate light source, hence, detected signals at frequencies other than the reference frequency are rejected by the lock-in amplifier. Alternatively, the contribution from each source to the detected signal can be isolated using a software that performs FFT. Systems which separate the signals by FFT are preferred over systems with lock-in detection, because they allow a large variety of source-detector arrangements requiring only minor changes to the software, and for a large number of channels an equivalent system would require a large number of lock-in amplifiers.

3.2 Optical Topography

3.2.1 Main systems

Wolf *et al* [2007], in their recent review on the progress of the NIRS and optical imaging instrumentation, provides an extensive list of the main recently developed systems. Optical topography systems are mostly CW systems that either record signals sequentially, activating one source at the time, or simultaneously, where each source is modulated at a different frequency. The first commercial optical topography system was the Hitachi ETG-100 system (with 24 channels, corresponding to 24 source-detector pairs which provide a measurement) [Yamashita *et al.*, 1999]. It incorporates eight laser light sources at 780 nm and eight at 830 nm, each source being modulated at a different frequency. Light is detected by eight avalanche photodiodes which are coupled to lock-in amplifiers that are locked to the respective source modulation frequency. The main advantage of the more recent Hitachi system ETG-7000 is its ability to perform measurements of the whole head using 40 sources and 40 detectors (120 channels). Another Hitachi system, the ETG-4000 (48 channels) system employs laser diodes at 695 nm and 830 nm.

The Hitachi optical topography systems have been used by researchers from the Hitachi Medical Corporation (Tokyo, Japan) and others to study the activity of normal brain and cerebral pathologies in adults, infants and newborns [Koizumi *et al.*, 2003; Otsuka *et al.*, 2007; Homae *et al.*, 2006; Koizumi *et al.*, 2005]. As an example, adult brain activation was studied during sleep, and the increase in the activation of the visual cortex during rapid eye movement (REM) sleep indicated brain activity related to dreaming [Igawa *et al.*, 2001]. In another study, the intake of alcohol did not significantly affect the haemodynamic activity detected in the visual cortex, when the subjects were submitted to repeated visual stimuli [Obata *et al.*, 2003].

Meanwhile, Watanabe *et al* [2002] used a Hitachi system to detect an increase in [HbO₂] and [HHb] during epileptic seizures. It has also been used to measure changes in haemoglobin of patients with psychiatric disorders, such as depression and schizophrenia [Suto *et al.*, 2004]. The Hitachi device was used to measure haemoglobin changes in the dorsal prefrontal cortex in children and adults playing video games for several hours, which showed a decrease in the oxy-haemoglobin concentration probably due to the player's attention on the video game [Matsuda & Hiraki, 2006]. Pena *et al* [2003] investigated the response of neonates to linguistic stimuli. Neonates at only few days of age responded significantly to normal speech, while showing a much smaller response to reversed speech and silence. Infants aged 2 to 4 months were submitted to visual and auditory stimuli in order to study the possibility of connections between different areas of the brain [Taga & Asakawa, 2007]. It was verified that infants at this age have separate cortical areas to process auditory and visual information.

Due to its commercial availability and well-established reputation in the field, the Hitachi systems are widely used by researchers in functional activation studies. However, most of the reported studies generally only describe global measurements of blood volume changes. The information from each channel is used to find the variation of cerebral oxy- and deoxy-haemoglobin independently. These systems are equipped with a simple imaging reconstruction software, where images are obtained by interpolating the measurements into the image space. Although it provides a visual tool to help researchers identify whether changes occur, the spatial resolution in the lateral direction is quite low and no attempt is made to reconstruct along the depth direction.

The CW4 system (developed at Photon Migration Lab, Massachusetts General Hospital, and TechEn Inc.) is frequency multiplexed and the detected signals are separated by Fourier transformation. This technology allows a large range of array designs by simply modifying the software, which represents an advantage over the previously described system that requires changes in the hardware. It has 16 laser diodes (8 lasers at 690 nm and 8 at 830 nm) and 16 avalanche photodiodes, and acquires data with a temporal resolution of 50Hz [Franceschini *et al.*, 2003]. Franceschini *et al* [2003] used the CW4 system to study the activation of the sensorimotor cortex in adults, using voluntary and nonvoluntary stimuli. The results were consistent with previous studies: they observed an increase in [HbO₂] and a decrease in [HHb] contralateral to the stimulated side, with the ipsilateral response appearing smaller than the contralateral response. Using a similar protocol but with 18 laser diodes instead, Franceschini *et al* [2004] were able to detect fast optical changes presumably from neuronal tissue (of the order of 50-100 ms) in approximately 60% of the subjects studied. More recently the CW5 system has been developed, consisting of 32 sources and 32 detectors, which enables the whole adult head to be covered in order to simultaneously monitor the prefrontal, motor, and visual cortices [Franceschini *et al.*, 2006]. Using this system they found that optical signals are correlated with physiological signals associated with the cardiac cycle, respiration, and blood pressure.

Bluestone *et al* [2001] measured the haemodynamic changes on the human forehead during a Valsalva manoeuvre, which is an expiration against a closed airway causing high intrathoracic pressure and low cardiac output, and found strong changes in [HbO₂] by comparison with

[HHb]. They used a CW system with frequency-encoded wavelengths, time multiplexed source illumination, lock-in detection and an acquisition rate of 3 Hz [Schmitz *et al.*, 2000]. This system has been improved recently [Schmitz *et al.*, 2002] and also the DYNOT system (manufactured by NIRx) has been introduced [Schmitz *et al.*, 2005]. The DYNOT system uses a time multiplexed scheme, where the sources are illuminated sequentially, which decreases the frame rate but has the advantage of increasing the dynamic range.

Another commercially available instrument is the optical multichannel monitoring (OMM) system, developed by Shimadzu Corporation. It can have up to 16 light source fibers and 16 detectors, each light source emitting at 780 nm, 805 nm, and 830 nm. This system was used by Hoshino *et al* [2005] to study subdural grid electrodes implanted in the motor cortex, which are used to identify functional brain areas before surgery, and successfully detected electrically induced changes with optical topography. Emotional response was evaluated by submitting 6 adults to visual stimuli consisting of attractive face versus unattractive faces [Mitsuda & Yoshida, 2005]. The left anterior frontal cortex showed an increase in [HbO₂] when subjects considered a face to be unattractive.

Chance *et al* [1998] used a system with 4 detectors and 9 sources at a single wavelength, which are switched sequentially, to image sensorimotor and cognitive activation in adults and neonates. Although it was an intensity based system they used a phase cancellation technique, so that the amplitude is null for the detector placed between two equidistant sources. The null is highly sensitive to small perturbations in the medium.

Zeff *et al* [2007] successfully monitored functional activation of the visual cortex using a high-density diffuse optical topography system. High density imaging arrays (24 sources and 28 detectors), high dynamic range and low inter-channel crosstalk are achieved by frequency, time and spatial encoding. The time encoding scheme decreases the cross-talk by illuminating nearby sources sequentially [Zeff *et al.*, 2007; White & Culver, 2009]. Frequency encoding assigns specific modulation frequencies allowing sources to be illuminated simultaneously and increase the acquisition rate. Spatial encoding allows distant sources to be illuminated simultaneously, which also increases the acquisition rate and reduces cross-talk. High dynamic range is also achieved by using isolated and dedicated detector channels. The frame rate was 12 Hz. Most of the available optical topography systems lack the technology necessary to perform high density measurements, which is highly desirable since it can improve resolution, localisation and contrast [Zeff *et al.*, 2007; White & Culver, 2009]. Nevertheless, systems like the University College London (UCL) optical topography system, which is described below, could be relatively easily adapted and programmed to accommodate high density imaging arrays.

As mentioned earlier, measurement of intensity alone does not allow the separation between the effects of absorption and scattering [Arridge & Lionheart, 1998]. To overcome this limitation FD and TD systems have been used for optical topography. An example of a FD system is the one used by Franceschini *et al* [2000] to study the adult motor cortex. The system is commercially produced by ISS Inc; it employs 16 laser diodes (758 nm and 830 nm) and measurements on all the 32 channels are acquired in 160 ms. Selb *et al* [2005] developed a TD system with a pulsed laser emitting in the range 750 nm to 850 nm and the signal is recorded



Figure 3.2: UCL topography system with 32 laser sources and 16 detectors.

by a intensified CCD camera, which performs a gated detection of the TPSF through narrow temporal windows.

3.2.1.1 UCL Optical topography system

The optical topography system developed at UCL (figure 3.2) has 32 diode laser sources (16 at 770 nm and 16 at 850 nm) [Everdell *et al.*, 2005]. The system can operate at other wavelengths by simply replacing the laser sources. Each source is modulated at a different frequency, ranging from 2 kHz to 4 kHz. A single 20 MHz quartz crystal oscillator provides all the multiplexing frequencies. Then the signal is converted into kHz frequencies by a frequency divider. Each laser has a mean power of approximately 2 mW and the intensity is controlled by a digital potentiometer through software, which was implemented in C using LabWindows CVI (National Instruments, Austin, Texas). A driver circuit is used to modulate the laser diode. Light passing through the tissue is collected by 16 avalanche photodiode detectors that detect light from all the sources at the same time. The signal is low-pass filtered to prevent aliasing with the 10 kHz Nyquist frequency of the analogue-to-digital converter, which samples each channel at 20 kHz. A Fast Fourier Transform (FFT) is performed on the detected data to isolate the signals from each source. This system allows great flexibility in the arrangement of the optodes and in the selection of source-detector pairs to be used, which only requires small changes in the software. With 16 detectors active the system can operate at 20 frames/s. Figure 3.3 shows a schematic diagram of the UCL topography system.

The UCL optical topography system has been used to study visual activation in 4 month-old infants [Blasi *et al.*, 2007]. The stimuli consisted of visual noise, face images and cartoons. Images of $\Delta[\text{HbO}_2]$ were reconstructed which showed an increase in HbO_2 concentration in different areas of the visual cortex for different stimuli. They verified that face stimuli induced haemodynamic changes across a larger volume than visual noise stimuli.

More recently language processing studies in babies up to 2 months-old have been performed. The probe was redesigned and divided in two, so that both hemispheres of

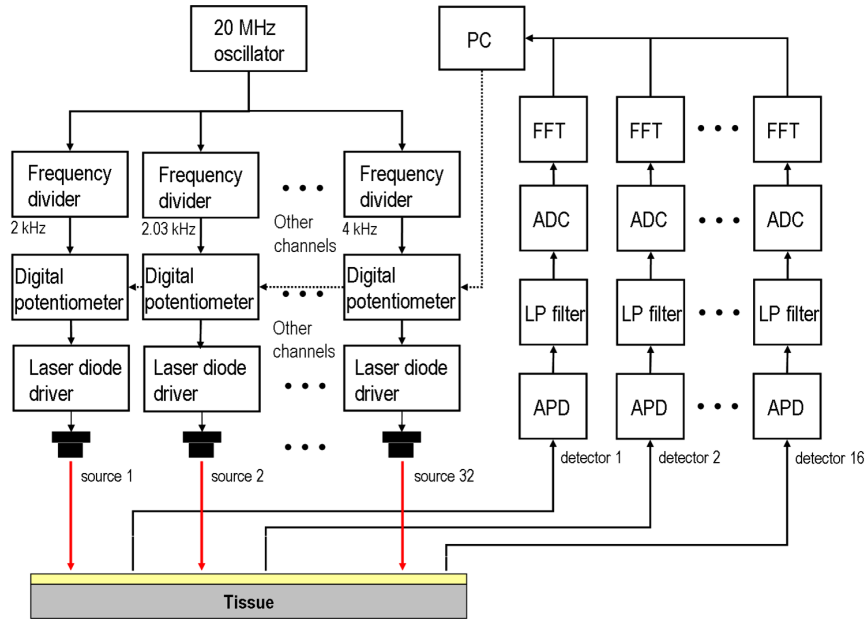


Figure 3.3: Block diagram of the UCL topography system: APD - avalanche photodiode detector, LP filter - low pass filter, ADC - analog-to-digital converter, FFT- fast Fourier transform.

the auditory cortex can be investigated at the same time [Branco, 2007]. In one of the studies the stimuli consisted of 6 different conditions: native language (English), foreign language (Arabic), emotional sounds (human vocalisations), animal vocalisations (monkey calls), scrambled controls, and silence. These investigations, and other language processing studies, are ongoing.

3.3 Probe configuration

As mentioned earlier in §3.1, the volume of tissue contributing to measurements is described by the PMDF, whose mean depth depends on the source-detector distance and is typically around half the optode spacing. Thus, for a source-detector separation of 30 mm, the sampled region is a banana-shaped region which is narrowest immediately below the source and detector, and broadest mid-way between them, centred at a depth of about 15 mm.

If the source-detector separations are kept small only the superficial areas of the tissue are sampled, but the advantage is that the signal is relatively high. On the other hand, if the separations are large, light will reach deeper areas but has to travel a bigger distance and will be highly attenuated, resulting in a lower intensity signal. At source-detector distances of approximately 4-5 cm the typical detected light levels on the head are low and noisy. Typically the scalp and skull are approximately 1 to 2 cm thick in an adult (depending on the region of the head), and it is assumed that these areas do not change during a measurement, hence it is necessary to choose a separation that allows light to reach areas of the brain where changes in haemoglobin concentration occur. The optode spacing should be chosen according to the type of study performed, and the spacings that do not contribute with information or even cause deterioration of images should be excluded.

The Hitachi system probes commonly employ the configurations shown in figure 3.4, where the dimensions, the number of sources, and the number of detectors can vary. The

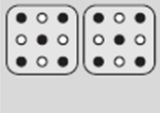


	Mode 1 (3 × 3)	Mode 2 (4 × 4)	Mode 3 (3 × 5)
Number of probes	Light sources: 10 Detectors: 8	Light sources: 8 Detectors: 8	Light sources: 8 Detectors: 7
Probe configuration			
	●: Light sources ○: Detectors		

Figure 3.4: Hitachi probe configurations [Kawaguchi *et al.*, 2001].

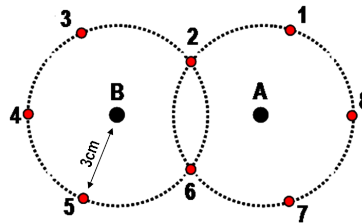


Figure 3.5: Array used by Franceschini *et al* [2000] to probe the motor cortex during a hand tapping protocol. The array has 16 sources (red circles) and two detectors (black circles).

separation between sources and detectors is 30 mm [Kawaguchi *et al.*, 2001]. Two 3×3 probes are used for studies where the activation within both brain hemispheres is compared. The 4×4 probe is used for measurements of the visual cortex or areas on one side of the brain. The 3×5 probe is useful for language studies, where simultaneous measurements of the Broca's area and Wernicke's area is required.

Franceschini *et al* [2000] used two circles of sources with a detector in the centre of each circle (see figure 3.5) to study the adult cortex during hand-tapping. In a different study of the sensorimotor cortex they used two parallelogramic shaped probes attached on each side of the head, containing 8 sources whose minimum distance to each of the 16 detectors was 3.0 cm, allowing 28 source-detector pairs [Franceschini & Boas, 2004]. Although this probe can cover a large area it is relatively sparse and there is only one source-detector separation, which resulted in images with poor spatial resolution.

An auditory study on neonates was performed by Nissila *et al* [2004] using an arc shaped probe consisting of 8 detectors around a laser source, leaving space for the ear. This shape ensures that the optodes are positioned on the assumed location of the auditory cortex.

A study undertaken by Yamamoto *al* [2002] aimed to find the probe configuration that can provide optimum lateral spatial resolution. Three different configurations were tested (refer to figure 3.6): the lattice arrangement (LA), the double-density arrangement (DA), and the quadruple-density arrangement (QA). The distance between a source-detector pair is 30 mm. Only measurements performed at this distance were used in the image reconstruction. The separation between each measuring point (midpoint between a source and detector pair) is 21 mm for LA, 15 mm for DA and 11 mm for QA. The different probes were tested on a phantom that simulated a three layer adult human brain, containing an absorbing object which represented the brain activation, and the DA achieved the highest lateral resolution. Even

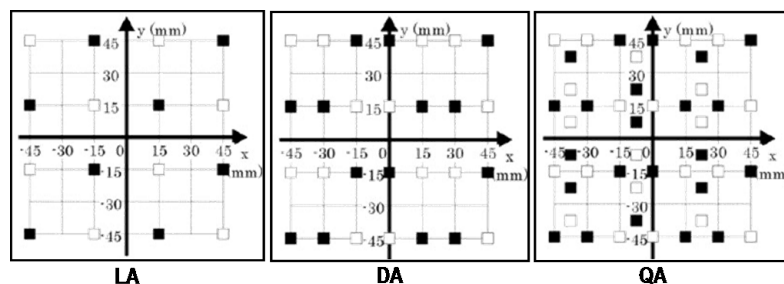


Figure 3.6: Three optode arrangements, where \square are the sources and \blacksquare the detectors: LA - lattice arrangement, DA - double-density arrangement and QA - quadruple-density arrangement [Yamamoto *et al.*, 2002].

though the QA is more dense than the DA it did not improve the lateral resolution of the image. Kawaguchi *et al* [2007] showed through simulations that overlapping measurements by the DA improve the lateral spatial accuracy.

Without overlapping measurements the lateral resolution is limited, since the only information available is that a change occurred somewhere between the source and the detector, whereas using overlapping measurements this ambiguity is reduced. Rectangular and hexagonal geometries with overlapping measurements were analysed by Boas *et al* [2004b]. The first employed 9 sources and 16 detectors (84 measurements), one source surrounded by 4 detectors, and the second 8 sources and 15 detectors (50 measurements), one source in the centre and 6 detectors on each vertex of the hexagon. The longer source-detector separation is 4.25 cm for both geometries and shorter separations are 1.9 cm for the rectangular probe and 2.5 cm for the hexagonal probe. They were evaluated using motor stimulation studies on adults and it was shown that overlapping measurements significantly improve the image lateral resolution and spatial accuracy. The rectangular geometry performs slightly better. Zhao *et al* [2006a] used a multi-centred hexagonal geometry with overlapping measurements (7 sources in the centre) which gave better results than the single-centred and rectangular geometries.

A probe employed by Toronov *et al* [2007] to image the visual cortex uses source-detector distances from approximately 20 to 30 mm, by placing 4 detectors in the centre of a circle containing 16 sources. A disadvantage of this probe is that it only has depth sensitivity in a small region. A different probe, which covers a larger depth range, was used to study the motor cortex, employing 16 sources and 2 detectors with two source-detector distances. A smaller distance (0.5 cm) was used to probe superficial layers and a larger one (2.8 cm) to sample brain cortical tissues [Toronov *et al.*, 2000]. Zeff *et al* [2007] imaged the visual cortex using a high-density (HD) array with 24 source positions and 28 detectors (see figure 3.7). With this array they were able to collect 348 measurement at four different source-detector distances (13, 30, 40 and 48 mm). They obtained images with impressively high contrast-to-noise, lateral resolution, and volumetric localisation. High density optical topography systems are a promising technology. Improvements in their dynamic range will allow the use of larger source-detector separations, which will provide measurements from deeper regions within the brain [Dehghani *et al.*, 2009]. Dehghani *et al* [2009] performed a simulation study where they added a fifth source-detector separation (5.4 cm) to the high-density probe of Zeff *et al* [2007].

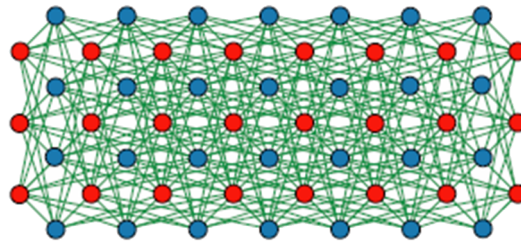


Figure 3.7: High-density array with 24 sources (red) and 28 detectors (blue). The measurement channels are represented by the green lines [Zeff *et al.*, 2007].

They used a realistic head model to analyse the influence on the measurement sensitivity and image reconstruction, of measurements using the first nearest neighbour (NN) source-detector separation and of increasing the maximum separation up to the fifth nearest neighbour. Their modelling predicts that first NN measurements can only interrogate the extracerebral layers. Including second NN measurements provides sensitivity to the outer surface of the brain. Adding other NN increases the distance within the brain at which brain activations can be measured and imaged. With the fifth NN measurements, cerebral activity at depths greater than 20 mm within the brain can be imaged.

Heiskala *et al* [2009] compared three different probe configurations: the hexagonal arrangement, the DA and HD arrangement. They simulated two absorption perturbations at different centre-to-centre separations but at the same depth from the imaging surface, located in the grey matter and outer region of the white matter. The reconstructions from simulated data obtained using the HD arrangement showed the best spatial accuracy and recovery of contrast. They have also shown that image quality can be further improved if mean time data are available.

In conclusion, overlapping measurements improve image resolution and accuracy. Without overlapping measurements it is difficult to identify where absorption changes occur in the brain; i.e. near the source, near the detector or somewhere in between. Depth resolution can also be improved by using different source-detector separations, since we are sampling both superficial and deeper areas of the brain. High density arrays are desired, they can cover a large range of depths and provide images with better quality. However, developments in the technology of the UCL optical topography system and other systems are still required.

3.3.1 UCL topography probe

The array designed for human adult and infant studies has 3 different source-detector distances: 1.43 cm, 1.78 cm and 2.20 cm (see figure 3.8) [Blasi *et al.*, 2007]. We have adapted the software so that each detector can receive signals from all 16 sources, making a total of 64 channels and 15 different source-detector separations. This increases the number of overlapping measurements and gives 16 different source-detector separations, as shown in figure 3.9.

Figure 3.9 shows all the possible channels, where the red lines represent the short distances (the same as in the previous figure) and the blue lines represent larger distances. By using multiple distances we hope to cover a great range of depths, which should improve depth

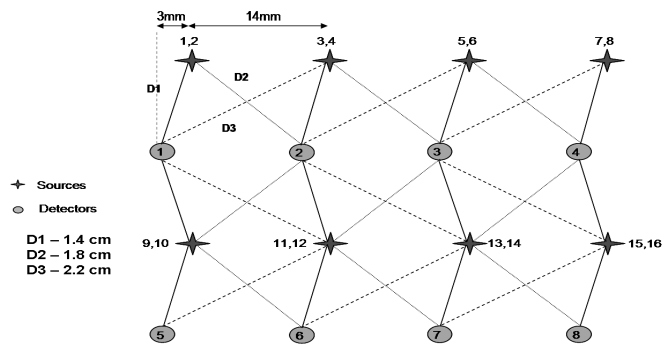


Figure 3.8: Array showing the 16 sources (8 at each wavelength), 8 detectors and 3 source-detector distances. D1-D3 are the source-detector separations.

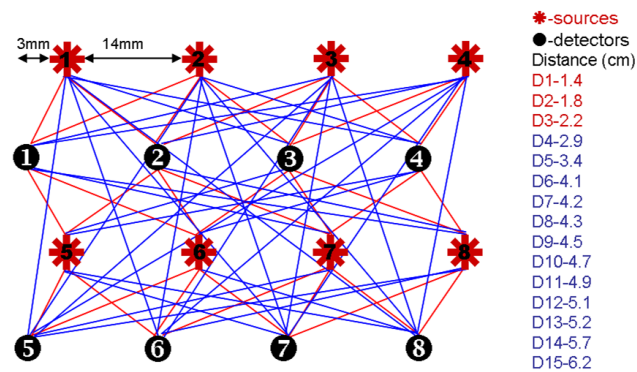


Figure 3.9: Array showing 64 channels. All the source-detector pairs collect data. The red and blue lines represent short and large separations respectively. D1-D15 are all the possible source-detector separations in cm.

resolution and spatial accuracy. The largest separations are unlikely to provide any useful information because of the high attenuation. However, these channels should not have a negative effect on the reconstructed images using the difference imaging approach (refer to §5.2).

3.4 Image reconstruction

Optical imaging of brain activity has potential applications in several clinical areas, for example diagnosis and study of trauma, depression, Alzheimer's, schizophrenia, stroke, haemorrhage, or even tumours. It is also clearly useful for understanding normal brain response to stimuli or mental tasks.

The images generated by the Hitachi systems assume that a change in intensity in a source-detector pair measurement has its origin midway between the source and detector, and there is no attempt to resolve in the depth direction. This assumption limits the lateral spatial resolution to the optode separation [Yamamoto *et al.*, 2002]. Boas *et al* [2001a] used a linear reconstruction method based on the diffusion approximation (refer to §5 for detailed explanation) and obtained more accurate measurements of the changes in haemoglobin concentration than using simple NIRS with the modified Beer-Lambert law (MBLL). The MBLL method often provides underestimated haemoglobin concentration changes, since it assumes that the medium is homogeneous with uniform changes throughout the sampled

volume and scattering is high but constant [Strangman *et al.*, 2002]. Also, the MBLL does not provide a framework for reconstructing images. Overlapping measurements can contribute to the same pixel in the image when linear reconstruction is used, which can increase spatial accuracy and resolution, as mentioned previously [Boas *et al.*, 2004a]. The quality and accuracy of the image can also be improved by including the derivation of the unknown coupling coefficients (refer to §5.2.2 for more detail) as part of the reconstruction process [Boas *et al.*, 2001c].

Optical topography systems detect diffusely reflected light from superficial brain tissue and in most cases 2D maps are generated to represent the haemodynamic changes. For studies of infants heads, light may be transmitted across the head, because the skull is thinner and their head is smaller, giving information about much deeper regions of the brain. Although 2D images can be sufficient for some type of studies, for example when there is only an interest to identify if an haemodynamic change occurs or not, they will not convey any depth information. Pseudo 3D images can be reconstructed by generating individual 2D maps at different depths [Cheng & Boas, 1998]. However, in order to generate true 3D images one has to solve an inverse problem that is largely under-determined in addition to being ill-posed (§5.2.2.1 explains how to deal with this type of problem). The main difficulty of 3D reconstruction is the small number of measurements compared to the large number of pixels in the image to be reconstructed.

Using a multi-distance probe Bluestone *et al* [2001] reconstructed 3D maps of the changes in the concentrations of oxy-haemoglobin, deoxy-haemoglobin, which occurred in an adult forehead during a Valsalva manoeuvre. They used the finite-element method (see §5.1.1.1.4.1) to solve the forward problem and an iterative scheme to reconstruct the images. The optode positions and head surface geometry were determined using photogrammetry. A similar approach was used to image the haemodynamic changes in a rat brain when carbon dioxide was inhaled [Bluestone *et al.*, 2004].

DOT images typically exhibit poor spatial resolution due to the diffuse nature of the photon migration, but the quality of the image can be improved by the use of prior information. One possible way to improve the images is to use MRI images to provide anatomical information. For example, Ntziachristos *et al* [2000] used information about the localisation of a tumour, which was extracted from an MRI image. Boas *et al* [2005] tried a different approach to test spatial accuracy: they performed Monte Carlo (see §5.1.1.2.1 for description of the method) simulations on a human head model and used a linear reconstruction approach, to which they applied a cortical constraint that forces the absorption changes to occur in the cortex, in an attempt to improve depth localisation accuracy. The constraint was obtained by segmentation of a structural MRI image. However, the absorption contrast was always displaced towards the surface of the cortex.

Corlu *et al* [2005] presented an algorithm which reconstructs $\Delta[\text{HbO}_2]$ and $\Delta[\text{HHb}]$ directly, without having to reconstruct absorption coefficient images first, which gives greater separation between absorption and scattering, and between chromophores (refer to §8 for more detail). The reconstructed variables are wavelength independent, and therefore data sets from different measurement wavelengths can be used simultaneously in the image reconstruction.

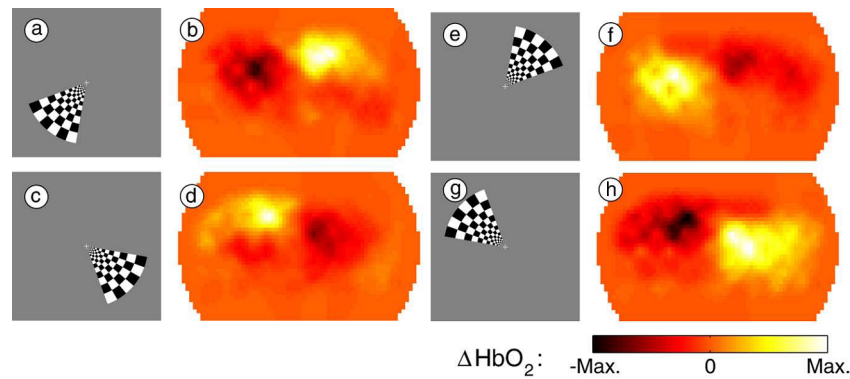


Figure 3.10: Images of activation in the visual cortex obtained using a high-density array with 24 sources and 28 detectors. The visual stimulus is a reversing, black and white radial grid (a,c,e,g). The maximal haemodynamic response appears in the opposite visual quadrant from the stimulus (b,d,f,h)[White & Culver, 2009].

However, the quality of the images is dependent on the choice of wavelengths, i.e., some combinations of wavelengths yield better images than others.

Superb images obtained by Zeff *et al* [2007] and White and Culver [2009] are a clear indicator of the advantages of high density optical topography. High contrast to noise ratio images with high spatial resolution enabled visualisation of retinotopic characteristics of the visual cortex (see figure 3.10), which so far had only been studied by other well established technologies such as PET and fMRI (these technologies are introduced in §4). The optical topography results are in good agreement with those obtained by the other technologies. Not only it is important to have an algorithm capable of producing accurate reconstructions, it is also extremely important to have good quality data.

Custo *et al* [2010] introduced an imaging protocol that consists of using an atlas-based anatomical model of the human head to solve the forward problem (refer to §5) and reconstruct 3D images, eliminating the need for subject specific anatomical MRI. The brain functional images reconstructed using the general anatomic head model showed activity in the same regions as those reconstructed using a subject-specific anatomical model. Therefore, even when an anatomical MRI is not available results can still be interpreted anatomically.

A common practise in optical imaging is to use a so-called difference imaging approach (refer to §5.2), where a data set is acquired before a change occurs and another one during or shortly after. The use of difference measurements reduces the influence of uncertainties in the geometry and of the initial estimate of background properties [Gibson *et al.*, 2005; Bluestone *et al.*, 2001]. However, this approach can only determine changes in the optical properties, and cannot yield absolute values.

Improving depth resolution and spatial accuracy is a complicated problem due to the decrease in sensitivity with depth, and most of the approaches involve using sophisticated image reconstruction algorithms in addition to overlapping measurements. Several methods have been proposed to increase image quality, and each one has its advantages and drawbacks. This thesis addresses some of the issues in image reconstruction of CW optical topography data and aims to provide tools that can help reduce the artefacts, and increase the quality and accuracy of the reconstructed images.

Imaging techniques in medicine

Current medical imaging techniques are briefly reviewed, with particular emphasis on brain imaging. Their strengths and weaknesses are analysed and compared to those of optical topography. Medical imaging can be divided into two main groups: anatomical and functional imaging. Anatomical imaging provides static images of anatomical structures. Functional imaging provides images of physiological changes of tissues.

4.1 X-ray computed tomography

X-ray imaging, known as radiography, records the structure of the body as a two-dimensional (2D) image, without any depth information, i. e., all the body structures superimpose in the final image. Bone structures produce high contrast, but soft tissues show very little contrast. The problem with brain imaging is that the soft tissue is surrounded by the highly absorbing skull, which makes it more difficult to view the low contrast tissues of brain. However, in angiography a contrast agent is injected into the bloodstream, which increases the contrast of blood vessels, making them visible on x-ray images.

X-ray computed tomography (CT) overcomes these contrast and depth problems [Michael, 2001]. It provides a series of 2D thin slices of the body by rotating the source of x-rays around the subject to obtain multiple projections. First generation CT scanners involved translating and rotating a narrow beam of x-rays about the subject. This was a very slow process, where data acquisition of each section could take several minutes. Nowadays, most CT scanner are of fourth generation. In fourth generation CT scanners a fan beam of x-rays is rotated around the patient and is detected by a ring of discrete detectors placed around the subject (figure 4.1). Complete sectional scans take only few seconds. The fifth generation CT scanners provide scan times of only few milliseconds. Instead of physically rotating the source, an electron beam is swept across a semicircular anode ring surrounding the subject, producing x-rays (figure 4.2).

The images are no longer obtained directly, as in radiography. Instead, CT requires image reconstruction algorithms that generate values of attenuation for each pixel of the image. One of the most common image reconstruction methods is called filtered backprojection. This method

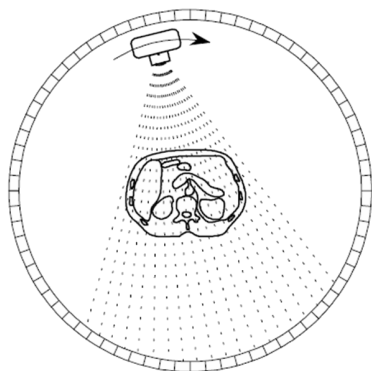


Figure 4.1: Fourth generation CT scanner geometry [Michael, 2001].

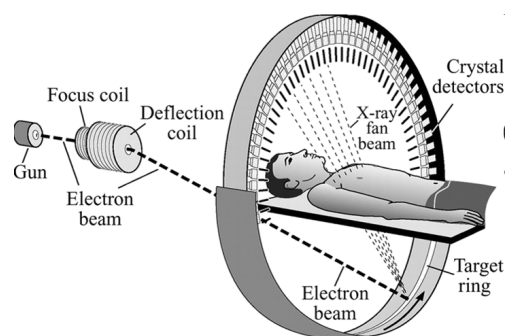


Figure 4.2: Fifth generation CT scanner [Wolbarst & Hendee, 2006].

is based on the simple backprojection, where each attenuation profile, recorded for each viewing angle, is projected across the image space and added together. However, the image formed is highly artefactual with degraded contrast. Convolving the attenuation profiles with a filter kernel before backprojection produces more accurate images. An advantage of this method is that the projections are processed individually, therefore, the image reconstruction can start before all the data has been acquired. Once the attenuation coefficients have been calculated for each pixel in the image, these values are converted to CT numbers, which are calculated relative to the value for water. Therefore, the CT number of water is zero. In order to enhance the contrast of the images, a limited range of CT numbers can be selected, depending on the type of tissue of interest.

CT is a non-invasive technology that provides high resolution anatomical images. Applications of CT brain imaging include detection of tumours [Conry & Reznick, 1987], hemorrhage [Brott *et al.*, 1997; Zimmermana *et al.*, 2006] and stroke [Mayer *et al.*, 2000]. Figure 4.3 shows a CT scan of an adult brain suffering from intracranial hemorrhage (ICH), where region 1 defines zones of ICH and regions 2 defines zones of edema (induced by ICH). The disadvantages of CT are the relative high dose of ionising radiation, and the high cost of the scan. Some patients have shown allergic reactions to certain contrast agents. CT scans should not be done on children and pregnant women.

4.2 Ultrasound

In Ultrasound (US) imaging a pulse of high frequency sound waves is emitted, typically from 1 to 10 MHz, and travels through the patients body until it meets boundaries between different structures in the body and is reflected [Pope, 1999]. Reflection occurs due to differences in the acoustic impedance Z of tissues, which is defined as

$$Z = \rho c, \quad (4.1)$$

where ρ is the tissue density and c is the speed of sound in the same tissue. If the mismatch between acoustic impedances is large, then the pulse is almost entirely reflected, whereas if they are similar, then the pulse is mostly transmitted across the boundary. The fraction of intensity

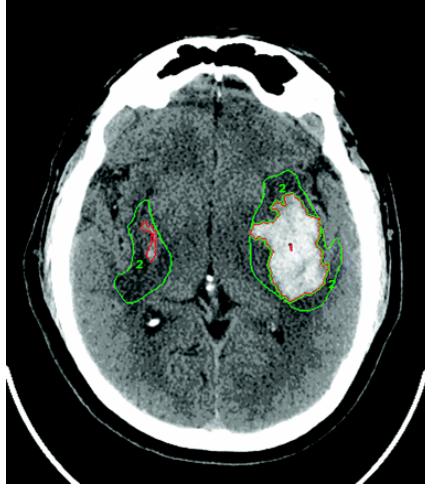


Figure 4.3: Computer tomography scan of an adult brain showing hemorrhages (1) and edema (2) [Zimmermana *et al.*, 2006].

reflected back, when the pulse is incident perpendicular to the boundary between two media, is given by

$$\alpha = \left(\frac{Z_2 - Z_1}{Z_2 + Z_1} \right)^2, \quad (4.2)$$

where Z_1 and Z_2 are the acoustic impedances of the media. The depth of the structure can be found using the time delay t between emission of the pulse and detection of the echo

$$d = \frac{ct}{2}. \quad (4.3)$$

The ultrasound pulse is produced and detected by a transducer. This device converts an electrical signal to sound waves and vice-versa, through the piezoelectric effect. In order to generate 2D images, the transducer can be physically moved or, alternatively, an array of transducers can be used, where all the pulses are produced simultaneously, or successively, producing a steered signal.

Doppler ultrasound systems can detect moving structures, by measuring the frequency changes of the reflected waves. It can be used to measure the velocity of blood flow through vessels.

Ultrasound is widely used in obstetrics, in abdominal studies, breast examinations and cardiology [McGahan & Goldberg, 2008]. Ultrasound of the neonatal brain can access the brain via the soft fontanelles. Neonatal brain ultrasound is a useful diagnostic tool for cerebral disorders, such as hemorrhage, leukomalacia, meningitis and congenital anomalies [Ancel *et al.*, 2006; Yikilmaz & Taylor, 2008; Schenk *et al.*, 2006]. However, ultrasound has limited use for imaging the adult brain due to the high impedance of the skull relative to the surrounding tissues. Transcranial duplex ultrasound, which combines the anatomical and doppler measurements, allows the examination of the neurovascular system of adults [Seidel *et al.*, 2000, 2004]. It uses the harmonic technology, based on the principle that harmonic frequencies of the original ultrasound signal are generated inside the tissue,

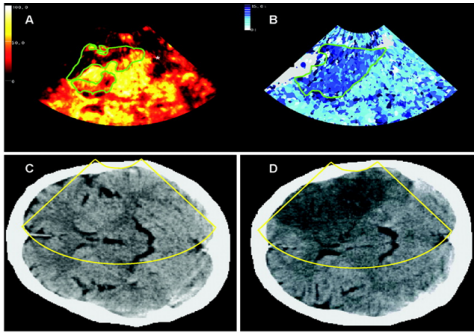


Figure 4.4: Harmonic imaging ultrasound (top) and cranial computed tomography (CT) (bottom) for comparison [Seidel *et al.*, 2004]. (A) Pixel-wise peak intensity image. (B) Time to peak intensity image. (C) CT scan 3 hours after symptom onset. (D) CT scan 30 hours after symptom onset.

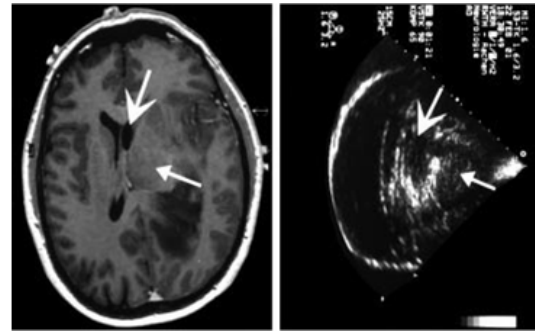


Figure 4.5: Ultrasound scan (right) of a cerebral tumour (small arrow) located near the lateral ventricle (large ventricle) [Harrer *et al.*, 2003]. Magnetic resonance scan of the tumour for comparison (left).

hence the harmonic signals only need to pass through the skull once before being detected [Hedrick & Metzger, 2005]. Penetration depth at low frequencies is approximately 10 cm, which is achieved at the expense of spatial resolution. This technique requires administration of an US contrast agent and imaging is limited to regions where the skull is thinner. US is also used in neurosurgery to identify tumours. Harmonic imaging after administration of an US contrast agent has been used to identify cerebral acute ischemic (insufficient blood flow) stroke. Figure 4.4 shows a colour-coded duplex US of a patient with middle cerebral artery occlusion [Seidel *et al.*, 2004]. The first image represents the pixel-wise peak intensity, where every pixel is set to the peak intensity within the data set. The second image is the time to peak intensity, in which each pixel is set to a value that represents the delay between the start and onset of the contrast enhancement. A similar technique has been used to identify different types of tumours in adult brains. An example can be seen in figure 4.5 [Harrer *et al.*, 2003].

The main advantages of US are its portability, moderate cost, mobility, the use of non-ionising energy and it can provide real-time images. The disadvantages are the strong reflection when passing through bone or air, the decrease in the image resolution with depth, and the potential risk that high frequencies can induce temperature rises that can cause tissue damage.

4.3 Magnetic Resonance Imaging

Magnetic Resonance Imaging (MRI) is a non-invasive imaging modality, with a high soft-tissue contrast. The patient is placed inside a strong magnetic field, which causes the randomly oriented magnetic moments of hydrogen nuclei (protons) to adopt one of the two spin states, such that they precess around the magnetic field vector \vec{B}_0 . The principles of MRI or Nuclear Magnetic Resonance (NMR), as it used to be referred to, are described by Mansfield and Morris [Mansfield & Morris, 1982].

In more detail, the nucleus of atoms with odd number of protons or odd number of neutrons have a property known as spin. As a consequence, the spinning charged particle generates a magnetic field. For example, hydrogen is quite abundant in the human body, and

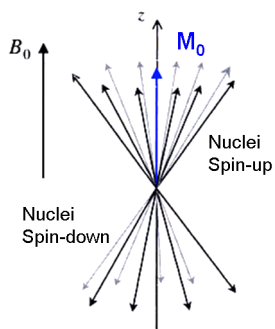


Figure 4.6: External magnetic field \vec{B}_0 produces the net magnetisation \vec{M}_0 in the direction of the magnetic field [McRobbie & E A Moore, 2007].

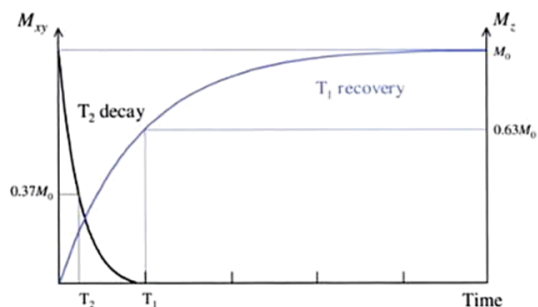


Figure 4.7: T1 and T2 relaxation times after a 90° RF pulse. T2 is much quicker than T1 [McRobbie & E A Moore, 2007].

the nucleus has an odd number of protons, hence the nucleus has a half-integer spin, which is $\frac{1}{2}$. If \vec{B}_0 is applied, then the hydrogen nuclei in the human body can assume two energy levels: spin-up (lower energy, oriented with a component parallel to the field) and spin-down (higher energy, oriented with a component antiparallel to the field). According to the Boltzmann distribution, slightly more nuclei will adopt the lower energy state (one in a million), so that the net magnetisation \vec{M}_0 points in the direction of the magnetic field (see figure 4.6). The individual proton magnetic moments also precess about \vec{B}_0 . The frequency of precession is known as the Larmor frequency, and is given by the equation

$$\omega_0 = \gamma \vec{B}_0, \quad (4.4)$$

where γ is the gyromagnetic ratio. For the hydrogen, $\gamma = 42.6$ MHz/Tesla, hence, in a magnetic field $\vec{B}_0 = 1$ Tesla the precession frequency of the proton is 42.6 MHz. Even though the individual protons precess, the net magnetisation \vec{M}_0 only has a component aligned with \vec{B}_0 , and therefore it does not precess.

If a radio frequency (RF) pulse is applied, whose frequency is ω_0 and magnetic field vector is perpendicular to \vec{B}_0 , the individual magnetic moments can precess in phase with each other, and \vec{M}_0 is given a component perpendicular to \vec{B}_0 . At the same time spins in the lower energy spin-up state flip to the higher energy spin-down state, causing the component of \vec{M}_0 parallel to \vec{B}_0 to decrease. This is called resonance. When the RF pulse is turned off the spins begin to dephase, and return to their original distribution of spin states. This recovery to the original state is characterised by the relaxation times T1 and T2 (figure 4.7). The first is called the longitudinal relaxation time and is related to the time it takes for the spin states to return to their initial distribution (figure 4.8). The relaxation time T2, or transverse relaxation time, characterises the time it takes for the spins to precess out of phase with each other (figure 4.8). However, due to inhomogeneities in the magnetic field \vec{B}_0 , the dephasing occurs faster than T2 and the term T2* is often used. Different tissues of the brain have different relaxation times.

Applying RF pulses with magnetic field gradients across each of the three orthogonal directions, provides spatially encoded data. The most common method uses a longitudinal gradient, where each slice along this direction has a different resonance frequency, therefore the

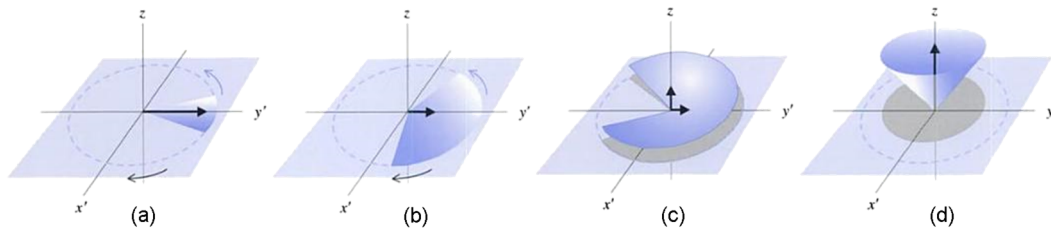


Figure 4.8: After the 90° RF pulse, T2 relaxation causes rapid decrease in the transverse magnetisation (a) and (b). At the same time, T1 causes a slower increase in the longitudinal magnetisation \vec{M}_0 (c) and (d) [McRobbie & E A Moore, 2007].

RF pulse with the same frequency only rotates the spins of a single slice. Therefore, the spins of the selected plane have the same precession frequency. Subsequently, a phase encoding gradient is applied across an orthogonal direction for a short period. When the gradient is turned on the spins along the direction of the gradient precess at different frequencies. When the gradient is turned off the spins have the same precession frequency but different phases, hence the name of the gradient. Immediately after a frequency encoding gradient is applied across the remaining direction. Each pixel of the selected plane has now a characteristic phase and frequency. A 2D Fourier transform of the plane can provide the location of the spins.

The magnetic properties of haemoglobin depend on its oxygenation level. Oxy-haemoglobin is diamagnetic, like most tissues in the brain, meaning that it develops a magnetic field opposite to the applied field \vec{B}_0 . Meanwhile, deoxy-haemoglobin is paramagnetic, meaning that it creates a magnetic field aligned with \vec{B}_0 . Functional MRI (fMRI) measures the effect of the perturbations in magnetic field produced by deoxy-haemoglobin on the surrounding protons, known as the Blood Oxygen Level Dependent (BOLD) effect. The presence of deoxy-haemoglobin produces a decrease in the relaxation time T2. If a subject being studied is presented with a stimulus, in the stimulated region of the brain an increase in the blood flow is observed, bringing more oxy-haemoglobin and removing deoxy-haemoglobin, thus T2 increases locally.

Matthews *et al* [1999] describe several application of fMRI. For example, this technique has been used to study regions of the brain related to motor, language and visual tasks. It is also used to diagnose brain disorders and for function neurosurgery, where the regions to be treated are stimulated. Figure 4.9 shows two fMRI brain maps, of the left and right hemisphere, showing significant activation of regions of the medial prefrontal cortex, posterior superior temporal sulcus, anterior cingulate cortex and inferior frontal gyrus [Pelphrey *et al.*, 2005]. The subjects were submitted to social perception stimuli, such as eye, mouth and hand movements. This study pays particular attention to the pSTS; mouth movements are located along the mid-posterior pSTS, while hand movements activate inferior and posterior regions and eye movements activate superior and posterior regions of the right pSTS.

The advantages of MRI are its high spatial resolution, the use non-ionising radiation and non-invasiveness. The disadvantages are associated with the presence of strong magnetic fields, the high cost of the equipment, its large dimensions and the requirement for the patient to stay still for relatively long periods of time, which is a particularly difficult task for infants and

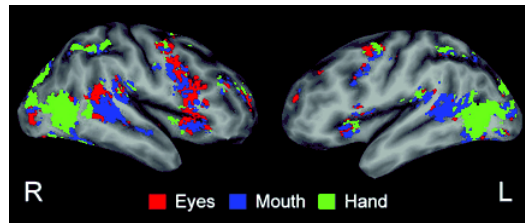


Figure 4.9: fMRI activation maps showing the regions of greatest response to social stimuli: eye, mouth and hand movements [Pelphrey *et al.*, 2005].

newborns (who often require the administration of anesthesia or sedatives).

4.4 Radionuclide Imaging

In radionuclide imaging a radioactive tracer is injected into the body and becomes involved in the chemical and physiological processes. The radionuclide decays and produces detectable γ -rays, which provide a measurement of the concentration of the tracer in the body [Cherry *et al.*, 2003]. Several different types of radionuclides are available that can trace different regions of the body, making this technique very versatile, since it can image the anatomy and function of different tissues and organs.

4.4.1 Single Photon Emission Computed Tomography

In Single Photon Emission Computed Tomography (SPECT) a gamma camera is rotated around the subject, and is used to record a series of projection profiles that can be combined to generate a 3D image [Powsner & Powsner, 2006]. The filtered backprojection method, described previously, can then be used to reconstruct sections of the body. Multiple projection views can be acquired simultaneously by using multi-headed gamma cameras, which accelerates the data acquisition. Photons originating in deeper parts of the body are more likely to be attenuated than those at the surface. Therefore, a correction using a sensitivity matrix is applied to the algorithm assuming that attenuation is uniform, which is more or less accurate for brain tissue.

Technetium-99m (Tc-99m) is a radioactive tracer that is widely available, with a short half-life of approximately 6 hours, and is used for functional brain imaging to measure blood flow. Generators that contain molybdenum-99 (Mo-99), called the parent nuclide, decay to produce Tc-99m. Mo-99 has a relatively long half-life (approximately 67 hours) and is produced in nuclear reactors.

4.4.2 Positron Emission Tomography

Positron Emission Tomography (PET) radionuclides decay emitting a positron, which travels few millimetres in the tissue until it finds an electron and annihilates, producing a pair of γ -rays (each with energy 511 keV) that travel in opposite directions [Powsner & Powsner, 2006]. The subject is surrounded by a ring of detectors and only coincident photons are used to generate images. Therefore, the annihilation must have occurred somewhere between the line that connects the two detectors that register the event. Image reconstruction can be performed using filtered backprojection and applying the attenuation correction. The resolution can be increased by measuring the time of flight of the coincident γ -rays, which informs where along the line between the detectors the annihilation occurred. Nevertheless, the resolution of PET images is

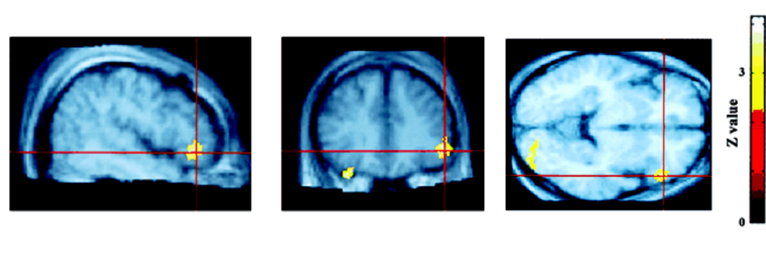


Figure 4.10: PET statistical maps showing regions involved in the assessment of facial emotion. Left: sagittal; Centre: coronal; Right: transverse images [Nakamura *et al.*, 1999].

limited by the distance travelled by the positron before it annihilates. Figure 4.10 shows a PET statistical map of significant activation in the right inferior frontal cortex, during the assessment of facial emotion [Nakamura *et al.*, 1999]. The activated areas denote a significantly increased CBF, which were then superimposed onto a mean MRI image produced from all the subjects.

It is necessary to have a cyclotron located near the imaging facility in order to produce the short half-life radionuclides used in PET - half-lives of the order of few minutes to a couple of hours. Fludeoxyglucose (F-18) is a tracer similar to glucose that is used to monitor brain functional activity. When a certain region of the brain is more active more glucose is consumed, and therefore, more radioactive glucose will also be consumed. Oxygen-15 (half-life ~ 2 min) labelled water or Carbon-11 (half-life ~ 20 min) labelled carboxy-haemoglobin are used to measure cerebral blood flow.

The main disadvantages of PET and SPECT are their high cost and the use of radioactive substances.

4.5 Electrical Impedance Tomography

Electrical Impedance Tomography (EIT) produces images of the conductivity (σ) of tissues. Different types of brain tissue have different conductivity values. For example, the conductivity of grey matter is 3.5 mS/cm and of white matter is 1.5 mS/cm [Bronzino, 2000].

Electrodes placed at the surface of the body apply a small alternating current into the tissue, which generates electrical potentials that are measured at the surface. The most common electrode driving arrangement is the adjacent pattern, where the current is injected through two adjacent electrodes, and the voltages are measured by the remaining pairs of adjacent electrodes (figure 4.11). The process is repeated for all the electrode pairs [Graham & Adler, 2007].

In order to obtain maps representing the changes in conductivity of different brain tissues, from measurements of voltage changes, it is necessary to solve an ill-posed inverse problem. The type of problem and the approaches used in the image reconstruction are very similar to those used in optical tomography (see §5).

EIT has been used to measure functional brain activity in the presence of a stimulus [Tidswell *et al.*, 2001] and localise regions of the brain responsible for epileptic seizures [Fabrizi *et al.*, 2006]. Changes in conductivity can occur due to changes in blood volume. Figure 4.12(a) shows a conductivity increase in the visual cortex while the subject performed a visual task [Bagshaw *et al.*, 2003]. Similarly, figure 4.12(b) shows the response in the motor cortex due to motor stimuli [Bagshaw *et al.*, 2003].

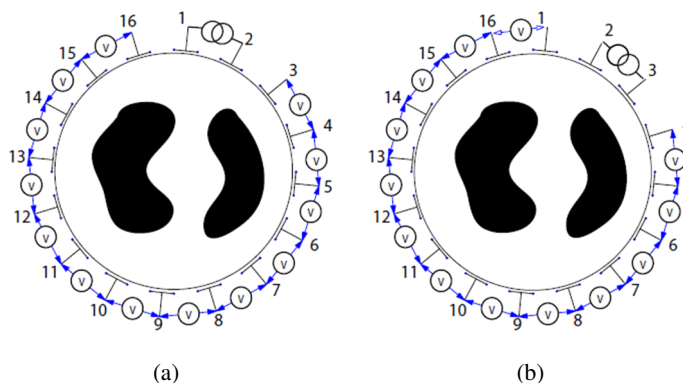


Figure 4.11: EIT adjacent drive pattern [Graham & Adler, 2007]. (a) Current is injected through electrode pair (1,2) and measured by the remaining pairs, except pairs (16,1) and (2,3). (b) Current is injected through electrode pair (2,3) and voltage differences are measured by the remaining pairs, except pairs (1,2) and (3,4).

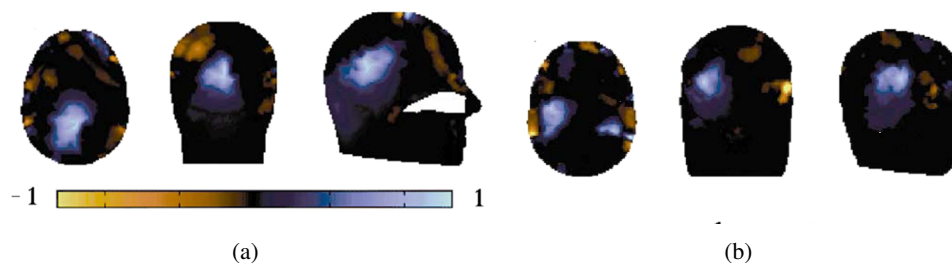


Figure 4.12: EIT transverse, coronal and sagittal images of the response to (a) visual and (b) motor stimuli [Bagshaw *et al.*, 2003].

This imaging modality produces images with relatively low spatial resolution, but it has a high temporal resolution and is based on inexpensive technology.

4.6 Comparison of brain imaging modalities

The main characteristics of the medical imaging modalities described in this chapter and optical topography (OT) using CW measurements are summarised in table 4.1. The values are estimates extracted from different literature sources. Spatial resolution is the smallest distance between two separate objects that can be successfully resolved by the imaging system. Temporal resolution reflects the ability of the imaging system to resolve successive events in time, and is effectively the inverse of the frame rate. However, note that in some cases, spatial resolution can be increased at the expense of temporal resolution and vice versa.

Table 4.1: Main characteristics of medical imaging modalities. Cost: \$ - < £300k, \$\$ - £300k – 1 million, \$\$\$ - > £1million.

modality	Spatial resolution	Temporal resolution	Type of information	advantages	disadvantages	cost
CT	0.5-1 mm	>50ms scan:5-15 min	<ul style="list-style-type: none"> Brain anatomy Blood vessels: angiography hemorrhage tumours, etc 	<ul style="list-style-type: none"> can image bone, blood and soft tissue simultaneously not very sensitive to motion availability 	<ul style="list-style-type: none"> ionising radiation may use contrast agents (possible allergic reaction) cannot differentiate tissues with similar density bulky equipment children and pregnant women should not be scanned 	\$\$
US	1-2 mm	20ms	<ul style="list-style-type: none"> Brain anatomy blood velocity: Doppler US hemorrhage blood vessel anomalies 	<ul style="list-style-type: none"> no ionising radiation real time imaging not sensitive to motion availability portable 	<ul style="list-style-type: none"> unsuitable when tissue densities are very different bones are obstacles limited for adult brain imaging newer US techniques, provide better resolution but require contrast agents 	\$
MRI	0.1-1 mm	not applicable scan:10-20 min	<ul style="list-style-type: none"> Brain anatomy Blood vessels: MR Angiogram tumours blood vessel anomalies, etc 	<ul style="list-style-type: none"> no ionising radiation high soft-tissue contrast cross-sectional views in any plane 	<ul style="list-style-type: none"> bulky equipment patient is required to stay still for long periods infants need to be sedated strong magnetic fields: no pacemakers or metal body parts claustrophobic may require contrast agents 	\$\$\$
fMRI	3-5 mm	1-4s	<ul style="list-style-type: none"> Brain function BOLD: deoxy-haemoglobin 	<ul style="list-style-type: none"> same as MRI 	<ul style="list-style-type: none"> same as MRI 	\$\$\$
SPECT	8-9 mm	10-50s scan:10-50 min	<ul style="list-style-type: none"> cerebral hemodynamics 	<ul style="list-style-type: none"> relative availability of radionuclides different cross-sectional views availability (more than PET) 	<ul style="list-style-type: none"> use of ionising radiation wait for tracer distribution bulky equipment 	\$\$
PET	5-8 mm	10-50s scan:10-50 min	<ul style="list-style-type: none"> brain metabolism cerebral hemodynamics 	<ul style="list-style-type: none"> different cross-sectional views 	<ul style="list-style-type: none"> same as SPECT expensive cyclotron on-site availability 	\$\$\$
EIT	1-3 cm	<1s	<ul style="list-style-type: none"> brain function 	<ul style="list-style-type: none"> no ionising radiation portability 	<ul style="list-style-type: none"> low sensitivity to deeper regions no anatomical information 	\$
OT	1-3 cm	<1s	<ul style="list-style-type: none"> brain function oxy-, deoxy- and total haemoglobin 	<ul style="list-style-type: none"> no ionising radiation portability 	<ul style="list-style-type: none"> sensitivity decreases with depth no anatomical information 	\$

Imaging modalities that provide structural information include x-rays CT, US and MRI. Functional imaging techniques include fMRI, SPECT, PET, EIT and OT. The structural imaging modalities like CT and MRI have the advantage of generating images with high spatial resolution. MRI can provide images with higher soft tissue contrast than CT, but it is also more expensive, which makes it less available. However, CT has the disadvantage of using ionising radiation. Contrast agents can be used to enhance contrast of blood vessels. However, these modalities do not provide functional information and involve large, expensive and immobile equipment.

US has very good temporal resolution but low spatial resolution. This technique is inexpensive and widely available. Imaging of the adult brain is confined to regions where the skull is thinner and requires the administration of US contrast agents. Larger field of views have lower spatial resolution.

Anatomical MRI images can be combined with functional information from fMRI acquired in the same session. fMRI can provide images of the BOLD signal and changes in cerebral blood flow (with contrast agent) with a good temporal and spatial resolution. PET and SPECT can also provide metabolic and functional measurement of cerebral activity, but with lower temporal and spatial resolution, but require exposure to radioactive substances. PET has better spatial resolution than SPECT, but it is more expensive since it requires a cyclotron on-site to produce the short half-life radionuclides used by this technique.

EIT and OT have high temporal resolution but limited spatial resolution. The image reconstruction techniques are very similar. They use inexpensive and portable technologies, and do not require immobilisation of the subject, like in PET and MRI, since the probe is attached to the head. Both modalities have low sensitivity to deeper regions below the surface, hence imaging is mostly confined to the cerebral cortex. EIT provides measurements of changes in cerebral blood volume, whereas OT provides simultaneous absolute or relative measurements of CBF, CBV, oxy-haemoglobin and deoxy-haemoglobin concentration. By selecting appropriate wavelengths OT can provide information about other wavelengths-dependent chromophore concentrations. This is also an advantage over fMRI which can only provide measures of deoxy-haemoglobin content without any information about oxy-haemoglobin.

Due to its low spatial resolution and depth sensitivity, optical imaging cannot replace anatomical imaging methods, such as MRI and CT. However, it can provide complementary functional information, which can be combined with images provided by these high resolution structural imaging modalities.

Image reconstruction

The image reconstruction process in optical topography aims to retrieve maps of changes in optical properties and chromophore concentrations occurring in the medium under study, based on diffuse reflection measurements performed at the surface. In order to do so, it is necessary to model the propagation of light, from a given light source, through a medium with specific optical properties, until it reaches a detection point. This is the forward problem, which is commonly described by the diffusion approximation. The inverse problem attempts to recover the internal optical properties from measured data, through the inversion of the forward problem. However, this is an ill-posed and under-determined problem, and regularisation techniques are required to obtain stable solutions. The changes in the chromophore concentrations can be calculated from the changes in optical properties, or reconstructed directly using the multispectral method. This chapter describes all the steps involved in reconstruction of images from optical topography data.

5.1 Forward problem

5.1.1 Radiative transfer equation

A commonly used model of photon propagation through tissue is the diffusion model [Boas *et al.*, 2001b; Arridge, 1999]. This model can only be applied when scattering is much higher than absorption, and consequently the photons undergo a random walk. The diffusion equation is derived from the radiation transport equation (RTE), which describes the flow of energy through a medium. This equation represents the balance of energy between the incident, outgoing, absorbed and scattered photons within the medium. The time-dependent RTE is given by

$$\frac{1}{c} \frac{\partial \phi(r, t, \hat{s})}{\partial t} + \hat{s} \cdot \nabla \phi(r, t, \hat{s}) = -(\mu_a + \mu_s) \phi(r, t, \hat{s}) + \mu_s \int_{4\pi} p(\hat{s}, \hat{s}') \phi(r, t, \hat{s}') d^2 \hat{s}' + q(r, t, \hat{s}), \quad (5.1)$$

where $\phi(r, t, \hat{s})$ is the radiance (energy per unit time, per unit solid angle and through a unit area) at position r , travelling in direction \hat{s} and at time t , $p(\hat{s}, \hat{s}')$ is the scattering phase function, which represents the probability that a photon travelling in direction \hat{s} is scattered into direction \hat{s}' (see §2), q is the radiation source and c is the speed of light (in a specific medium it is divided by its refractive index) [Arridge, 1999; Kaltenbach & Kaschke, 1993]. For a small volume (δV) around the point r , the first term on the left-hand side represents the radiance change with time of photons travelling in direction \hat{s} through δV , and the second term represents the change due to energy flow. The first term on the right-hand side represents the radiance loss owing to absorption and scattering, the second term represents the gain in radiance due to scattering from all directions into direction \hat{s} and the last term is the light source distribution introduced at position r per unit time, per unit volume, per unit area in direction \hat{s} .

The RTE is computationally demanding due to the directional dependence of its terms. Expanding the terms that depend on direction \hat{s} into spherical harmonics yields an infinite series of equations, and truncating the expansion to the first N term gives the P_N approximation to the RTE.

5.1.1.1 Diffusion equation

The diffusion equation can be obtained using the first-order approximation P_1 to the RTE. Derivation of the diffusion equation is explained in more detail by Arridge [1999]. Expanding the \hat{s} direction dependent terms $\phi(r, t, \hat{s})$ and $q(r, t, \hat{s})$ into spherical harmonics $Y_{l,m}$ leads to:

$$\phi(r, t, \hat{s}) = \sum_{l=0}^{\infty} \sum_{m=-l}^l \left(\frac{2l+1}{4\pi} \right)^{\frac{1}{2}} \phi_{l,m}(r, t) Y_{l,m}(\hat{s}) \quad (5.2)$$

$$q(r, t, \hat{s}) = \sum_{l=0}^{\infty} \sum_{m=-l}^l \left(\frac{2l+1}{4\pi} \right)^{\frac{1}{2}} q_{l,m}(r, t) Y_{l,m}(\hat{s}). \quad (5.3)$$

In the diffusion approximation the terms $\phi_{l,m}, q_{l,m}=0$ for $l > 1$. Expressing \hat{s} in terms of spherical harmonics and using the photon density Φ , which is

$$\Phi(r, t) = \int_{4\pi} \phi(r, t, \hat{s}) d\hat{s}, \quad (5.4)$$

and the photon current J , which is

$$J(r, t) = \int_{4\pi} \hat{s} \phi(r, t, \hat{s}) d\hat{s}, \quad (5.5)$$

the approximation is obtained

$$\phi(r, t, \hat{s}) \approx \frac{1}{4\pi} \Phi(r, t) + \frac{3}{4\pi} \hat{s} \cdot J(r, t). \quad (5.6)$$

Similarly, for the source term the approximation leads to

$$q(r, t, \hat{s}) \approx \frac{1}{4\pi} q_0(r, t) + \frac{3}{4\pi} \hat{s} \cdot q_1(r, t), \quad (5.7)$$

where q_0 represents the isotropic component of the source

$$q_0(r, t) = \int_{4\pi} q(r, t, \hat{s}) d\hat{s}, \quad (5.8)$$

and q_1 the dipole component

$$q_1(r, t) = \int_{4\pi} \hat{s} \cdot q(r, t, \hat{s}) d\hat{s}. \quad (5.9)$$

Substituting equations (5.6) and (5.7) in the RTE, equation (5.1), and integrating over all angles, leads to the continuity equation

$$\left(\frac{1}{c} \frac{\partial}{\partial t} + \mu_a \right) \Phi(r, t) + \nabla \cdot J(r, t) = q_0(r, t). \quad (5.10)$$

Multiplying both sides of the RTE by \hat{s} and integrating over all angles gives

$$\left(\frac{1}{c} \frac{\partial}{\partial t} + \mu_a + \mu'_s \right) J(r, t) + \frac{1}{3} \nabla \Phi(r, t) = q_1(r, t). \quad (5.11)$$

Equations (5.10) and (5.11) can be further simplified into a single equation that contains $\Phi(r, t)$ only. This is done by assuming that the source is isotropic, $q_1(r, t) = 0$ and that the photon current J changes slowly, $\frac{\partial J(r, t)}{\partial t} = 0$. Equation (5.11) reduces to Fick's Law

$$J(r, t) = -\kappa(r) \nabla \Phi(r, t), \quad (5.12)$$

where the diffusion coefficient κ is defined as

$$\kappa(r) = \frac{1}{3(\mu_a(r) + \mu'_s(r))}. \quad (5.13)$$

The diffusion approximation is obtained by substituting equation (5.12) into the continuity equation (5.10). The time-dependent diffusion equation is then given by

$$-\nabla \cdot \kappa(r) \nabla \Phi(r, t) + \mu_a(r) \Phi(r, t) + \frac{1}{c} \frac{\partial \Phi(r, t)}{\partial t} = q_0(r, t). \quad (5.14)$$

The assumption $\frac{\partial J(r, t)}{\partial t} = 0$ is valid when

$$\left(\frac{\ell'}{c} \right) \left(\frac{1}{J(r, t)} \frac{\partial J(r, t)}{\partial t} \right) \ll 1, \quad (5.15)$$

where $\ell' = \frac{1}{\mu_a + \mu'_s}$ is the transport mean free path. This means that the fractional change in the photon current over the time it takes for photons to travel ℓ' is quite small [Wang & Wu, 2007]. Equation (5.15) can be written as

$$\frac{1}{c} \frac{\partial J(r, t)}{\partial t} \ll J(r, t) (\mu_a + \mu'_s). \quad (5.16)$$

The P_1 approximation considers the radiance as almost isotropic, which results in an isotropic photon density, which is achieved in highly scattering media, $\mu_a \ll \mu'_s$. Therefore, collimated sources cannot be modelled using the diffusion equation since it will not be accurate near the boundaries, since a minimum travelling length ℓ' is necessary before entering the diffusion

regime. However, boundary and source conditions can help improve the accuracy in these areas.

Fourier transforming equation (5.14) gives the frequency-domain diffusion equation

$$-\nabla \cdot \kappa(r) \nabla \Phi(r, \omega) + \mu_a(r) \Phi(r, \omega) + \frac{i\omega}{c} \Phi(r, \omega) = q_0(r, \omega). \quad (5.17)$$

Setting the modulation frequency ω to zero provides the time-independent form of the diffusion equation, which is described as

$$-\nabla \cdot \kappa(r) \nabla \Phi(r) + \mu_a(r) \Phi(r) = q_0(r). \quad (5.18)$$

5.1.1.1.1 Boundary conditions

The solution of the diffusion equation in a domain Ω requires the use of boundary conditions applied to the boundary $\partial\Omega$, where the sources s_i and the detectors d_i are placed. Different types of conditions can be applied. One the simplest and most common is the Dirichlet boundary condition [Schweiger *et al.*, 1995]

$$\Phi(r, t) = 0, \quad \forall r \in \partial\Omega, \quad (5.19)$$

hence all the photons travelling across $\partial\Omega$ will be absorbed. Therefore, it is equivalent to surrounding the domain Ω with a perfect absorber.

For a non-absorbing surrounding medium a more adequate boundary condition can be applied, which requires that photons from outside Ω cannot travel into the domain Ω , except at source positions s_i , that is,

$$\phi(r, t, \hat{s}) = 0, \quad \forall r \in \partial\Omega, \hat{s} \cdot \hat{n} < 0, \quad (5.20)$$

where \hat{n} is the unit normal vector on $\partial\Omega$ pointing outward [Arridge, 1999; Schweiger *et al.*, 1995]. However, the diffusion equation cannot fully satisfy this boundary condition, since it deals with the integral of radiance over all directions and not only a single direction. Instead, the total photon current flowing into the domain Ω is set to zero

$$\int_{\hat{s} \cdot \hat{n} < 0} \hat{s} \phi(r, t, \hat{s}) d\hat{s} = 0, \quad \forall r \in \partial\Omega, \quad (5.21)$$

which, within the diffusion approximation, leads to the Robin boundary condition given by [Arridge, 1999; Schweiger *et al.*, 1995]

$$\Phi(r, t) + 2\kappa(r) \frac{\partial \Phi(r, t)}{\partial \hat{n}} = 0, \quad \forall r \in \partial\Omega. \quad (5.22)$$

This is equivalent to a non-scattering surrounding medium, hence, all the diffusive sources are inside Ω . The Robin boundary condition is a linear combination of a Dirichlet condition (first term) and Neumann condition (second term). If the surrounding medium has a different refractive index than the domain Ω , it is necessary to take into account the reflection of light back into the medium Ω that occurs at the boundary. The boundary condition in equation (5.21) becomes

$$\int_{\hat{s} \cdot \hat{n} < 0} \hat{s} \phi(r, t, \hat{s}) d\hat{s} = \int_{\hat{s} \cdot \hat{n} > 0} \hat{s} R \phi(r, t, \hat{s}) d\hat{s}, \quad \forall r \in \partial\Omega, \quad (5.23)$$

where R is the boundary reflection factor. From experimental measurements of R for different refractive indices, Egan and Hilgeman [1979] found the polynomial approximation $R \approx -1.4399n_{\Omega}^{-2} + 0.7099n_{\Omega}^{-1} + 0.6681 + 0.0636n_{\Omega}$, where n_{Ω} is the refractive index of the medium. This expression is only valid when one of the media has a refractive index equal to one. For the diffusion approximation this boundary condition takes the form [Schweiger *et al.*, 1995].

$$\Phi(r, t) + 2\kappa(r)A \frac{\partial \Phi(r, t)}{\partial \hat{n}} = 0, \quad \forall r \in \partial\Omega, \quad (5.24)$$

where the parameter A is generally defined as $A = \frac{1+R}{1-R}$.

5.1.1.1.2 Source conditions

The light sources s_i on the boundary $\partial\Omega$ are usually simulated using the collimated source model or the diffuse boundary source model [Schweiger *et al.*, 1995]. A collimated source is represented by a diffuse point source of the form

$$q_0 = \delta(r - r_s), \quad (5.25)$$

where r_s is at depth $1/\mu'_s$ (reduced mean free path) below the surface, approximately the distance at which the propagation of light becomes isotropic. Recall from § 2.1 that the probability of survival of a photon in a random walk after travelling a distance z follows an exponential decay, $\exp(-\mu'_s z)$.

If the anisotropic source q_1 in equation (5.11) is kept, then after applying Fick's law (equation (5.12)) the new source term \tilde{q}_0 in the diffusion equation becomes [Arridge, 1999]

$$\tilde{q}_0(r, t) = q_0(r, t) - 3\nabla \cdot k(r)q_1(r, t). \quad (5.26)$$

In the diffuse boundary source model, the sources are inward directed diffuse photon currents Γ_s placed on the boundary at positions $\partial\Omega_s \subset \partial\Omega$. The Robin boundary condition in equation (5.24) becomes

$$\Phi(r, t) + 2\kappa(r)A \frac{\partial \Phi(r, t)}{\partial \hat{n}} = \begin{cases} 0, & \forall r \in \partial\Omega \setminus \partial\Omega_s \\ -4\Gamma_s w(s_i), & \forall r \in \partial\Omega_s \end{cases}, \quad (5.27)$$

where the function $w(s_i)$ indicates if the source s_i is on ($w(s_i) = 1$) or off ($w(s_i) = 0$). The source is usually defined as a Gaussian profile in order to model light emitted by an optical fiber.

The quantity measured by the detectors d_i located on the boundary is the exitance (flux), defined as

$$\Gamma(r, t) = \int_{4\pi} \phi(r, t, \hat{s}) \hat{s} \cdot \hat{n} d\hat{s}. \quad (5.28)$$

Using Fick's law the exitance can be written as

$$\Gamma(r, t) = \hat{n} \cdot J(r, t) = -\kappa(r) \frac{\partial \Phi(r, t)}{\partial \hat{n}}, \quad (5.29)$$

which can be further simplified using equation (5.24) to

$$\Gamma(r, t) = \frac{1}{2A} \Phi(r, t). \quad (5.30)$$

The time- independent equations can be obtained by integrating over time the boundary and source conditions. Alternatively, it can be done by transforming to the frequency domain and taking the DC component.

5.1.1.1.3 Analytical solutions - Green's functions

The analytical solution of the diffusion equation is a Green's function when the source is a delta function at position r' and time t' , $\delta(r', t')$. A detailed description of the analytical methods can be found in [Patterson *et al.*, 1989; Arridge *et al.*, 1992]. The analytical solution of the diffusion equation is obtained by convolving its Green's function with the source term

$$\Phi(r, t) = \int_{-\infty}^{\infty} \int_{\Omega} g^{(\Phi)}(r, r', t - t') q_0(r', t') dr' dt', \quad (5.31)$$

where $g^{(\Phi)}$ is the time-domain Green's function solution of the diffusion equation. In the frequency-domain

$$\Phi(r, \omega) = \int_{\Omega} G^{(\Phi)}(r, r', \omega) q_0(r', \omega) dr', \quad (5.32)$$

where $G^{(\Phi)}$ is the corresponding Green's function solution. If the source is assumed to be a delta function, i.e. point source, the solution to the diffusion equation reduces to the Green's function: $\Phi(r, t) = g^{(\Phi)}(r, t)$ or $\Phi(r, \omega) = G^{(\Phi)}(r, \omega)$.

For an infinite homogeneous medium and an isotropic point source, the solutions of the diffusion equation in the temporal- and frequency domain are given, respectively, by

$$g_{inf}^{(\Phi)}(r, t) = \frac{1}{(4\pi c\kappa(t - t'))^{3/2}} \exp\left(-\frac{|r - r'|^2}{4c\kappa(t - t')} - \mu_a c(t - t')\right), \quad (5.33)$$

$$G_{inf}^{(\Phi)}(r, \omega) = \frac{\exp(-i\omega t')}{2|r - r'|(2\pi)^{3/2} c\kappa} \exp\left(-|r - r'| \left(\frac{(\mu_a + i\omega/c)}{\kappa}\right)^{1/2}\right). \quad (5.34)$$

For the semi-infinite space the point source is assumed to be at a depth $z_0 = 1/\mu'_s$. The method of images is used to find the solution for this geometry, by placing a negative image source outside the medium at $-z_0$, which forces the photon density to zero $\Phi = 0$ on the boundary at $z=0$ (see figure 5.1). The solutions are given by

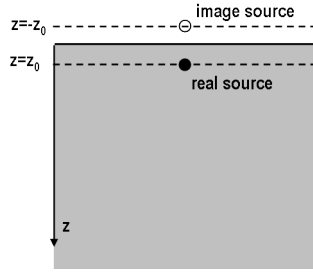


Figure 5.1: Geometry used to derive the Green's function solution to the diffusion equation for a semi-infinite homogeneous medium. The isotropic point source is indicated by the filled circle. The boundary condition is satisfied by adding a negative source indicated by the open circle [Patterson *et al.*, 1989].

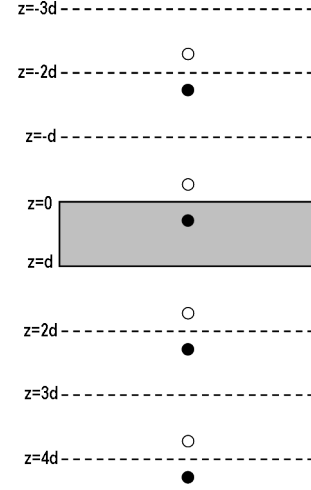


Figure 5.2: Geometry used to derive the Green's function solution to the diffusion equation for a finite homogeneous slab. The boundary conditions are met by adding multiple positive and negative sources [Patterson *et al.*, 1989].

$$g_{semi-inf}^{(\Phi)}(r, t) = \frac{1}{(4\pi c\kappa(t-t'))^{3/2}} \exp\left(-\mu_a c(t-t') - \frac{\rho^2}{4c\kappa(t-t')}\right) \times \left[\exp\left(\frac{-(z-z_0)^2}{4c\kappa(t-t')}\right) - \exp\left(\frac{-(z+z_0)^2}{4c\kappa(t-t')}\right) \right], \quad (5.35)$$

$$G_{semi-inf}^{(\Phi)}(r, \omega) = \frac{\exp(-i\omega t')}{2(2\pi)^{3/2} c\kappa} \left[\frac{\exp\left(\left(\frac{-\mu_a + i\omega/c}{\kappa}\right) (\rho^2 + (z-z_0)^2)\right)^{1/2}}{(\rho^2 + (z-z_0)^2)^{1/2}} - \frac{\exp\left(\left(\frac{-\mu_a + i\omega/c}{\kappa}\right) (\rho^2 + (z+z_0)^2)\right)^{1/2}}{(\rho^2 + (z+z_0)^2)^{1/2}} \right], \quad (5.36)$$

where $\rho = \sqrt{(x-x')^2 + (y-y')^2}$.

For the Robin boundary condition with boundary reflectivity (equation (5.24)), the photon density is forced to zero on the extrapolated boundary, shifted outwards from $\partial\Omega$. The extrapolated boundary is shifted by

$$z_R = 2\kappa \frac{1+R}{1-R}. \quad (5.37)$$

The solution for this boundary condition can be found using the method of images, with the source at $z = z_0$ and negative image source placed at $z = -z_0 - 2z_R$.

For an infinite slab geometry of thickness d , the method of multiple images is used to set the photon density to zero on the boundaries $z = 0$ and $z = d$. This condition can be met by

placing an infinite number of positive sources at location $z = 2Nd + z_0$ and negative sources at locations $z = 2Nd - z_0$, where N is an integer. The first four positive and negative sources are shown in figure 5.2. The solutions are similar to equations (5.35) and (5.36), but for the multiple source locations and summing all source terms, i.e. for the time-domain

$$g_{slab}^{(\Phi)}(r, t) = \frac{1}{(4\pi c\kappa(t-t'))^{3/2}} \exp\left(-\mu_a c(t-t') - \frac{\rho^2}{4c\kappa t}\right) \times \sum_{N=-\infty}^{N=\infty} \left[\exp\left(-\frac{(z-2Nd-z_0)^2}{4c\kappa(t-t')}\right) - \exp\left(-\frac{(z-2Nd+z_0)^2}{4c\kappa(t-t')}\right) \right], \quad (5.38)$$

For solutions to other geometries such as cylinders and spheres refer to [Arridge *et al.*, 1992]. Analytical solutions are limited to simple geometries and homogeneous optical properties. Numerical methods of solving the diffusion equation are generally necessary for more complex media.

5.1.1.1.4 Numerical methods

5.1.1.1.4.1 Finite element method

The finite element method (FEM) can be used to solve the diffusion equation numerically. In FEM the domain Ω is divided into P elements, joined at N nodes [Arridge, 1999; Schweiger *et al.*, 1995]. The solution of the diffusion equation Φ in FEM is the approximate piecewise function

$$\Phi^h(r) = \sum_{j=1}^N \Phi_j u_j(r), \quad (5.39)$$

where u_j are the basis functions, which inform how the function varies between nodes. This function decreases from one at the j th node to zero to the surrounding nodes.

The Galerkin approach allows this discretisation by converting the differential diffusion equation into an integral equation. The weak formulation requires the residual R to be zero over the domain, where R in the frequency-domain is

$$R(r, \omega) = \left(-\nabla \cdot \kappa(r) \nabla + \mu_a(r) + \frac{i\omega}{c} \right) \Phi^h(r, \omega) - q_0(r, \omega), \quad (5.40)$$

and, therefore, the sum of the inner product of R with the basis functions is required to be zero

$$\int_{\Omega} u_i(r) R(r, \omega) d\Omega = \int_{\Omega} u_i(r) \left[\left(-\nabla \cdot \kappa(r) \nabla + \mu_a(r) + \frac{i\omega}{c} \right) \Phi^h(r, \omega) - q_0(r, \omega) \right] = 0. \quad (5.41)$$

The Green's first identity theorem for any functions u, w is derived from the relation

$$\nabla \cdot (\kappa u \nabla w) = u \nabla \cdot (\kappa \nabla w) + \kappa \nabla u \cdot \nabla w, \quad (5.42)$$

and by applying the divergence theorem

$$\int_{\Omega} u \nabla \cdot \kappa \nabla w d\Omega = \int_{\partial\Omega} u \kappa \hat{n} \cdot \nabla w d(\partial\Omega) - \int_{\Omega} \kappa \nabla u \cdot \nabla w d\Omega. \quad (5.43)$$

Applying the divergence theorem to the term containing κ in equation (5.41)

$$\int_{\Omega} u_i(r) \nabla \cdot (\kappa \nabla \Phi^h(r, \omega)) d\Omega = \int_{\partial\Omega} u_i(r) \kappa \hat{n} \cdot \nabla \Phi^h(r, \omega) d(\partial\Omega) - \int_{\Omega} \kappa \nabla u_i(r) \cdot \nabla \Phi^h(r, \omega) d\Omega, \quad (5.44)$$

and replacing into equation (5.41) leads to

$$\begin{aligned} & \int_{\Omega} \kappa \nabla u_i(r) \cdot \nabla \Phi^h(r, \omega) d\Omega - \int_{\partial\Omega} u_i(r) \kappa \frac{\partial \Phi^h(r, \omega)}{\partial \hat{n}} d(\partial\Omega) \\ & + \int_{\Omega} \mu_a \Phi^h(r, \omega) u_i(r) d\Omega + \frac{i\omega}{c} \int_{\Omega} \Phi^h(r, \omega) u_i(r) d\Omega \\ & = \int_{\Omega} q_0 u_i(r) d\Omega. \end{aligned} \quad (5.45)$$

This equation can be written in a matrix form

$$(K(\kappa) + C(\mu_a) + i\omega B) \Phi(\omega) = Q(\omega), \quad (5.46)$$

where the matrix elements are given by

$$K_{i,j} = \int_{\Omega} \kappa \nabla u_i(r) \nabla u_j(r) d\Omega, \quad (5.47)$$

$$C_{i,j} = \int_{\Omega} \mu_a u_i(r) u_j(r) d\Omega, \quad (5.48)$$

$$B_{i,j} = \int_{\Omega} \frac{1}{c} u_i(r) u_j(r) d\Omega, \quad (5.49)$$

$$Q_j = \int_{\Omega} u_i(r) q_0(r, \omega) d\Omega, \quad (5.50)$$

$$\Phi = [\Phi_1, \Phi_2, \dots, \Phi_N]^T. \quad (5.51)$$

Similarly, in the time domain the matrix equation can be expressed as

$$(K(\kappa) + C(\mu_a)) \Phi(t) + B \frac{\partial \Phi(t)}{\partial t} = Q(t). \quad (5.52)$$

For a Dirichlet boundary condition the basis functions associated with boundary nodes are set to zero, except for source nodes. If the Robin boundary condition with reflection and a diffuse point source (equation (5.24)), for any point $r_{\partial\Omega}$ on the boundary, is expressed as

$$\frac{\partial \Phi^h(r_{\partial\Omega}, \omega)}{\partial \hat{n}} = f(r_{\partial\Omega}) \Phi(r_{\partial\Omega}, \omega) + g(r_{\partial\Omega}), \quad (5.53)$$

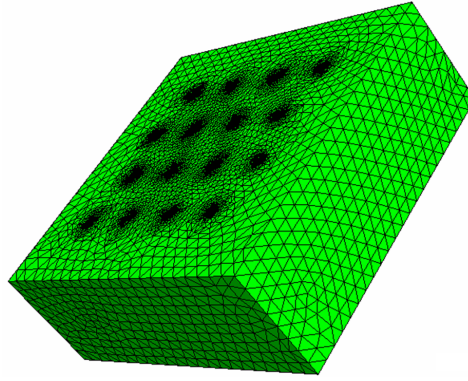


Figure 5.3: A slab finite element mesh. The triangles are the elements and the vertexes the nodes. The regions of higher node density correspond to the location of the optodes.

the source term Q vanishes and equation (5.46) becomes

$$(K(\kappa) + C(\mu_a) + F(A) + i\omega B) \Phi(\omega) = G(\omega), \quad (5.54)$$

where

$$F = \int_{\partial\Omega} \frac{1}{2A} u_i(r) u_j(r) d(\partial\Omega), \quad (5.55)$$

$$G = \int_{\partial\Omega} \frac{-2\Gamma_s \omega(s)}{A} u_i(r) d(\partial\Omega). \quad (5.56)$$

For an isotropic source $G(\omega) = Q(\omega)$. The exitance at the measurement points is given by equation (5.30), which is the measurement operator \mathcal{M} in the FEM forward problem for the i th source, defined as [Schweiger *et al.*, 1995]

$$F_i = \mathcal{M}[\Phi_i^h]. \quad (5.57)$$

The FEM framework is based on discrete points, the nodes. Different optical properties can be assigned and complex boundary shapes can be modelled with high accuracy. The basis functions ensure the continuity of the solution, interpolating the values at the nodes through all the points connecting between them. The meshes used in FEM usually have higher node density near the boundary, in particular near the sources and detectors (optodes) for optical imaging, since the rate change of photon density is greatest in this region. This is the approach used by the software package TOAST (Temporal Optical Absorption and Scattering Tomography), which has been developed by Prof. Arridge and Dr. Schweiger at University College London (UCL). Figure 5.3 shows a mesh generated with the meshing software NETGEN by Schöberl [1997].

5.1.1.1.4.2 Finite difference method

The finite difference method (FDM) finds the solution to the diffusion equation on a regular grid, by approximating the differential equations to finite difference quotients [Arridge, 1999; Klose & Hielscher, 1999]. This method is limited to rectangular geometries, since the nodes are assumed to be equally spaced. To model accurately irregular boundaries the number of nodes

can be increased at the expense of computation time. The solutions are calculated at the nodes and the function between nodes is not known. The FDM is relatively easy to implement but the approximate solution is less accurate for complex structures and regions near the boundary.

5.1.1.1.5 Validity of the diffusion approximation

The main conditions for the applicability of the diffusion approximation are that the absorption coefficient is significantly smaller than the reduced scattering coefficient, $\mu_a \ll \mu'_s$, and that the measurement point is at least $1/\mu'_s$ away from the source.

Light propagation cannot be modelled using the diffusion approximations in the presence of non-scattering regions, like the cerebrospinal fluid (CSF) surrounding the brain. Methods such as Monte Carlo (described later in this section) can provide accurate numerical solutions but are computationally time demanding. A common approach is to use hybrid methods, where the diffusion approximation is applied to highly scattering regions and another method models the non-scattering regions. Hybrid methods include the Monte Carlo-diffusion method [Hayashi *et al.*, 2003], the RTE-diffusion method [Tarvainen *et al.*, 2005] and the radiosity-diffusion method [Arridge *et al.*, 2000; Dehghani *et al.*, 2000]. The radiosity theory calculates the flux Γ_i at the point r on the outer surface, due to contributions from an inner surface $\partial\Omega'$. For the frequency-domain, under the Robin boundary condition with reflectivity, the flux Γ_i is given by [Arridge *et al.*, 2000; Dehghani *et al.*, 2000]

$$\Gamma_i(r) = \int_{\partial\Omega'} \frac{\phi(r') \cos(\theta_1) \cos(\theta_2)}{|r' - r|^2} \frac{1}{2A} \exp[-(|r' - r|)(\mu_a + i\omega/c)] dr' \quad (5.58)$$

where I is the radiance at point r' located on the surface $\partial\Omega'$, θ_1 is the angle between the travelling direction at position r' and the normal at the same point, and θ_2 is the equivalent for position r .

5.1.1.2 Numerical solutions of the RTE

5.1.1.2.1 Monte Carlo method

The Monte Carlo (MC) simulation is a stochastic method that models individual photons travelling through the medium that undergo discrete absorption and scattering events [Wang & Wu, 2007]. It can model the propagation of photons through complex geometries and regions with different optical properties. MC is relatively simple to implement, the required parameters are the absorption coefficient μ_a , the scattering coefficient μ_s and the scattering phase function $p(\hat{s}, \hat{s}')$. Individual photons are injected into the medium and travel in a straight line until they are scattered. The probability of a photon being scattered, after travelling a distance τ , is randomly generated according to an exponential distribution: $P(\tau) = \mu_s \exp(-\mu_s \tau)$. The cumulative probability distribution of a scattering event is given by

$$\int_0^\tau P(\tau) d\tau = 1 - \exp(-\mu_s \tau) = r, \quad (5.59)$$

where r is a randomly selected number from a uniform distribution in the range $0 \leq r \leq 1$. Therefore, a scattering event will take place after a distance

$$\tau = -\frac{1}{\mu_s} \ln(1-r) = -\frac{1}{\mu_s} \ln(r). \quad (5.60)$$

The direction of propagation after scattering is given by a polar angle θ , whose probability is described by the phase function, and an azimuthal angle Ψ

$$\Psi = 2\pi r_1, \quad (5.61)$$

$$\int_0^\theta p(\theta') d\theta' = r_2, \quad (5.62)$$

where r_1 and r_2 are uniformly distributed random numbers in the range $0 \leq r_1, r_2 \leq 1$.

Photon paths terminate when the photon is either completely absorbed or when it reaches the detection point. Absorption is included in the model between successive scattering events through a weight function. The photon with a weight w_i after a scattering event, reduces its contribution before it is scattered again, according to the weight function w

$$w = w_i \exp(-l\mu_a), \quad (5.63)$$

where l is the distance travelled since the previous scattering event.

One of the main disadvantages of MC methods is the required computation time. In order to obtain statistically significant results, millions of photons which suffer hundreds of interactions need to be followed. However, it can provide solutions of the RTE with great accuracy, including for media where the diffusion approximation fails.

5.1.1.2.2 Kubelka-Munk theory

The Kubelka-Munk (KM) theory is a simplification of the RTE. It considers two opposite directed fluxes, the forward (Γ_+) and backward flux (Γ_-), propagating through a multi-layered medium, where each layer is homogeneous [Kubelka, 1948, 1954]. Consider a homogeneous slab, of thickness d , where light propagates in the normal direction. The rate of change of the fluxes, as a function of depth z , is given by

$$d\Gamma_+ = -(K+S)\Gamma_+ dz + S\Gamma_- dz \quad \Rightarrow \quad \frac{d\Gamma_+}{dz} = -(K+S)\Gamma_+ + S\Gamma_-, \quad (5.64)$$

$$d\Gamma_- = S\Gamma_+(-dz) - (K+S)\Gamma_-(-dz) \quad \Rightarrow \quad \frac{d\Gamma_-}{dz} = -S\Gamma_+ + (K+S)\Gamma_-, \quad (5.65)$$

where K and S are the KM coefficients for absorption and scattering, respectively. For diffuse light these parameters are $K = 2\mu_a$ and $S = \mu_s$. Solving these differential equations for the boundary conditions $\Gamma_+(z=0) = \Gamma_0$ and $\Gamma_-(z=d) = 0$, where Γ_0 is the flux entering at $z=0$, yields

$$\Gamma_+ = \Gamma_0 \frac{\sinh(bSz)}{a\sinh(bSd) + b\cosh(bSd)}, \quad (5.66)$$

$$\Gamma_- = \Gamma_0 \frac{a \sinh(bSz) + b \cosh(bSz)}{a \sinh(bSd) + b \cosh(bSd)}, \quad (5.67)$$

where $a = \frac{S+K}{S}$ and $b = \sqrt{a^2 - 1}$.

This two-flux model is useful due to its simplicity, but is limited to diffuse fluxes. The four-flux theory can provide more accurate solutions, since it includes opposite directed collimated fluxes [Ishimaru, 1999]. The complexity and accuracy of the KM method can be increased by using multiple fluxes, but it will still be limited to simple homogeneous slab geometries.

5.2 Inverse problem

The inverse problem involves finding the optical properties x of the medium, μ_a and κ , from a set of measurements y made at the surface. The forward problem can be expressed as

$$y = F(x). \quad (5.68)$$

The optical properties can be obtained by solving the inverse problem

$$x = F^{-1}(y). \quad (5.69)$$

The forward operator is non-linear but it can be linearised and reduced to a so-called difference method, based on the assumption that the real optical properties are close to an initial guess x_0 . Therefore the measured data y is close to the modelled data y_0 [Arridge, 1999, 1997]. The forward problem can be expressed in a Taylor series

$$y = y_0 + F'(x_0)(x - x_0) + F''(x_0)(x - x_0)^2 + \dots, \quad (5.70)$$

where F' and F'' are the first and second Fréchet derivatives, represented by the Jacobian (J) and Hessian (H) matrices, respectively. If the higher order terms are ignored the problem reduces to the linear problem valid for small perturbations occurring in the medium

$$\Delta y = J \Delta x. \quad (5.71)$$

Therefore, the change in the optical properties $\Delta x = x - x_0$ can be obtained from the change in the measurements $\Delta y = y - y_0$ by inverting the matrix J : $\Delta x = J^{-1} \Delta y$.

5.2.1 Jacobian calculation

Let Φ be a solution to the diffusion equation in the frequency-domain (equation (5.17)), and let $\bar{\Phi} = \Phi + \delta\Phi$ be the solution for the same source, but in the presence of a perturbation in the optical properties given by [Arridge, 1999]

$$\bar{\mu}_a = \mu_a + \delta\mu_a, \quad (5.72)$$

$$\bar{\kappa} = \kappa + \delta\kappa. \quad (5.73)$$

Substituting into equation (5.17) leads to

$$\begin{aligned}
& -\nabla \cdot (\kappa(r) + \delta\kappa(r)) \nabla (\Phi(r, \omega) + \delta\Phi(r, \omega)) + (\mu_a(r) + \delta\mu_a(r)) (\Phi(r, \omega) + \delta\Phi(r, \omega)) \\
& + \frac{i\omega}{c} (\Phi(r, \omega) + \delta\Phi(r, \omega)) = q_0(r, \omega).
\end{aligned} \tag{5.74}$$

Assuming that all the second order terms in the previous equation are approximately zero, subtracting the original diffusion equation, and rearranging by keeping the terms corresponding to the homogeneous equation on the left hand side and perturbations terms on the right side, yields

$$-\nabla \cdot \kappa(r) \nabla \delta\Phi(r, \omega) + \mu_a(r) \delta\Phi(r, \omega) + \frac{i\omega}{c} \delta\Phi(r, \omega) = \nabla \cdot \delta\kappa(r) \nabla \Phi(r, \omega) - \delta\mu_a(r) \Phi(r, \omega). \tag{5.75}$$

Using the definition of the Green's function in equation (5.32), the perturbed solution $\delta\Phi$ of the diffusion equation is

$$\begin{aligned}
\delta\Phi(r, \omega) &= \bar{\Phi}(r, \omega) - \Phi(r, \omega) \\
&= \int_{\Omega} G^{(\Phi)}(r, r', \omega) \nabla \cdot \delta\kappa(r') \nabla \Phi(r', \omega) dr' - \int_{\Omega} G^{(\Phi)}(r, r', \omega) \delta\mu_a(r') \Phi(r', \omega) dr'.
\end{aligned} \tag{5.76}$$

Applying the divergence theorem in equation (5.43) to the term containing a scattering perturbation and the boundary condition $\Phi(r, \omega) = 0$ for $r \in \partial\Omega$, yields the Born approximation [Arridge, 1999]

$$\delta\Phi(r, \omega) = - \int_{\Omega} \delta\kappa(r') \nabla G^{(\Phi)}(r, r', \omega) \cdot \nabla \Phi(r', \omega) dr' - \int_{\Omega} \delta\mu_a(r') G^{(\Phi)}(r, r', \omega) \Phi(r', \omega) dr'. \tag{5.77}$$

For a δ source located at position r_s the solution can be written as

$$\begin{aligned}
\delta\Phi(r, r_s, \omega) &= - \int_{\Omega} \delta\kappa(r') \nabla G^{(\Phi)}(r, r', \omega) \cdot \nabla G^{(\Phi)}(r', r_s, \omega) dr' \\
&\quad - \int_{\Omega} \delta\mu_a(r') G^{(\Phi)}(r, r', \omega) G^{(\Phi)}(r', r_s, \omega) dr'.
\end{aligned} \tag{5.78}$$

However, the flux is the measured quantity at the detection point r_d on the boundary $\partial\Omega$, given by equation (5.29). Similarly, the perturbed flux $\delta\Gamma$ for a δ source is given by

$$\begin{aligned} \delta\Gamma(r_d, r_s, \omega) = & - \int_{\Omega} \delta\kappa(r') \nabla G^{(\Gamma)}(r_d, r', \omega) \cdot \nabla G^{(\Phi)}(r', r_s, \omega) dr' \\ & - \int_{\Omega} \delta\mu_a(r') G^{(\Gamma)}(r_d, r', \omega) G^{(\Phi)}(r', r_s, \omega) dr', \end{aligned} \quad (5.79)$$

where $G^{(\Gamma)}$ is the result of substituting the Green's function of the diffusion approximation into equation (5.29).

The Rytov approximation considers logarithmic changes in the measurements: $\delta(\log(\Gamma)) = \delta\Gamma/\Gamma$. Therefore, the Rytov approximation is the normalised Born approximation given by [Arridge, 1999]

$$\begin{aligned} \delta\Gamma(r_d, r_s, \omega) = & \frac{-1}{G^{(\Gamma)}(r_d, r_s, \omega)} \left(\int_{\Omega} \delta\kappa(r') \nabla G^{(\Gamma)}(r_d, r', \omega) \cdot \nabla G^{(\Phi)}(r', r_s, \omega) dr' \right. \\ & \left. + \int_{\Omega} \delta\mu_a(r') G^{(\Gamma)}(r_d, r', \omega) G^{(\Phi)}(r', r_s, \omega) dr' \right). \end{aligned} \quad (5.80)$$

The reciprocity theorem states that the measurement flux at r_d due to an isotropic source at r' is equal to the measurement of the photon density at r' due to a source at position r_d [Arridge & Schweiger, 1995], i.e.

$$G^{(\Gamma)}(r_d, r', \omega) = G^{(\Phi)+}(r', r_d, \omega), \quad (5.81)$$

where $G^{(\Phi)+}$ represents the adjoint Green's function, whose direction of propagation is the opposite to the normal Green's function.

The Jacobian represents the sensitivity of all the source-detector pair measurements to local perturbations in the optical properties of the medium. For the Born approximation together with the adjoint Green's function, the expression for the Jacobian is

$$J(r_d, r', r_s, \omega) = \begin{pmatrix} \frac{\partial\Gamma(r_d, r_s, \omega)}{\partial\delta\mu_a(r')} \\ \frac{\partial\Gamma(r_d, r_s, \omega)}{\partial\delta\kappa(r')} \end{pmatrix} = \begin{pmatrix} -G^{(\Phi)+}(r', r_d, \omega) G^{(\Phi)}(r', r_s, \omega) \\ -\nabla_{r'} G^{(\Phi)+}(r', r_d, \omega) \cdot \nabla_{r'} G^{(\Phi)}(r', r_s, \omega) \end{pmatrix}. \quad (5.82)$$

This is a complex matrix, where the real part represents the changes in the amplitude of the measurements and the imaginary part the changes in phase.

In the FEM framework, one can consider $\vec{\Phi}^{(i)}(\omega)$ to be the solution vector in the mesh for a source at position s_i and let $\vec{\Phi}_{Adj}^{(j)}(\omega)$ be the solution vector in the mesh for an adjoint source at position d_j on the boundary [Arridge & Schweiger, 1995]. The Jacobian can be calculated for a variety of measurement types (for more information refer to [Arridge & Schweiger, 1995; Arridge, 1995]). For the intensity solutions taken at zero frequency, the Jacobian for an absorption perturbation is

$$J_{\delta\mu_a}(s_i, d_j, r') = \left(\vec{\Phi}^{(i)} \times \vec{\Phi}_{Adj}^{(j)} \right) (r'). \quad (5.83)$$

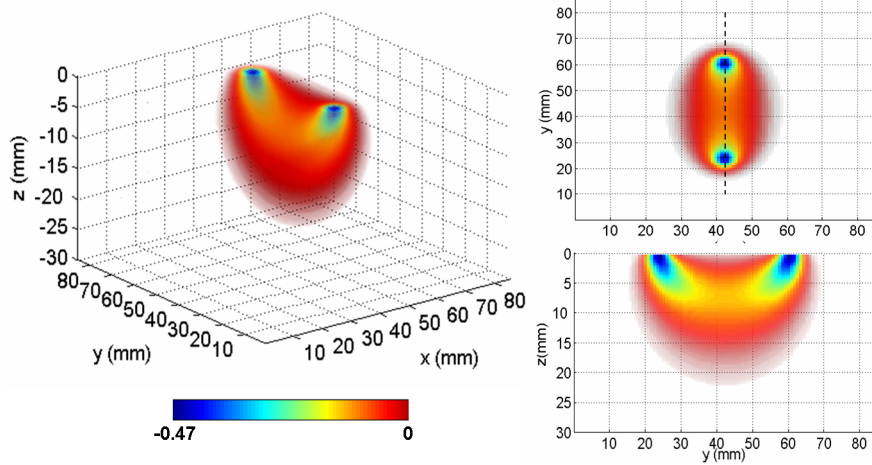


Figure 5.4: Absorption photon measurement density function for a slab geometry with dimensions $85 \text{ mm} \times 85 \text{ mm} \times 30 \text{ mm}$, optical properties $\mu_a = 0.01 \text{ mm}^{-1}$ and $\mu'_s = 1 \text{ mm}^{-1}$, and an optode separation of 35 mm (left). All the values are negative; the hot colours are values closer to zero and cold colours are more negative and show areas of higher sensitivity (units in mm). Top view of the PMDF (top right), where the dashed line marks the cross sectional plane of the lateral view (bottom right).

Similarly, the Jacobian for a perturbation in the diffusion coefficient is given by

$$J_{\delta\kappa}(s_i, d_j, r') = \left(\nabla \vec{\Phi}^{(i)} \times \nabla \vec{\Phi}_{Adj}^{(j)} \right) (r'). \quad (5.84)$$

Each row of the Jacobian matrix represents the sensitivity of a source-detector pair given an infinitesimal perturbation in the optical property. Each sensitivity map is known as a photon measurement density function (PMDF), and indicates to which regions the measurement is more sensitive, which can be interpreted as the probability of an infinitesimal perturbation at a location within the medium being detected. The columns of the Jacobian represent the perturbation in the data due to a change in the optical property at a node. Therefore, the Jacobian is a matrix whose rows are the measurements m and columns the nodes n . The PMDF has its maxima near the source and detector positions. The volume sensitivity is narrowest near the optodes and broadest in the middle, as shown in figure 5.4.

Using the adjoint method to calculate the Jacobian matrix saves computation time. Instead of finding the flux for each perturbed node, the same information can be obtained by replacing the measurement point with an adjoint source and the photon density is calculated for all nodes, as illustrated in figure 5.5.

5.2.2 Linear reconstruction

Linear image reconstruction aims to find the changes in the optical properties and hence changes in the chromophore concentrations from the changes in a set of measurements, by inversion of the Jacobian matrix

$$\Delta x = J^{-1} \Delta y. \quad (5.85)$$

Using the Rytov approximation the changes in log intensity are linearised. The reconstruction using logarithmic intensity gives improved images compared to intensity measurements, since

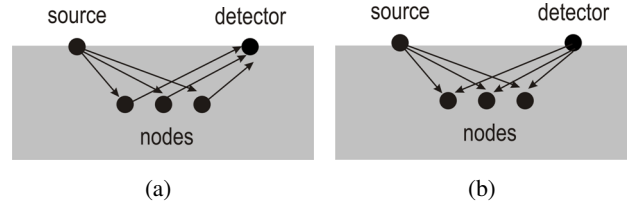


Figure 5.5: Simplified diagrams illustrating the determination of the sensitivity function. (a) Direct solution, where the forward problem needs to be solved for each absorption or scattering perturbation in the medium. (b) Adjoint method, where the sensitivity function is obtained in two steps: by calculating the forward and adjoint problem.

it compensates for the high sensitivity near the surface by increasing the sensitivity to inner regions [Arridge & Hebden, 1997].

By applying some normalisation to equation (5.85) the quality of the image may improve [Gibson *et al.*, 2006]. Because μ_a and κ have different units and range of values, Δx and J are normalised to the mean optical parameters. Also, in order to have equivalent contributions from all nodes, a row normalisation can be performed, where each row is divided by its mean. Finally, the standard deviation σ of the log amplitude can be considered a measure of the level of confidence in the data points, so we apply a correction by dividing by σ on both sides of the equation. The noise is considered to be Poisson distributed with standard deviation $\sigma = \sqrt{y}$ [Arridge & Schotland, 2009]. According to the Central Limit Theorem, if the number of detected photons is large enough, then the detected intensity approximates a Gaussian distribution [Arridge & Schotland, 2009].

The source and detector coupling coefficients can also be derived if they are included as unknowns in the image reconstruction. The Jacobian matrix becomes $\hat{J} = [JSD]$ and $\hat{\Delta}x = [\Delta x_1, \dots, \Delta x_{Nv}, \log S_1, \dots, \log S_{Ns}, \log D_1, \dots, \log D_{Nd}]$, where Nv is the number of voxels, Ns the number of sources and Nd the number of detectors. The elements of matrix [S D] are either one or zero, indicating which source and detector correspond to each measurement. Further details can be found in [Boas *et al.*, 2001c].

Image reconstruction is an under-determined problem, since the number of unknowns is much larger than the number of measurements and it is ill-posed.

5.2.2.1 Tikhonov regularisation

Consider a generic discrete ill-posed problem of the form

$$Ax = b, \quad (5.86)$$

where A is an $m \times n$ matrix and b is a vector of dimension m that corresponds to the measurements. A problem is well-posed in the Hadamard sense if the following conditions are true [Hadamard, 1902]:

1. A solution exists;
2. The solution is unique;

3. The solution depends continuously on the data.

A problem that does not satisfy one or more of these conditions is termed ill-posed. In discrete ill-posed problems the matrix A is ill-conditioned and its singular values decay gradually to zero, without a clear gap in the singular values spectrum. Because the data b are contaminated by noise, added to the fact that A is ill-conditioned, a small change in the data may result in a large change in the solution x , and consequently a direct solution of the least squares problem, $\min_x \|Ax - b\|_2$, may not yield a sensible result. By using regularisation methods we are able to reduce the influence of noise.

A well known regularisation method used to obtain more stable solutions is the Tikhonov regularisation, which is defined as [Tikhonov & Arsenin, 1977]

$$\min_x \|Ax - b\|_2^2 + \lambda^2 \|x\|_2^2, \quad (5.87)$$

where $\|Ax - b\|_2$ is the least square or residual norm, $\|x\|_2$ is the regularised or solution norm and λ is the regularisation parameter. The solution of (5.87) as a function of λ is given by

$$x_\lambda = A^T (AA^T + \lambda^2 I)^{-1} b, \quad (5.88)$$

where A^T is the transpose of matrix A and I is the identity matrix. Applying the Tikhonov regularisation to the optical imaging problem gives

$$\Delta x_\lambda = J^T (JJ^T + \lambda^2 I)^{-1} \Delta y. \quad (5.89)$$

5.2.2.2 Reconstruction methods

Increasing the number of measurement wavelengths in the conventional approach to optical topography image reconstruction only increases the number of unknowns, since absorption and scattering need to be reconstructed for each wavelength and subsequently chromophore concentrations can be obtained. The multispectral method proposed by Corlu *et al* [2005] directly reconstructs the chromophore concentrations and scattering factors which are wavelength independent, meaning that the number of unknowns remains constant when the number of measurement wavelengths increases, representing a clear advantage over the conventional approach.

5.2.2.2.1 Conventional Method

The Jacobian matrix J of the conventional method, as described previously, represents the sensitivity of a measurement to changes in the optical properties within the medium. For CW measurements only intensity I is measured, hence the Jacobian is a $M \times N$ matrix with the form

$$J_\lambda(\mu_a, \kappa) = \begin{pmatrix} \frac{\partial I_1(\lambda)}{\partial \mu_{a1}} & \frac{\partial I_1(\lambda)}{\partial \mu_{a2}} & \cdots & \frac{\partial I_1(\lambda)}{\partial \mu_{aN}} & \frac{\partial I_1(\lambda)}{\partial \kappa_1} & \frac{\partial I_1(\lambda)}{\partial \kappa_2} & \cdots & \frac{\partial I_1(\lambda)}{\partial \kappa_N} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial I_M(\lambda)}{\partial \mu_{a1}} & \frac{\partial I_M(\lambda)}{\partial \mu_{a2}} & \cdots & \frac{\partial I_M(\lambda)}{\partial \mu_{aN}} & \frac{\partial I_M(\lambda)}{\partial \kappa_1} & \frac{\partial I_M(\lambda)}{\partial \kappa_2} & \cdots & \frac{\partial I_M(\lambda)}{\partial \kappa_N} \end{pmatrix}, \quad (5.90)$$

where each row represents a source-detector pair measurement and each column a perturbation node. Changes in the measured data ΔI are related to changes in the optical properties Δx through $\Delta I = J_\lambda \Delta x$, where $x = [\mu_a; \kappa]$. Optical properties are obtained by solving an ill-posed inverse problem at each wavelength and then combined to calculate the chromophore concentrations using equation (2.3). If measurements are performed at two different wavelengths and only contributions of oxy-haemoglobin HbO_2 and deoxy-haemoglobin HHb are considered, then equation (2.3) can be written as

$$\begin{aligned}\mu_a^{\lambda_1} &= \varepsilon_{HbO_2}^{\lambda_1} [HbO_2] + \varepsilon_{Hb}^{\lambda_1} [Hb] \\ \mu_a^{\lambda_2} &= \varepsilon_{HbO_2}^{\lambda_2} [HbO_2] + \varepsilon_{Hb}^{\lambda_2} [Hb]\end{aligned}\quad (5.91)$$

Note that both absorption coefficient and chromophore concentrations (denoted by square brackets) represent relative changes, which is not explicit in the equation for simplicity. Solving this system of two equations yields the chromophore concentrations

$$\begin{aligned}[HbO_2] &= \frac{\mu_a^{\lambda_1} \varepsilon_{Hb}^{\lambda_2} - \mu_a^{\lambda_2} \varepsilon_{Hb}^{\lambda_1}}{\varepsilon_{HbO_2}^{\lambda_1} \varepsilon_{Hb}^{\lambda_2} - \varepsilon_{HbO_2}^{\lambda_2} \varepsilon_{Hb}^{\lambda_1}} \\ [HHb] &= \frac{\mu_a^{\lambda_2} \varepsilon_{HbO_2}^{\lambda_1} - \mu_a^{\lambda_1} \varepsilon_{HbO_2}^{\lambda_2}}{\varepsilon_{HbO_2}^{\lambda_1} \varepsilon_{Hb}^{\lambda_2} - \varepsilon_{HbO_2}^{\lambda_2} \varepsilon_{Hb}^{\lambda_1}}.\end{aligned}\quad (5.92)$$

5.2.2.2 Multispectral Method

The multispectral method involves reconstructing chromophore concentrations and scatter parameters using all measurements at all wavelengths simultaneously. The derivatives with respect to absorption in equation (5.90) are related to derivatives with respect to chromophore concentrations by

$$\tilde{J}(c_i, \lambda) = \frac{\partial I(\lambda)}{\partial c_i} = \frac{\partial I(\lambda)}{\partial \mu_a(\lambda)} \frac{\partial \mu_a(\lambda)}{\partial c_i} = \frac{\partial I(\lambda)}{\partial \mu_a(\lambda)} \varepsilon_i(\lambda) = J_\lambda(\mu_a) \varepsilon_i(\lambda). \quad (5.93)$$

Similarly, for scattering parameters a and b

$$\tilde{J}(a, \lambda) = \frac{\partial I(\lambda)}{\partial a} = \frac{\partial I(\lambda)}{\partial \kappa(\lambda)} \frac{\partial \kappa(\lambda)}{\partial a} = \frac{\partial I(\lambda)}{\partial \kappa(\lambda)} \frac{\partial \kappa(\lambda)}{\partial \mu'_s(\lambda)} \frac{\partial \mu'_s(\lambda)}{\partial a}, \quad (5.94)$$

$$\tilde{J}(b, \lambda) = \frac{\partial I(\lambda)}{\partial b} = \frac{\partial I(\lambda)}{\partial \kappa(\lambda)} \frac{\partial \kappa(\lambda)}{\partial b} = \frac{\partial I(\lambda)}{\partial \kappa(\lambda)} \frac{\partial \kappa(\lambda)}{\partial \mu'_s(\lambda)} \frac{\partial \mu'_s(\lambda)}{\partial b}. \quad (5.95)$$

Using the diffusion coefficient $\kappa = 1/3(\mu_a + \mu'_s)$, the previous equations become

$$\tilde{J}(a, \lambda) = J_\lambda(\kappa) (-3\kappa(\lambda)^2) (\lambda^{-b}), \quad (5.96)$$

$$\tilde{J}(b, \lambda) = J_\lambda(\kappa) (-3\kappa(\lambda)^2) (\mu'_s(\lambda)) (-\ln \lambda). \quad (5.97)$$

The multispectral method combines equation (5.93), (5.96) and (5.97) for n measurement wavelengths as follows

$$\begin{pmatrix} \Delta I_{\lambda_1} \\ \Delta I_{\lambda_2} \\ \vdots \\ \Delta I_{\lambda_n} \end{pmatrix} = \begin{pmatrix} \tilde{J}(c_1, \lambda_1) & \tilde{J}(c_2, \lambda_1) & \dots & \tilde{J}(c_i, \lambda_1) & \tilde{J}(a, \lambda_1) & \tilde{J}(b, \lambda_1) \\ \tilde{J}(c_1, \lambda_2) & \tilde{J}(c_2, \lambda_2) & \dots & \tilde{J}(c_i, \lambda_2) & \tilde{J}(a, \lambda_2) & \tilde{J}(b, \lambda_2) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \tilde{J}(c_1, \lambda_n) & \tilde{J}(c_2, \lambda_n) & \dots & \tilde{J}(c_i, \lambda_n) & \tilde{J}(a, \lambda_n) & \tilde{J}(b, \lambda_n) \end{pmatrix} \begin{pmatrix} \Delta c_1 \\ \Delta c_2 \\ \dots \\ \Delta c_i \\ \Delta a \\ \Delta b \end{pmatrix}, \quad (5.98)$$

The matrix \tilde{J} has dimensions (number of measurements \times number of wavelengths) \times (number of chromophores and scatter parameters \times number of nodes).

5.2.3 Non-Uniqueness problem

Arridge and Lionheart [1998] have shown that the solution of the image reconstruction problem, using CW measurements at a single wavelength, is non-unique. This means that several pairs of distributions of μ_a and μ'_s can yield identical data. As a consequence, it is difficult to separate μ_a and μ'_s simultaneously and so-called cross-talk occurs.

Following the demonstration of non-uniqueness by Arridge and Lionheart [1998], the domain Ω is divided into two sub-domains Ω_0 and Ω_1 , hence $\Omega = \Omega_0 \cup \Omega_1$. For the sub-domain Ω_0 the source term q_0 is always zero. Whereas, the sub-domain Ω_1 surrounding Ω_0 , contains the isotropic source within a distance of the surface $z_0 = 1/\mu'_s$.

Simplifying the diffusion equation (equation (5.17)) with the change of variables $\gamma^2 = \kappa$ and $\Psi = \gamma\Phi$, the following Helmholtz-type equation is obtained

$$-\nabla^2 \Psi(\omega) + \eta(\omega)\Psi(\omega) = \frac{q_0}{\gamma}, \quad (5.99)$$

where $\eta(\omega) = \eta_0 + i\omega\xi$, with

$$\eta_0 = \left(\frac{\nabla^2 \gamma}{\gamma} \right) + \frac{\mu_a}{\gamma^2}, \quad \xi = \frac{1}{c\gamma^2} \quad (5.100)$$

Two sets of functions (κ, μ_a) and $(\tilde{\kappa}, \tilde{\mu}_a)$ will return the same solution Ψ if the following conditions are met

- Condition 1: $\tilde{\eta} = \eta \quad \forall \Omega$,
- Condition 2: $\tilde{\kappa} = \kappa \quad \forall \Omega_1$.

Note that any different set $\tilde{\kappa}$, also gives a different ξ and consequently this violates the first condition. However, for the dc case ($\omega = 0$) condition 1 becomes $\tilde{\eta}_0 = \eta_0$. Consider that a function α is added to μ_a and β to κ such that

$$\begin{aligned} \tilde{\kappa} &= \kappa + \beta \\ \tilde{\mu}_a &= \mu_a + \alpha. \end{aligned} \quad (5.101)$$

If condition 2 holds, then the two data sets provide equal solutions if

$$\left(\frac{\nabla^2(\kappa + \beta)^{1/2}}{(\kappa + \beta)^{1/2}} \right) + \frac{\mu_a + \alpha}{(\kappa + \beta)} = \left(\frac{\nabla^2(\kappa)^{1/2}}{(\kappa)^{1/2}} \right) + \frac{\mu_a}{(\kappa)}, \quad (5.102)$$

which is true when

$$\alpha = (\kappa + \beta) \left(\left(\frac{\nabla^2 \kappa^{1/2}}{\kappa^{1/2}} \right) + \frac{\mu_a}{\kappa} - \left(\frac{\nabla^2(\kappa + \beta)^{1/2}}{(\kappa + \beta)^{1/2}} \right) \right) - \mu_a, \quad (5.103)$$

for any β satisfying $\beta = 0$ in the sub-domain Ω_1 .

The results demonstrate that there are an infinite set of functions that give the same data. Therefore, there is no unique solution for the CW inverse problem. Multispectral reconstruction can help reduce the non-uniqueness problem as will be shown in §8.

5.2.4 Non-linear reconstruction

If the reconstruction problem cannot be linearised or if absolute values for the optical properties are required, then a non-linear approach has to be used. Non-linear methods use iterative processes that update J in order to minimise the difference between the measured data and the data calculated by the forward model. The image reconstruction software TOAST developed at UCL uses the FEM for the forward model and can apply a non-linear algorithm to find the optical properties that best fit the data (refer to <http://web4.cs.ucl.ac.uk/research/vis/toast/> for a complete list of references).

Linear reconstruction has the advantage of being computationally fast in comparison with non-linear and iterative reconstructions and, hence, is preferred.

5.3 Summary

Optical topography is mostly used to monitor haemodynamic activity occurring in the brain. However, due to the limited penetration depth of light in tissue it can only image superficial cortical regions. In order to obtain spatial maps representing the optical properties and chromophore concentrations of tissues beneath the imaging probe it is necessary to solve an ill-posed and under-determined problem. Regularisation methods, such as the zero-order Tikhonov regularisation, are required to solve this type of problem. It is necessary to find the optimal balance between the noise present in the image and its accuracy, which is controlled by the regularisation parameter. The next chapter analyses different methods used to select the regularisation parameter.

In optical topography the measurement sensitivity is higher nearer the surface than at deeper regions, which leads to limited depth information. This decrease in sensitivity can be compensated by introducing a spatially variant regularisation parameter in the image reconstruction. The influence of using a spatially variant regularisation parameter will be analysed later in this thesis.

Selection of regularisation parameter

In order to reconstruct a 3D optical topography image it is necessary to solve an ill-posed inverse problem where the number of unknowns is much greater than the number of measurements. Tikhonov regularisation, and all regularisation methods in general, replace an ill-posed problem with an approximate but less ill-posed problem, which ensures the existence and stability of the solution. The *residual norm* measures the proximity of the original data to the simulated data which is obtained when the regularised solution is used. This proximity cannot be smaller than the average of the perturbations present in the data b . The *solution norm* is a measure of the noise in the solution. The regularisation parameter λ finds a balance between these norms, so that the least squares solution approximates the true solution. If λ is too large the inverse problem is only slightly related to the original unregularised problem and important information in the data b may not be used. If λ is chosen to be small the solution will be highly dominated by noise.

One attempts to find a good solution to the problem by choosing a suitable regularisation parameter. It is important to have a systematic and reliable method for computing an optimum regularisation parameter and different methods have been proposed. In §6.2 the most common methods are summarised. One cannot expect to find a method that will always find a good λ for every problem. In this chapter, two specific problems that have common features are studied in anticipation that at least one of the methods will perform well in both cases. First, the methods are applied to a deblurring problem, which is ill-posed, but its solution is known, and the amount of noise and blur in the data can be controlled. Second, the methods are applied to optical topography data. This is a more complex problem than deblurring since in addition to the ill-posedness it is also under-determined (less measurements than unknowns), it is not an easy task to estimate the solution when no prior information is available about the object, and there are multiple sources of noise. This work has been published in [Correia *et al.*, 2009b].

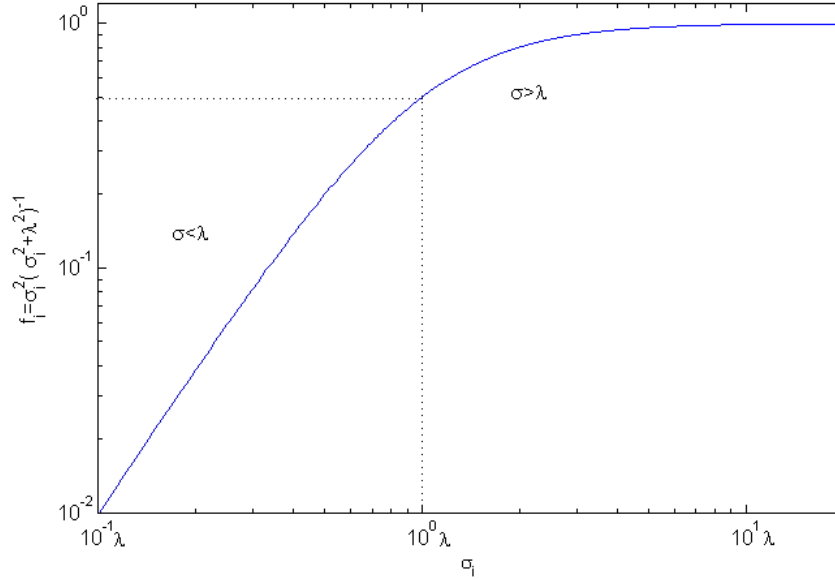


Figure 6.1: Tikhonov filter factors f_i

6.1 Singular Value Decomposition

The numerical technique known as Singular value decomposition (SVD) is employed here to illustrate the effect of Tikhonov regularisation (refer to §5.2.2.1) on the solution.

Let the SVD of matrix A be given by

$$A = USV^T = \sum_{i=1}^n u_i \sigma_i v_i^T, \quad (6.1)$$

where the left and right singular vectors u_i and v_i are orthonormal, and the singular values σ_i , which are the diagonal elements of S , appear in the following order: $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$.

The Tikhonov regularised solution is a function of the regularisation parameter λ , and the norm can be expressed in terms of SVD as [Hansen, 2001, 1998]

$$\|x_\lambda\|_2^2 = \sum_{i=1}^n \left(f_i \frac{u_i^T b}{\sigma_i} \right)^2, \quad (6.2)$$

where $f_i = \frac{\sigma_i^2}{\sigma_i^2 + \lambda^2}$ are the Tikhonov filter factors.

When the singular values are smaller than λ the filter factor is approximately σ_i^2/λ^2 which means that the filter factors are small, as illustrated in figure 6.1. This means that the singular vectors for small singular values are filtered, having a less significant contribution on the regularised solution. These singular vectors, which in the examples used here can be represented as images, reflect spatial variations. Smaller singular vectors correspond to low spatial frequencies, whereas higher singular vectors correspond to large spatial frequencies. Because the frequency of the singular vectors increases as the singular values decrease, by using the Tikhonov regularisation the high frequency oscillatory vectors are filtered. Note that

it is the division by small singular values, in the non-filtered solution, that amplifies the high frequency components of the data b . We can see in figure 6.1 that for singular values larger than the regularisation parameter, the filter factors approach one, which results in a significant contribution from low-frequency components to the regularised solution.

The residual norm in terms of SVD is

$$\|Ax_\lambda - b\|_2^2 = \sum_{i=1}^n ((1 - f_i) u_i^T b)^2, \quad (6.3)$$

where $1 - f_i = \frac{\lambda^2}{\sigma_i^2 + \lambda^2}$, and it is easy to see that for filter factors close to one the residual norm will be small, and vice-versa.

To ensure that the regularised solution approximates the exact solution, the discrete Picard condition (DPC) has to be satisfied for the given ill-posed problem [Hansen, 1990].

Definition 1 *The data vector b satisfies the DPC if the data space coefficients $|u_i^T b|$ on average decay to zero faster than the singular values σ_i .*

This condition ensures that the solution components related to small singular components are not over amplified. But in real problems, measured data b are generally contaminated by different types of noise and it usually does not fully satisfy the condition. If the DPC holds partially, until a point where the data space coefficient levels off and becomes dominated by errors, a valid solution should be obtained by any regularisation method. On the other hand, if the DPC is violated for a given problem one should question the validity of the solution.

6.2 Selection Methods

The influence of the regularisation parameter on the reconstructed image is fundamental. We have examined the most common methods, described below, in order to identify which one is the most suitable for ill-posed inverse problems such as imaging using optical topography. The methods are divided in two main groups: *generic*, which can be applied to all inverse problems in general, and *image based*, which are methods that use information extracted from the images to calculate the regularisation parameter. The generic methods are divided in two subgroups: the methods that use *noise statistics* and methods that *do not use noise statistics*.

6.2.1 Generic methods

6.2.1.1 Methods that *do not* use noise statistics

6.2.1.1.1 Heuristic Method

The Heuristic Method is a very popular method to find the regularisation parameter, mostly due to its simplicity [Graham & Adler, 2006]. It involves human selection by eye, where a subject generates several solutions for different regularisation parameters and selects the regularisation parameter which, in his/her opinion, results in the most acceptable reconstruction. However, this method is subjective, which is a disadvantage especially when there is no information available about the solution. It can lead to different λ choices by different subjects. It is also common to define the regularisation parameter equal to estimate of the noise present in the data [Gibson *et al.*, 2006].

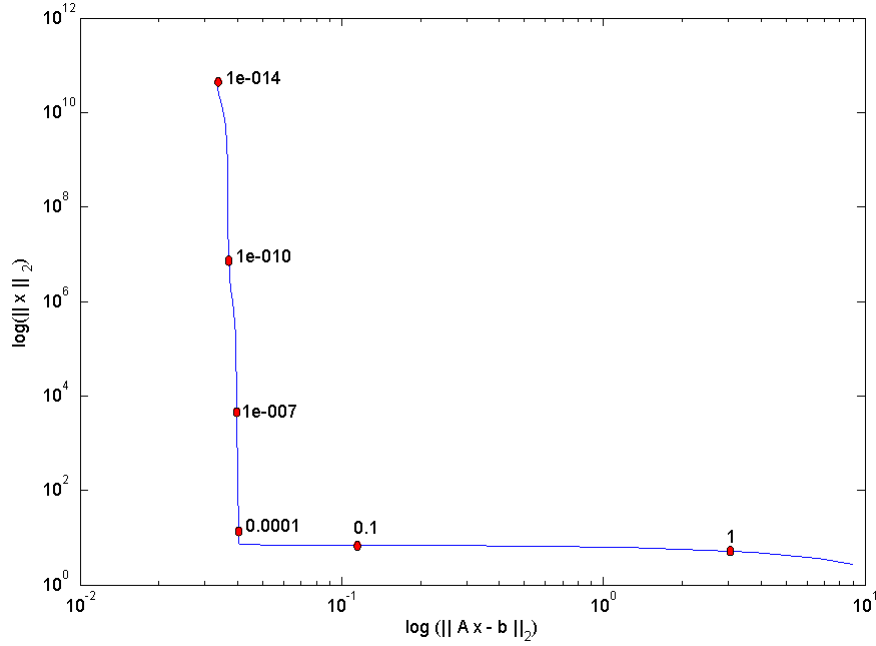


Figure 6.2: Example of an L-curve. The numbers correspond to the regularisation parameters for each point marked by the dots.

6.2.1.1.2 L-curve

The L-curve is perhaps the most commonly employed quantitative method for finding the regularisation parameter when solving ill-posed problems [Hansen, 2001, 1998; Hansen & O’Leary, 1993; Hansen, 1992]. The L-curve is a log-log plot, for different regularisation parameters, of the solution norm $\|x_\lambda\|_2$ against the residual norm $\|Ax_\lambda - b\|_2$ (see example in figure 6.2). We seek the minimisation of these two quantities, which theoretically corresponds to the corner of the graph, i.e., the point of maximum curvature.

The formulation described in [Hansen, 2001, 1998] is used to define

$$\eta = \|x\|_2^2, \quad \rho = \|Ax - b\|_2^2 \quad (6.4)$$

and

$$\hat{\eta} = \log \eta, \quad \hat{\rho} = \log \rho. \quad (6.5)$$

The point of maximum curvature of the L-curve k_λ is given by

$$k_\lambda = 2 \frac{\hat{\rho}''_\lambda \hat{\eta}'_\lambda - \hat{\rho}'_\lambda \hat{\eta}''_\lambda}{\left((\hat{\rho}'_\lambda)^2 + (\hat{\eta}'_\lambda)^2 \right)^{3/2}}, \quad (6.6)$$

where prime represents differentiation with respect to λ .

From equations (6.2), (6.3), (6.4), (6.5) and (6.6) the following expression is obtained

$$k_\lambda = 2 \frac{\eta_\lambda \rho_\lambda}{\eta'_\lambda} \frac{\lambda^2 \eta'_\lambda \rho_\lambda + 2\lambda \eta_\lambda \rho_\lambda + \lambda^4 \eta_\lambda \eta'_\lambda}{(\lambda^4 \eta^2 + \rho^2)^{3/2}}. \quad (6.7)$$

The behaviour of the solution norm for different regularisation parameters discussed earlier can be visualised on the L-curve (refer again to figure 6.2). For large λ the solutions are mostly influenced by the regularisation error. It is easy to see that for large values of λ , the solution norm $\|x_\lambda\|_2$ approaches the regularised exact solution $\|x_{\lambda,exact}\|_2$, where $x_{\lambda,exact} = A_\lambda^I b_{exact}$, with A_λ^I representing the Tikhonov regularised pseudoinverse of A and b_{exact} the unperturbed data vector. The residual norm $\|Ax_\lambda - b\|_2$ increases with λ . For small λ the contributions from the perturbation errors e in b , where the data $b = b_{exact} + e$, are more significant and $\|x_\lambda\|_2$ increases as λ decreases, whereas $\|Ax_\lambda - b\|_2$ becomes approximately $\|e\|$. The region where both $\|x_\lambda\|_2$ and $\|Ax_\lambda - b\|_2$ contribute equally corresponds to the L-curve corner, which is the maximum of k_λ .

6.2.1.1.3 f-slope

The f-slope is the plot of the solution norm $\|x_\lambda\|_2$ versus $\mu = \ln(1/\lambda)$ [Wu, 2003]. By using the inverse of λ in the x axis the values are kept positive. This method involves finding the "insensitive" or flattest part of the curve and thereby selecting a value of λ which corresponds to the smallest difference between adjacent solution norms [Wu, 2003].

A discrete least squares problem, that partially satisfies the DPC, has a regularised solution where the error $\|x_\lambda - x_{exact}\|_2$ is minimum if λ is close to σ_l , and where σ_i with $i > l$ decays faster than the data space coefficients $|u_i^T b|$. The exact solution x_{exact} is the unregularised solution to the unperturbed problem, $x_{exact} = A^I b_{exact}$, where A^I represents the pseudoinverse of the matrix A . The f-slope finds a value $\lambda = e^{-\mu}$ that minimises the derivative of the solution norm with respect to μ and the regularisation parameter given by this method should also be close to σ_l . In fact, this derivative suppresses contributions from σ_i/λ which are very large or small. This method has the drawback that it only analyses the solution norm from the least squares problem and not the data norm.

6.2.1.1.4 Generalised Cross-Validation (GCV)

The GCV is used in smoothing splines, where λ is the smoothing parameter, and in inverse problems where regularisation is needed. This method is based on the principle that, if a data point is omitted, then from the regularised solution obtained from this reduced data set, one should be able to estimate the missing data value [Wahba, 1977; Golub *et al.*, 1979]. The optimal regularisation parameter is the one that minimises the function [Hansen, 1998; Wahba, 1977; Golub *et al.*, 1979; Craven & Wahba, 1979; Wahba, 1990]

$$GCV(\lambda) = \frac{RN}{\tau^2} = \frac{\|Ax_\lambda - b\|_2^2}{(\text{trace}(I - AA_\lambda^I))^2}, \quad (6.8)$$

where RN is the residual norm. The denominator is the weighting factor of the function, which

can be written in terms of the filter factors as $\tau = \text{trace}(I - \sum_{i=1}^n f_i)$. The filter factors sum represents the degrees of freedom (df) for the signal, which basically is the number of singular values used in the regularised solution, hence the term τ represents the df for noise [Wahba, 1990]. The denominator favours residual norms whose regularised norm contains contributions from a small number of singular values, resulting in a smaller GCV value. Thus, minimising the GCV function provides a solution that can fit the data using the smallest possible number of parameters, thereby minimising the contribution from small singular values.

It has been shown [Craven & Wahba, 1979] that the GCV minimum is similar to the minimiser of the true mean square error $\|Ax_\lambda - b_{exact}\|_2^2$, which cannot be calculated if neither b_{exact} nor error variance are known. Although this seems to be a good argument, the GCV method relies only on the transition point of the residual norm, where its value starts increasing rapidly from an almost constant value, without considering the changes in the solution norm.

6.2.1.1.5 Quasi-Optimality Criterion (QOC)

The regularisation parameter is found by minimising the function [Hansen, 1998; Hanke & Hansen, 1993; Hanke & Raus, 1996; Kitagawa *et al.*, 2001]

$$Q_\lambda = \left\| \lambda^2 \frac{dx_\lambda}{d(\lambda^2)} \right\|_2 = \left(\sum_{i=1}^n \left(f_i (1 - f_i) \frac{u_i^T b}{\sigma_i} \right)^2 \right)^{1/2}. \quad (6.9)$$

The iterated Tikhonov regularisation computes successive approximations to the solution whose accuracy increases at each iteration. If $x^{(0)}$ is the starting vector, which can be any vector but usually $x^{(0)} = 0$, then the iterative solution is [Golub, 1965]

$$x^{(q+1)} = (A^T A + \lambda^2 I)^{-1} (A^T b + \lambda^2 x^{(q)}). \quad (6.10)$$

If the solution of the first iteration $x^{(1)}$ is defined to be the regularised solution x_λ , then the second iterate is given by

$$(A^T A + \lambda^2 I)^{-1} (A^T b + \lambda^2 x_\lambda) = x_\lambda - \lambda^2 \frac{dx_\lambda}{d(\lambda^2)}, \quad (6.11)$$

where

$$\frac{dx_\lambda}{d(\lambda^2)} = - (A^T A + \lambda^2 I)^{-1} x_\lambda. \quad (6.12)$$

From equation (6.11) it is easy to see that the minimisation of Q_λ minimises the correction to x_λ in the solution. Note that when the correction applied to the iterated Tikhonov approximation x_{it} is small this means that the solution is closer to the exact solution, hence $\|x_\lambda - x_{it}\| \approx \|x_\lambda - x_{exact}\|$. This method is similar to the f-slope method in the sense that it is also based on the error estimate of $\|x_\lambda - x_{exact}\|$.

6.2.1.2 Methods that use noise statistics

6.2.1.2.1 Unbiased Predictive Risk Estimator (UPRE)

The UPRE method seeks to minimise the predictive risk defined as follows [Vogel, 2002]

$$\frac{1}{m} \|p_\lambda\|_2^2 = \frac{1}{m} \|Ax_\lambda - Ax_{exact}\|_2^2. \quad (6.13)$$

Writing the Tikhonov regularised solution as $x_\lambda = A_\lambda b$, the solution error \mathbf{e}_λ can be defined as

$$\mathbf{e}_\lambda = x_\lambda - x_{exact} = (A_\lambda A - I) x_{exact} + A_\lambda \mathbf{e}. \quad (6.14)$$

The predictive error can be expressed as

$$p_\lambda = A\mathbf{e}_\lambda = (AA_\lambda - I) Ax_{exact} + AA_\lambda \mathbf{e}. \quad (6.15)$$

The influence matrix can be defined as

$$B_\lambda = AA_\lambda. \quad (6.16)$$

Inserting B_λ into equation (6.15), then the predictive error becomes

$$p_\lambda = (B_\lambda - I) Ax_{exact} + B_\lambda \mathbf{e}. \quad (6.17)$$

The regularised residual is defined as

$$r_\lambda = Ax_\lambda - b = (B_\lambda - I) Ax_{exact} + (B_\lambda - I) \mathbf{e}. \quad (6.18)$$

Comparing equation (6.18) with (6.17) one can verify that there is a new term $I\mathbf{e}$. We assume that B_λ is symmetric and that the noise signal \mathbf{e} is white noise, hence the expected value $E(\mathbf{e}) = 0$ and for $1 \leq i, j \leq m$

$$E(\mathbf{e}_i, \mathbf{e}_j) = \begin{cases} \sigma_{ij}^2, & i=j \\ 0, & i \neq j \end{cases}, \quad (6.19)$$

where σ^2 is the variance. The expected value of the mean squared norm of r_λ is thus given by

$$E\left(\frac{1}{m} \|r_\lambda\|_2^2\right) = \frac{1}{m} \|(B_\lambda - I) Ax_{exact}\|_2^2 + \frac{\sigma^2}{m} \text{trace}(B_\lambda^2) - \frac{2\sigma^2}{m} \text{trace}(B_\lambda) + \sigma^2. \quad (6.20)$$

Note that the two first terms of this equation are the expected value of the predictive risk. The UPRE is defined as

$$U_\lambda = \frac{1}{m} \|r_\lambda\|_2^2 + \frac{2\sigma^2}{m} \text{trace}(B_\lambda) - \sigma^2. \quad (6.21)$$

The UPRE method selects the regularisation parameter that minimises U_λ . This method depends on the accuracy of the estimate of the noise variance. The minimiser of U_λ is not

necessarily the same minimiser of the predictive risk and it does not ensure that the solution error is small.

6.2.1.2.2 Discrepancy Principle (DP)

The regularisation parameter obtained with the DP method satisfies the following condition [Vogel, 2002]

$$DP = \frac{1}{m} \|Ax_\lambda - b\|_2^2 - \sigma^2 \approx 0. \quad (6.22)$$

If it is assumed that $x_{exact} \approx x_\lambda$, then $\|Ax_\lambda - b\|_2^2 \approx I\mathbf{e}$ and therefore $E\left(\frac{1}{m} \|Ax_\lambda - b\|_2^2\right) \approx \sigma^2$. This method also depends on the quality of the noise statistics; in principle if the noise variance is known the method should find accurate solutions, whereas when variance is only an estimate the DP method does not necessarily return $\lambda \approx \lambda_{opt}$, where λ_{opt} is the optimal regularisation parameter.

6.2.1.2.3 Normalised cumulative periodogram (NCP)

This method favours the regularisation parameter for which the residual vector resembles white noise [Hansen *et al.*, 2006]. It is derived from the periodogram, which is the power spectrum of the residual and is obtained by taking the squares of the absolute values of the discrete Fourier transform (dft) for half of the residual vector length. Given a vector r , which here represents the residual vector of length N , the periodogram is computed as

$$P_k = |dft(r)_k|^2, \quad k = 0, \dots, N/2. \quad (6.23)$$

The normalised cumulative periodogram (NCP) is the cumulative periodogram vector normalised by the sum of its elements

$$NCP_k = \frac{\sum_{i=0}^k P_i}{\sum_{i=0}^{N/2} P_i}. \quad (6.24)$$

If the residuals are pure white noise then the power spectrum is flat, since the variance is uniform, and the NCP is a straight line from the coordinate point (0,0) to (0.5,1). The selected regularisation parameter is the one that minimises the distance of the NCP to a straight line [Hansen *et al.*, 2006]. Therefore, unlike the UPRE and DP methods, one does not need to directly estimate the noise variance.

Recall that the residual vector is

$$r_\lambda = Ax_\lambda - b = \sum_{i=1}^n u_i(1 - f_i)u_i^T b = U(I - F)U^T b, \quad (6.25)$$

and that white noise is invariant to orthogonal and unitary transformations, such as the Fourier transform. Recall that, given a random white noise vector e , the covariance is given by $cov(e) = \sigma^2 I$, where σ^2 is the variance. By applying the dft to the white noise vector will not change its covariance, $cov(dft(e)) = \sigma^2 I$, and since U is an orthogonal matrix then, also $cov(U^T e) = \sigma^2 I$. If the data vector $b = b_{exact} + e$, hence the covariance of $dft(r_\lambda)$ will be

$$\begin{aligned} \text{cov}(dft(r_\lambda)) &= \text{cov}(dft(U(I - F)U^T b)) \\ &= U(I - F)U^T \text{cov}(e)U(I - F)U^T = \sigma^2 U(I - F)^2 U^T. \end{aligned} \quad (6.26)$$

From the previous analysis on filter factors (refer to §6.1), when λ is too small the solution is dominated by high frequency noise components. When the regularisation parameter is optimal λ_{opt} the residual is white-noise with the low-frequencies removed, i.e., high-pass filtered white noise.

6.2.2 Image based methods

6.2.2.1 Fixed Noise Figure (NF)

The Fixed Noise Figure is defined as the ratio between the signal to noise ratio of the measurements and the signal to noise ratio of the image [Graham & Adler, 2006; Adler & Guardo, 1996; Adler, 2004]

$$NF = \frac{SNR_{data}}{SNR_{image}} = \left| \left(\frac{\bar{b}}{\sigma_b} \right) \right| / \left| \left(\frac{\bar{x}_\lambda}{\sigma_{x_\lambda}} \right) \right|, \quad (6.27)$$

where \bar{x}_λ represents the mean value of the solution image, σ_{x_λ} is the standard deviation of image, and \bar{b} and σ_b represent the equivalent parameters for data. The regularisation parameter is found by plotting NF as a function of λ , and selecting the optimal NF value whose λ returns the most acceptable solution. The NF gives a measure of the image reconstruction performance, and the curve allows a comparison between different regularised solutions.

In principle the selected NF values for experimental data should be ≤ 1 , since Tikhonov regularisation smooths the image, eliminating high-frequency components and hence reducing its standard deviation. This method replaces the selection of a regularisation parameter as in the Heuristic method, by the selection of a fixed NF value. The advantage is that for the same experimental set-up one can use the same NF value which should result in good reconstructions for all the data.

6.2.2.2 Blur Radius (BR)

Adler and Guardo [1996] defined the Blur Radius as a measure of the resolution achieved by electrical impedance tomography (EIT) using the following expression

$$BR = \frac{r_z}{r_0} = \sqrt{\frac{A_z}{A_0}}, \quad (6.28)$$

where r_0 and A_0 are the total image radius and area respectively, and r_z and A_z are the radius and area of the region containing half the total image amplitude. In a mesh structure each element has a distinct area A_{el} , calculated from the position of the vertices. Thus $A_0 = \sum A_{el}$ and $A_z = \sum A_{el > \delta}$, where the threshold δ is defined so that only areas of elements with intensity above half the maximum intensity contribute to the summation. The minimum value of

the BR function corresponds to the maximum resolution and hence to the optimal regularisation parameter λ .

For images with square pixels, where each element has equal area, this equation can be expressed as [Wheeler *et al.*, 2002; Avis & Barber, 1994]

$$BR = \sqrt{\frac{p_{>FWHM}}{p_t}}, \quad (6.29)$$

with p_{FWHM} representing the number of pixels with intensity above half the maximum intensity and p_t is the total number of pixels. This measure of resolution has a limitation, since it depends on the existence of a single point-like feature in the image so that the Full Width Half Maximum (FWHM) of the point-spread function (PSF) can be found.

When contrast is low or no isolated feature exists, instead of using FWHM that may lead to a erroneous result, the threshold can then be defined as the background value, and BR becomes

$$BR2 = \sqrt{\frac{p_{>bkq}}{p_t}}, \quad (6.30)$$

where now $p_{>bkq}$ is the number of pixels with intensity above the background intensity. This no longer relies on an isolated feature: all the pixels that appear with a different intensity from that of the background are counted. A disadvantage of this is that it considers every variation from background as being part of the signal, even though it could be noise.

6.2.2.3 Full Width Half Maximum

This method is applicable to image data containing a single isolated region of contrast. The FWHM of the region is calculated for different regularisation parameters. The FWHM is found by selecting all the pixels of the row/column that include the pixel with maximum value in the image (interpolating where necessary and/or rejecting non-correlated pixels). It is a figure of merit much like the BR; the major difference between these two methods is that while the BR includes all the pixels in the image with intensity above half maximum intensity the FWHM only includes pixels from a row/column in the image. By considering the FWHM for all the pixels the BR method has higher probability of including pixels that represent noise. For a three-dimensional image the Full Volume Half Maximum (FVHM) is used, which is considered to be $FVHM = FWHM_x \cdot FWHM_y \cdot FWHM_z$.

Regińska [1996] has shown that the minimum of the function

$$\Psi_\alpha(\lambda) = \|x\|_2 \cdot \|Ax_\lambda - b\|_2^\alpha, \text{ where } \alpha > 0, \quad (6.31)$$

is similar to the point with maximum curvature of the L-curve. It can be also useful to analyse the behaviour of $FWHM_w = FWHM \cdot \Psi_\alpha$, where both quantities are normalised to vary from 0 to 1. The optimal λ corresponds to the minimum value of this function, where one expects to achieve a compromise between the size of the feature, the noise and accuracy of the solution, i.e., FWHM should represent the true size of the feature, with minimal noise and blurring. The

function Ψ_α can be considered to be a measure of accuracy, or a weighting factor; if FWHM is small then Ψ_α needs to be small. A value of α is chosen to force the quantities to be of the same magnitude.

6.2.2.4 Contrast Ratios

The first of two contrast ratios described here involves plotting the contrast to noise ratio (CNR) as a function of the regularisation parameter [Culver *et al.*, 2003], where noise is the solution norm and contrast is defined as

$$C = \frac{I_{max} - I_{bkg}}{I_{bkg}}, \quad (6.32)$$

where I_{max} is the peak value of the image and I_{bkg} is averaged intensity of the background. One seeks the regularisation parameter that results in a maximum contrast to noise ratio.

Using the function Ψ_α as a weighting factor, following the same arguments presented for the previous method, a second ratio can be defined $C_w = \text{CNR} \cdot \Psi_\alpha^{-1}$, which should have a maximum value whose corresponding λ compromises between high CNR and accuracy of the solution.

A potential disadvantage of this method is that it favours contrast over resolution.

6.2.2.5 Criteria

This study seeks a method for selecting the regularisation parameter for optical topography. We begin by showing and describing the use and performance of the various methods, but reject some immediately because they are not suitable for optical topography. Our criteria are:

1. The method should not require any subjective input from the user.
2. The method should only require knowledge which is available during clinical optical topography. For example, it should not require knowledge of the size of the feature being examined.
3. The method should not rely on the presence of particular features in the image. For example, it should not assume there is a single, spatially isolated change.

6.3 Regularisation parameter selection

The above methods were initially applied to a deblurring problem. This is a test problem for which the solution is known, which is simple to generate and is ill-posed. Thereafter, the same methods were applied to experimental optical topography data, and the consistency between the results for the two problems were examined.

6.3.1 Deblurring

Deblurring is an example of an ill-posed inverse problem. We start with an image which is blurred and noisy, and we perform deblurring and denoising processes to estimate the true image. For the deblurring problem A is the blurring matrix, x is the original image, b is the blurred image and finally δ is additive Gaussian noise. To generate the matrix A , and the images x and b , we used the function *blur* from the Matlab package Regularisation Tools [Hansen, 1994, 2007]. The image x has 50×50 pixels (see figure 6.3). The matrix A is basically a

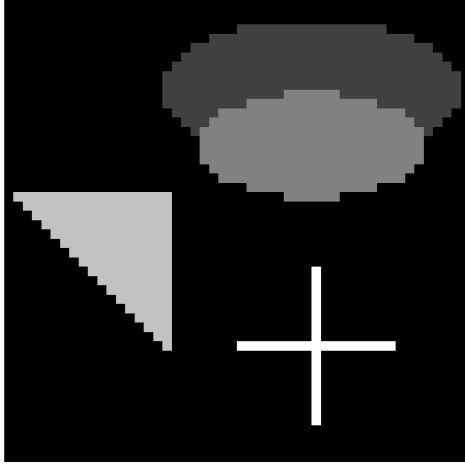


Figure 6.3: Original test image.

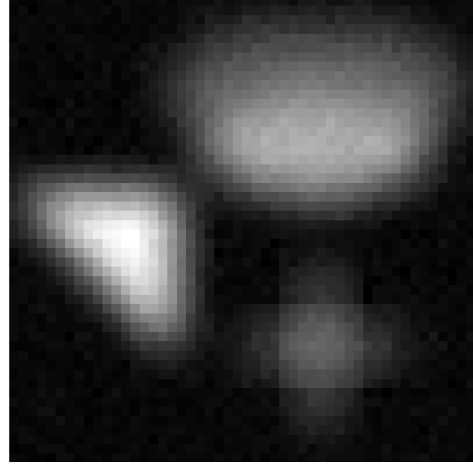


Figure 6.4: Blurred image with 5% added Gaussian noise.

symmetric matrix $50^2 \times 50^2$ with Toeplitz blocks, which is determined from $A = T \otimes T$, where \otimes represents the Kronecker product and T is a Toeplitz matrix that contains elements of a Gaussian point spread function with variance σ^2 , which models the blurring effect [Hansen, 2002].

Definition 2 A Toeplitz matrix T is a $n \times n$ matrix, where its elements $t_{ij} = t_{i-j}$, i.e., it is a matrix of the form

$$T = \begin{pmatrix} t_0 & t_{-1} & t_{-2} & \dots & t_{-(n-1)} \\ t_1 & t_0 & t_{-1} & \dots & t_{-(n-2)} \\ t_2 & t_1 & t_0 & \dots & t_{-(n-3)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ t_{n-1} & t_{n-2} & t_{n-3} & \dots & t_0 \end{pmatrix}.$$

The advantage of this structure, in particular when dealing with large-scale problems, is that it can be represented by $2n - 1$ different elements instead of storing n^2 elements. If the matrix elements become $t_i = 0$ for $|i| > l$, T is termed a *banded* Toeplitz matrix, with band l that defines the number of diagonals, with nonzero elements, from the main diagonal that are stored in the matrix T . The Toeplitz matrix used in the *blur* function is symmetric, hence it contains information in n different elements.

Definition 3 The Kronecker product of a $n \times n$ matrix T and a $\bar{n} \times \bar{n}$ matrix \bar{T} is a $n\bar{n} \times n\bar{n}$ matrix with the block structure

$$T \otimes \bar{T} = \begin{pmatrix} t_{11}\bar{T} & \dots & t_{1n}\bar{T} \\ \vdots & \ddots & \vdots \\ t_{n1}\bar{T} & \dots & t_{nn}\bar{T} \end{pmatrix}.$$

We added different levels of Gaussian noise to the blurred image b , from 5% to 40% and for 500 noise realisations each. For example, if 5% Gaussian noise is added to b , this means that $\delta = 0.05 * \frac{\|b\|_2}{\sqrt{n*n}} * randn(n * n, 1)$. For this test we have set $\sigma = 3$ and $band = 5$ (figure 6.4).

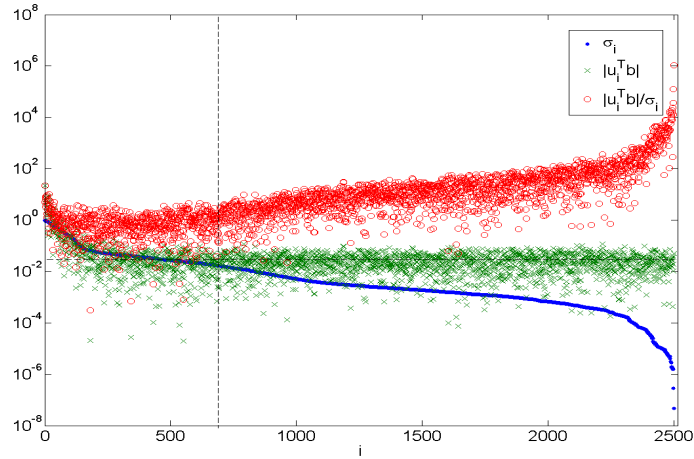


Figure 6.5: Picard condition for the deblurring problem, where the vertical dashed line marks the beginning of $|U_i^T b| < \sigma_i$ and the horizontal line represents the noise level. DPC is satisfied for $i \lesssim 695$.

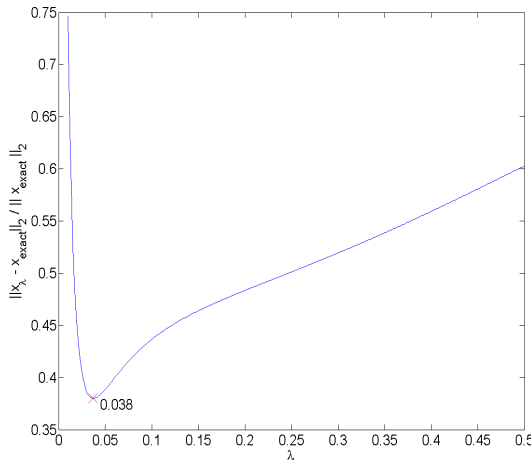


Figure 6.6: Relative error for the regularised solution for a single realisation of 5% noise. The λ_{opt} value is marked by the cross.

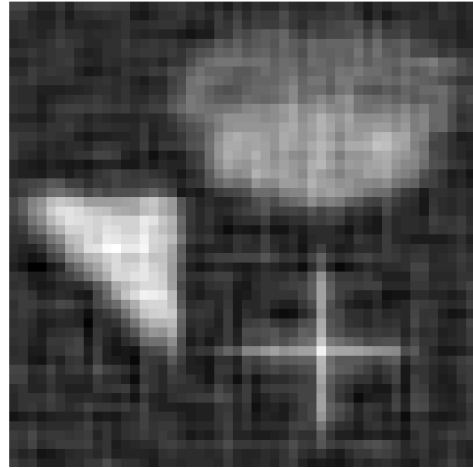


Figure 6.7: Reconstructed image with $\lambda_{opt} = 0.038$.

To deblur the image, i.e., to perform the inverse problem, we need to use the properties of the Kronecker product. If the SVD of T is $T = U_b S_b V_b^T$, then the approximate solution \hat{x} to the original image x is given by the following equation

$$\hat{x} = V_b \frac{S_t (U_b^T b U_b)}{S_t^2 + \lambda^2} V_b^T, \quad (6.33)$$

where $S_t = \text{diag}(S_b) \text{diag}(S_b^T)$.

Figure 6.5 shows that the DPC is satisfied for $i \lesssim 695$. For higher values of i the DPC is no longer fulfilled and $|U_i^T b|$ reaches the noise level. Since the DPC is at least partially satisfied it means that we can find a solution which approximately recovers the real solution.

The unblurred and noise-free image of the test problem is known, and using this information we can find the regularised solution x_λ which is closest to the real solution x_{exact} .

method	λ	ε
<i>Optimal</i>	0.038	0.380
Heuristic	[0.02,0.08]	0.448
L-curve	0.031	0.387
f-slope	0.071	0.411
Fixed NF	[0.02,0.08]	0.448
BR	0.005	1.490
BR2	0.035	0.381
GCV	0.031	0.385
QOC	0.122	0.450
FWHM	[0.035,0.082]	0.422
FWHM $\cdot \Psi$	0.019	0.460
CNR	0.012	0.638
CNR $\cdot \Psi^{-1}$	0.019	0.460
DP	0.052	0.391
UPRE	0.031	0.385
NCP	0.055	0.481

Table 6.1: Regularisation parameter λ and relative error ε obtained using the selection methods for the deblurring problem for a single noise realisation of 5% noise.

In order to do so, we define the optimal regularisation parameter λ_{opt} to be the one that minimises the relative error $\varepsilon = \|x_\lambda - x_{exact}\|_2 / \|x_{exact}\|_2$. Figure 6.6 shows the relative error for different regularisation parameters, for a single noise realisation of 5% noise, and the minimum which occurs for $\lambda_{opt} = 0.038$. The corresponding deblurred and denoised image is shown in figure 6.7.

6.3.2 Results

The results obtained for the deblurring problem are described for each selection method. First the results for this problem with 5% noise are shown and described in order to illustrate the features of each method. The results and the relative errors are presented in table 6.1. Thereafter, the results for different levels of added noise are analysed, but only for the methods that satisfy the three criteria defined in §6.2.2.5. The regularisation parameter was chosen from a set of 1000 logarithmically spaced points between 10^{-4} and $10^{-0.3}$.

6.3.2.1 Example of methods and results - 5 % noise

Heuristic method Images were generated for different values of λ . It was not easy to identify a single value, and we conclude that acceptable results are obtained for $\lambda \in [0.02, 0.08]$.

L-curve The L-curve exhibits a point of maximum curvature which is easily identifiable (figure 6.8(a)). Using equation (6.7) we can find the curvature for different regularisation parameters, which is illustrated in figure 6.8(b). The point of maximum curvature is found at $\lambda = 0.031$.

f-slope The f-slope curve clearly shows a flat part which is less sensitive to perturbations (figure 6.9). We search on this part of the curve for the smallest slope, which occurs at $\lambda = 0.071$.

Fixed NF The $\log(NF)$ decreases with increasing λ (see figure 6.10). We selected 5 points on the curve with values approximately between $NF=1$ and $NF=2$. We use the corresponding regularisation parameters to deblur the images, the reconstructed image whose relative error is smallest corresponds to the optimal value of NF, which was obtained for $NF=1.28$ ($\lambda = 0.04$). But we know already from the Heuristic method that $\lambda \in [0.02, 0.08]$ which corresponds to $NF \in [1.22, 1.43]$. Here we have used information from another method to find NF, but as described previously the NF value is usually found first and then the images are reconstructed for the corresponding fixed NF values.

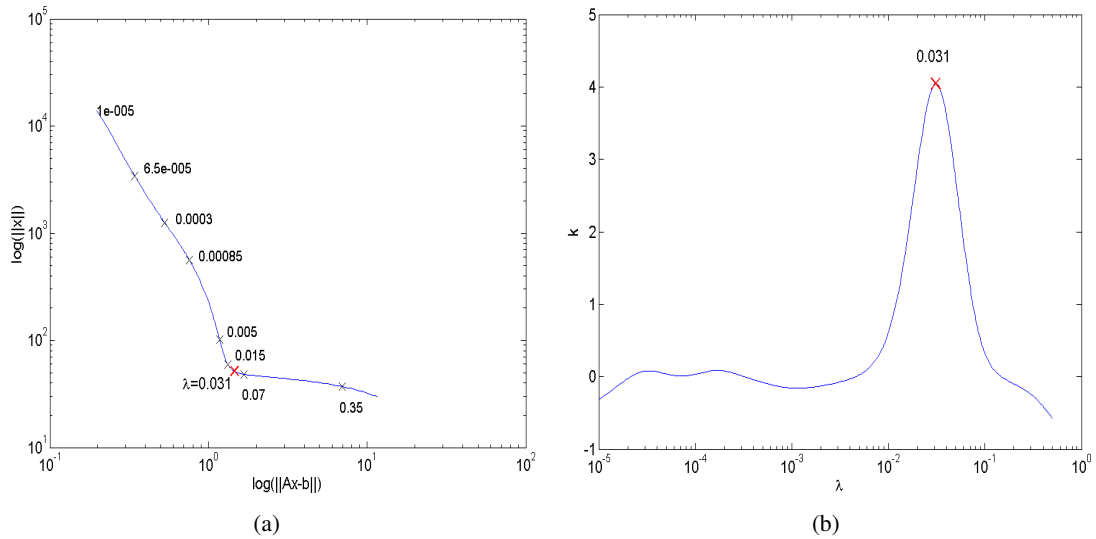


Figure 6.8: (a) L-Curve method and (b) curvature for different values of λ for the deblurring example.

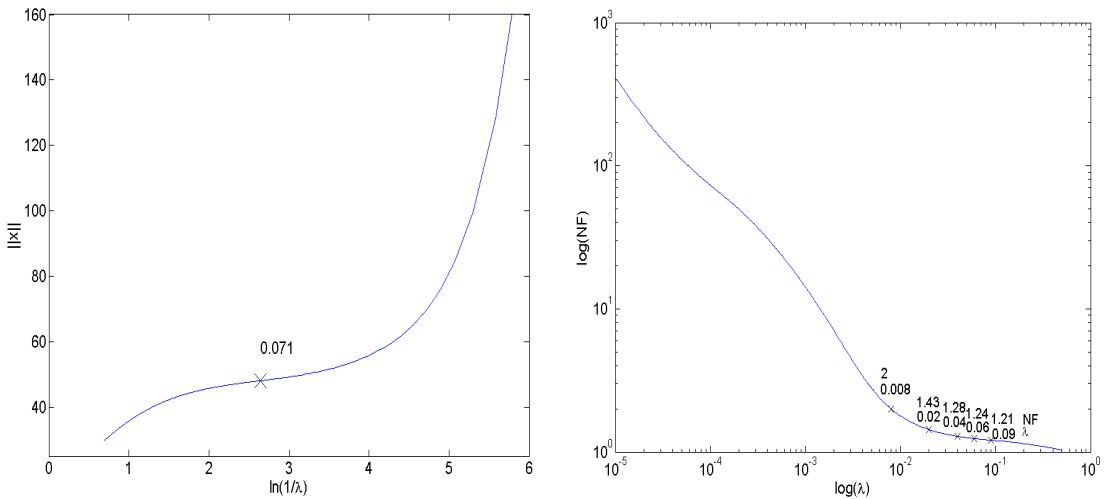


Figure 6.9: f-slope as a function of $\ln(1/\lambda)$.

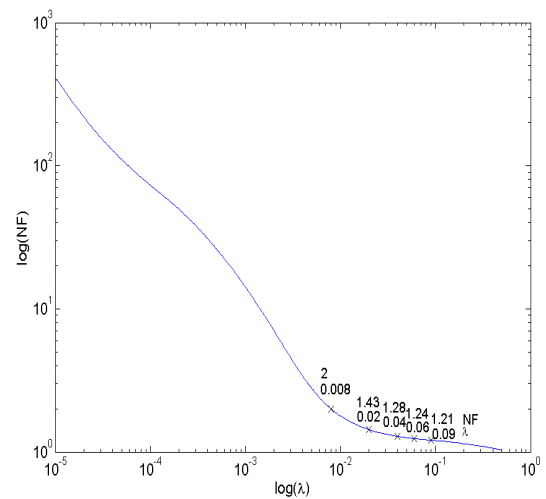


Figure 6.10: Fixed NF as a function of λ . For \times the top numbers are the NF values and the bottom numbers the corresponding λ value.

BR We calculated the BR using equation (6.29) and found several minima (see figure 6.11(a)). Although the local minimum at $\lambda = 0.017$ results in an acceptable deblurred image, since this minimum is not so prominent we select $\lambda = 0.005$ as the result for this method (figure 6.11(a)), which is extremely small. When calculating BR using equation (6.30), the resolution of the image is better when the regularisation parameter has a value of 0.036, as shown in figure 6.11(b). We find that the second definition is more effective for this task.

GCV The minimum value for this function was easily identifiable, although it is located on a relatively flat region of the curve, as shown in figure 6.12. The minimum value for the GCV function is found at $\lambda = 0.029$.

QOC As shown in figure 6.13, the QOC function decreases with increasing λ and the minimum value for this method occurs for $\lambda = 0.12$, a value which is higher than the values found with the previous methods.

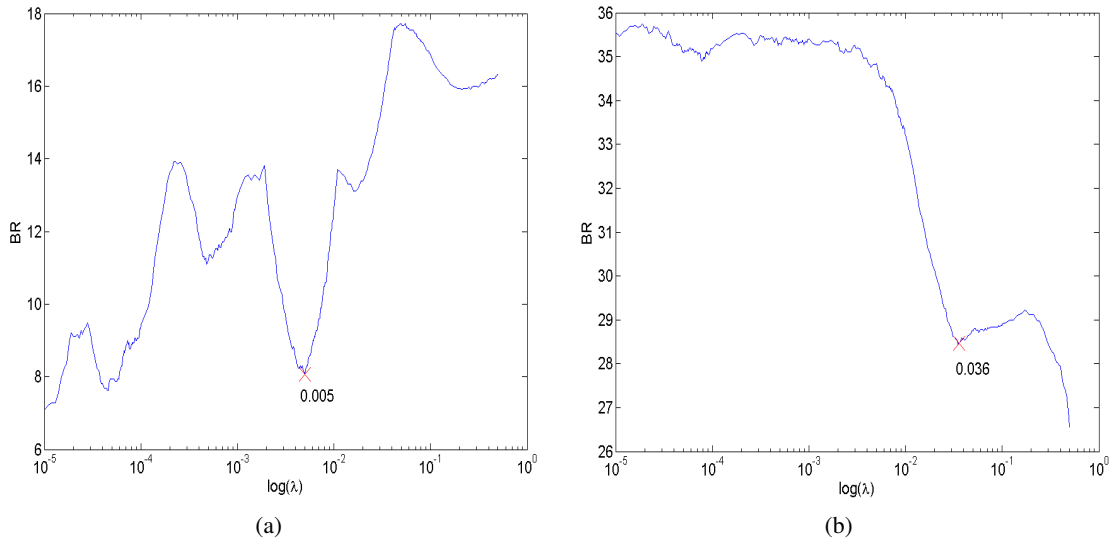


Figure 6.11: (a) Blur radius method calculated using the number of pixels with intensity above half maximum and (b) BR with number of pixels with intensity above the background intensity.

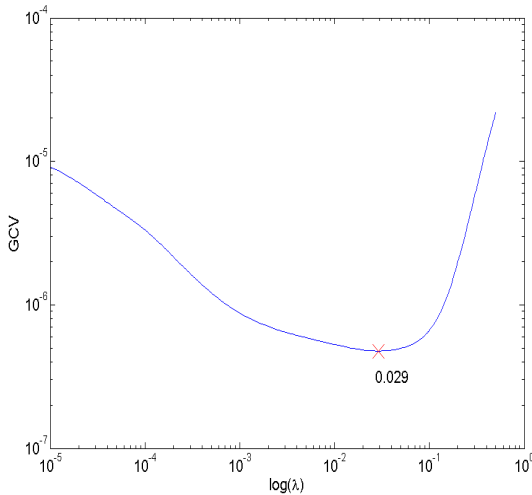


Figure 6.12: Generalised-cross validation function and its minimum.

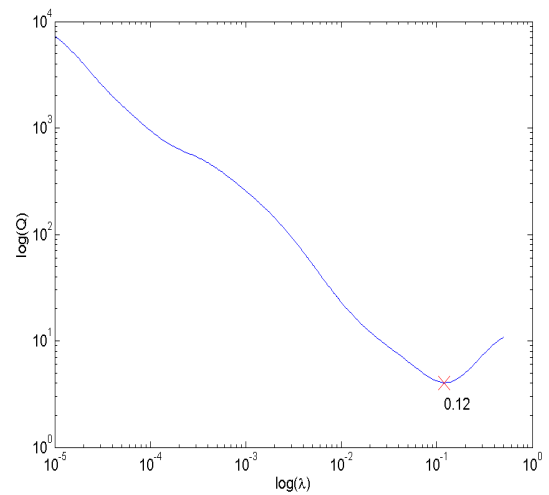


Figure 6.13: Quasi-optimality criterion and its minimum.

FWHM To evaluate this method we use the green ellipse to estimate the FWHM, where we selected all the pixels along the horizontal row of pixels through the centre of the ellipse. We compute the intensity by averaging 63 pixels around the central pixel of our region of interest (ROI).

The analysis of the FWHM together with the ROI intensity for different λ allows us to have a more qualitative idea of what the regularisation is actually doing to the image. From figure 6.14(a) it can be seen that it is not possible to identify a single regularisation parameter, but instead we can define a preferred range. For $\lambda \in [0.012, 0.091]$ the FWHM does not change significantly and its value is close to the real value (see figure 6.14(a)). On the other hand, as shown in figure 6.14(b), the intensity within the ROI is higher for $\lambda \in [0.035, 0.082]$. As the latter interval is contained in the former, the regularisation parameter should be selected from that interval.

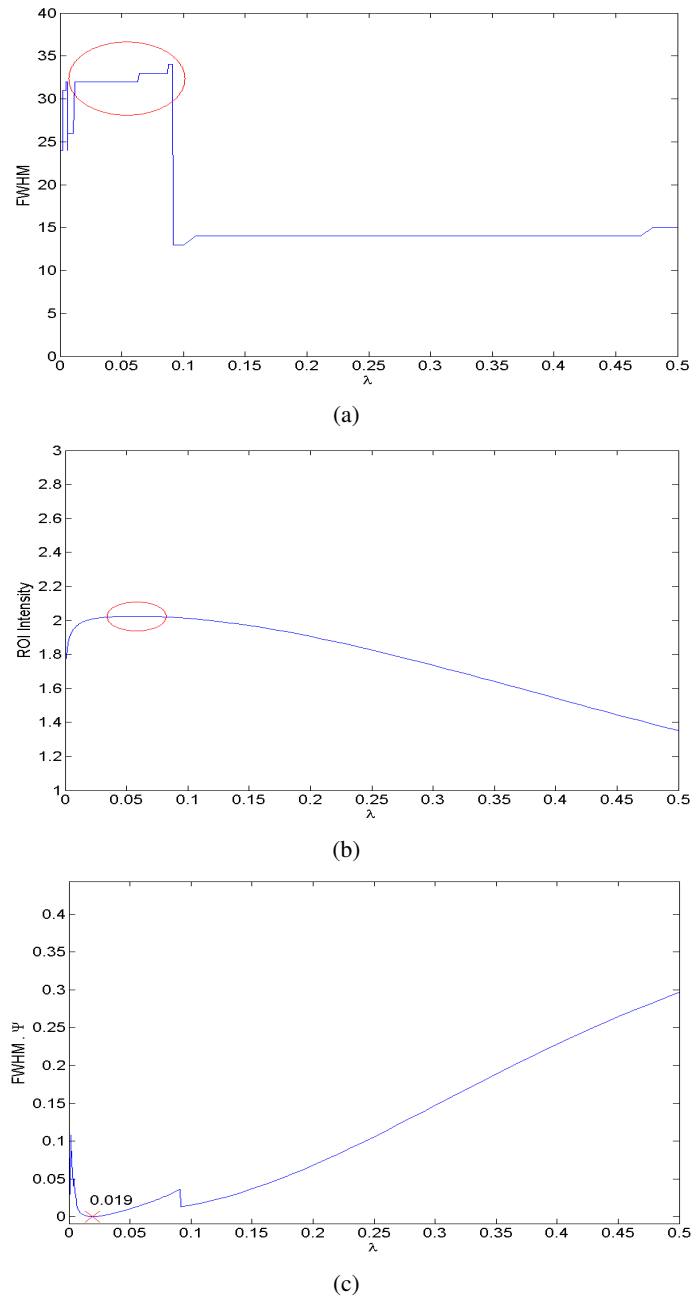


Figure 6.14: (a) FWHM, (b) maximum intensity of ROI and (c) weighted FWHM of the deblurred images. The circles represent the intervals that provide sensible results.

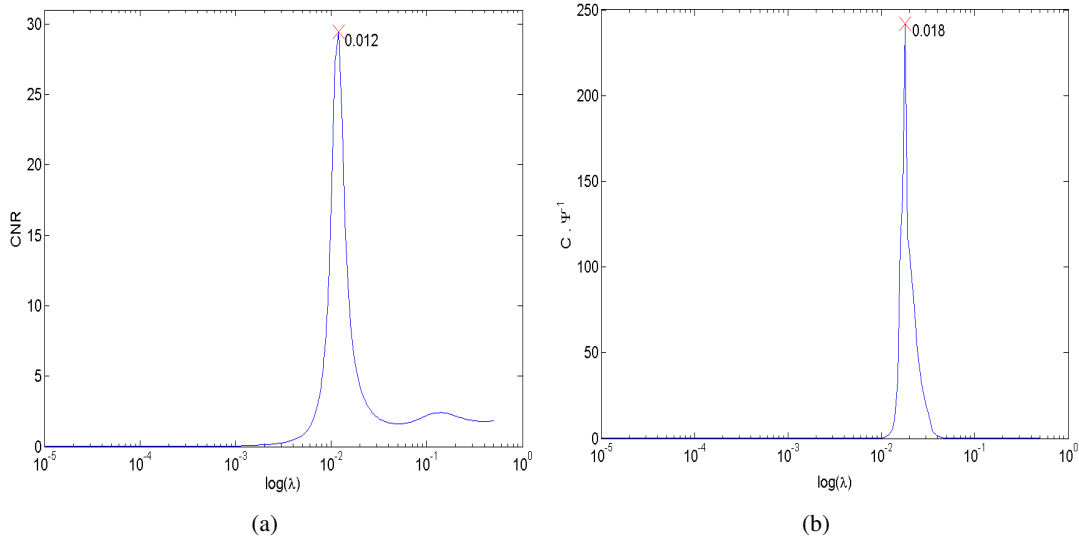


Figure 6.15: (a) Contrast to noise ratio curve with maximum located at $\lambda = 0.012$ and (b) C_w curve with maximum at $\lambda = 0.018$.

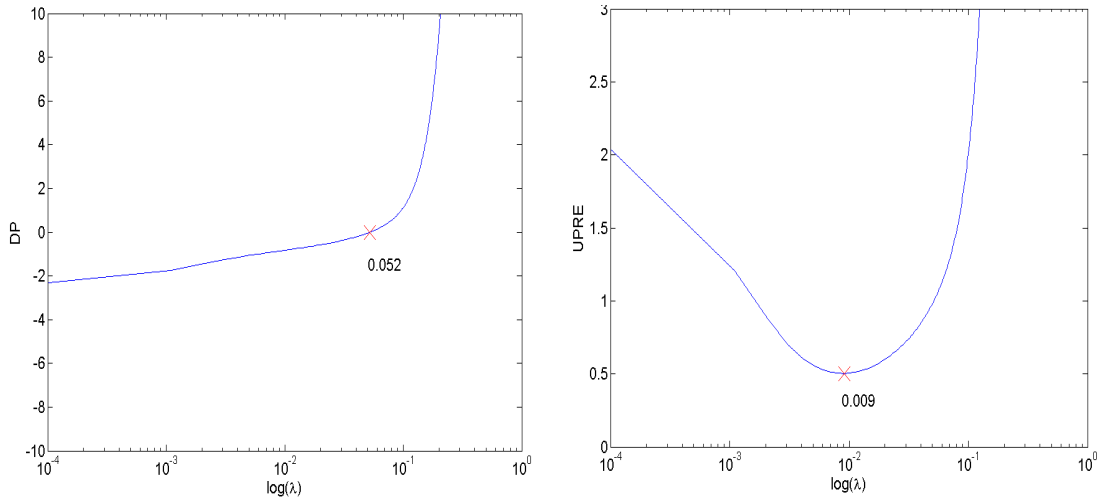


Figure 6.16: DP method where the cross marks the zero of the function.

Figure 6.17: UPRE function and its minimum.

Multiplying the FWHM by the solution norm and the residual norm, with α set to 0.58, we obtain a single λ value at 0.019 (figure 6.14(c)).

CNR The obtained CNR curve is shown in figure 6.15(a), where it is possible to locate a sharp global maximum at $\lambda = 0.012$. The value for I_{bkg} was found by averaging a single row of pixels located in one of the extremities of the image, which means that 50 pixels were included. Using the weighting factor Ψ_α we obtain the curve C_w and its maximum point occurs for $\lambda = 0.019$, which can be seen in figure 6.15(b). These two curves and their results have shown to be very similar.

DP The residual norm equals the noise variance at $\lambda = 0.052$ (figure 6.16). The noise variance was approximated from the noise norm as $\sigma^2 \simeq \frac{1}{m} \|\mathbf{e}\|_2^2$.

UPRE Figure 6.17 shows the UPRE function which is minimised for $\lambda = 0.009$. To calculate the function UPRE it is necessary to find the $\text{trace}(B_\lambda)$ which is given by the filter factors.

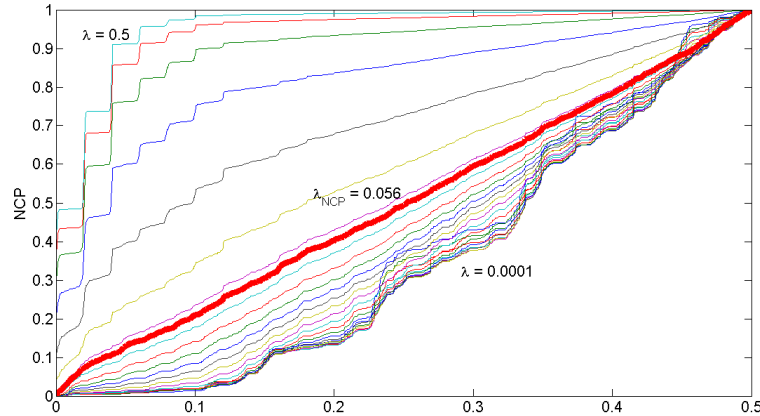


Figure 6.18: Normalised cumulative periodograms for different regularisation parameters. The red straight line is the NCP for the λ selected using the NCP method.

Finding the SVD of the full blur matrix is computational intensive, hence the singular values are calculated from SVD of matrix T as $S = \text{diag}(U_b \otimes U_b)^T A (V_b \otimes V_b)$.

NCP The residual vector resembles white noise when the regularisation parameter $\lambda = 0.056$. The NCP function for this λ value is represented by the red straight line in figure 6.18. When the regularisation parameter is small the residual becomes dominated by high frequency signals, whereas for larger λ low frequency components dominate.

6.3.2.2 All noise levels and realisations

The DPC is partially satisfied for all the different noise levels, which indicates that if the regularisation parameter λ is accurately selected then the regularised solution x_λ will approximate the exact solution x_{exact} .

The BR, Fixed NF and FWHM methods do not satisfy the criteria defined in §6.2.2.5, therefore these are excluded from the following analysis. For all noise realisation, the mean optimal regularisation parameters, and corresponding standard deviations, for the nine selection methods are shown for different noise levels in table 6.2. The corresponding relative errors calculated using the mean values in table 6.2 are shown in table 6.3 and in figure 6.19.

Heuristic Although the heuristic method is subjective and non-repeatable, we retain this method in our analysis as a measure against which to compare other more objective methods. In every noise case, the selected range using this method included λ_{opt} . The corresponding ranges of errors are quote in table 6.3. Note that for better comparison with the other methods, we chose to display its central value and error, where the latter gives the range limits. The range of λ included λ_{opt} , however the regularisation parameter errors are consistently higher than for most of the other methods, illustrating the irreproducibility of the heuristic method.

L-curve The L-curve did generally exhibit a single, easily identifiable point of maximum curvature. The predicted λ agree closely with λ_{opt} . For the deblurring problem, an image reconstructed with 5 % noise was slightly undersmoothed, with a visible noise component. However, all the results were acceptable, as confirmed by the small relative errors (table 6.3) and by the small regularisation parameter error (table 6.2). The latter reflects the small sensitivity

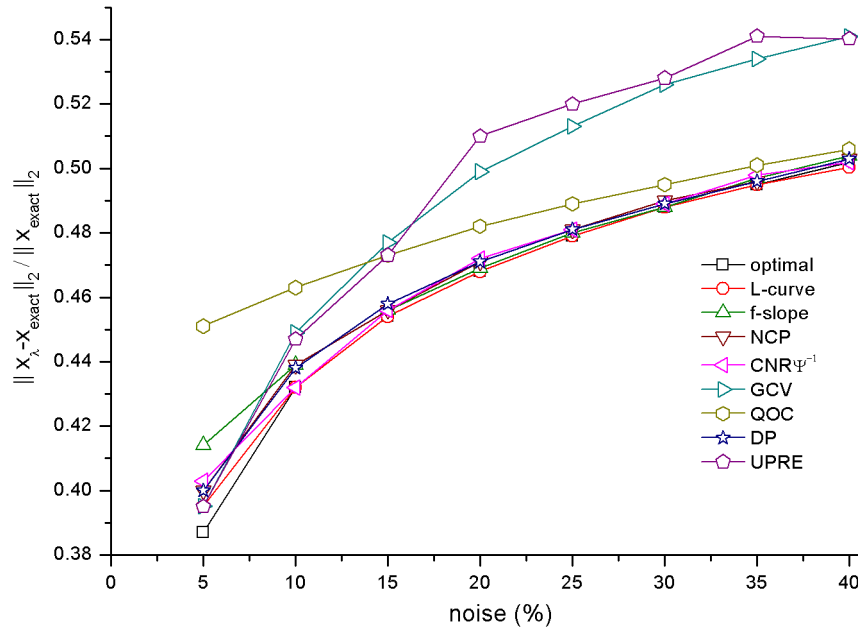


Figure 6.19: Mean relative error of λ_{opt} (\square), L-curve (\circ), f-slope (Δ), NCP (∇), $\text{CNR} \cdot \Psi^{-1}$ (\diamond), GCV (\triangleright), QOC (\circ), DP (\star) and UPRE (\diamond) methods for different noise levels.

of this method to perturbations of similar nature in the data.

f-slope This method seems to overestimate λ at low noise levels but selects λ close to λ_{opt} at higher noise levels.

GCV The selected λ was similar to λ_{opt} at low noise levels, 5% and 10% noise, but was rather low at higher noise levels.

QOC The QOC regularisation parameter prediction is much higher than λ_{opt} , however the relative error becomes closer to that of the optimal solution relative error ε_{opt} as the noise increases.

CNR The CNR method, for all noise levels, yields regularisation parameters which are lower than λ_{opt} and generated much greater relative errors than for any of the other methods, and the errors associated with λ are as large as those corresponding to the heuristic methods. The addition of the Ψ parameter led to λ which were closer to λ_{opt} than for CNR alone. The method still tended to underestimate λ for low noise levels, but for higher noise levels, the selection of λ was good. Note that the regularisation parameter uncertainty is particularly large for noise in the data between 15% and 25%.

DP This method appeared to overestimate values of λ ; nevertheless its relative error $\varepsilon_{DP} \approx \varepsilon_{opt}$ particularly when noise $\geq 15\%$.

UPRE The performance of UPRE appears to be similar to that of GCV, with good estimates of λ at low noise levels but an underestimated λ at higher noise levels.

NCP This method and the DP have similar performances. However, this method seems to be more sensitive to the different noise realisations, in particular for noise levels above 25%.

Method	5%	10%	15%	20%
Optimal	0.039 ± 2.6%	0.066 ± 3.3%	0.090 ± 4.0%	0.111 ± 4.2%
Heuristic	0.050 ± 60.0%	0.125 ± 60.0%	0.130 ± 53.8%	0.190 ± 57.9%
L-Curve	0.030 ± 2.7%	0.065 ± 1.7%	0.090 ± 1.6%	0.110 ± 1.5%
GCV	0.030 ± 4.0%	0.047 ± 5.1%	0.060 ± 5.6%	0.071 ± 5.8%
UPRE	0.030 ± 3.3%	0.048 ± 4.4%	0.062 ± 4.4%	0.073 ± 4.8%
DP	0.054 ± 3.2%	0.087 ± 3.2%	0.113 ± 3.3%	0.133 ± 3.1%
NCP	0.056 ± 6.4%	0.088 ± 7.6%	0.113 ± 8.5%	0.134 ± 8.7%
f-slope	0.070 ± 1.7%	0.089 ± 1.7%	0.104 ± 1.8%	0.116 ± 1.7%
QOC	0.121 ± 1.5%	0.142 ± 1.7%	0.160 ± 1.8%	0.176 ± 1.7%
CNR	0.014 ± 35.0%	0.018 ± 57.2%	0.025 ± 60.0%	0.033 ± 61.8%
CNR · Ψ^{-1}	0.022 ± 3.6%	0.049 ± 4.5%	0.077 ± 13.2%	0.105 ± 11.3%
Method	25%	30%	35%	40%
Optimal	0.131 ± 4.5%	0.148 ± 4.2%	0.164 ± 4.3%	0.178 ± 4.0%
Heuristic	0.190 ± 57.9%	0.300 ± 66.7%	0.300 ± 66.7%	0.350 ± 71.4%
L-Curve	0.128 ± 1.5%	0.145 ± 1.5%	0.159 ± 1.5%	0.172 ± 1.3%
GCV	0.081 ± 6.1%	0.090 ± 5.7%	0.099 ± 5.8%	0.108 ± 5.0%
UPRE	0.083 ± 5.1%	0.092 ± 4.8%	0.101 ± 4.9%	0.109 ± 5.0%
DP	0.151 ± 3.2%	0.167 ± 3.1%	0.181 ± 3.2%	0.195 ± 3.1%
NCP	0.152 ± 9.1%	0.168 ± 9.6%	0.183 ± 9.3%	0.195 ± 9.7%
f-slope	0.127 ± 1.8%	0.138 ± 1.7%	0.147 ± 1.8%	0.155 ± 1.7%
QOC	0.189 ± 1.7%	0.202 ± 1.7%	0.213 ± 1.6%	0.223 ± 1.7%
CNR	0.039 ± 55.4%	0.044 ± 58.7%	0.049 ± 53.4%	0.051 ± 56.4%
CNR · Ψ^{-1}	0.129 ± 9.7%	0.154 ± 2.4%	0.183 ± 2.8%	0.190 ± 2.8%

Table 6.2: Regularisation parameters λ obtained using the selection methods for the deblurring problem with different noise levels.

6.3.3 Discussion

As described in §6.3.1, the optimal regularisation parameter for the deblurring problem is $\lambda_{opt} = 0.038$. Most methods analysed assume that no prior statistical information about the noise is needed to find the regularisation parameter. Although there are methods that rely on the knowledge of the noise variance, which is known in the deblurring problem, in most real situations this value is unknown.

We note that the heuristic, the fixed NF, the BR and the FWHM methods do not meet the suitability criteria. They identify a range of regularisation parameters, rather than a single value. They are also subjective and non-repeatable, which means that the selected interval of values may depend on the person analysing the method and there is no way to define the extremes of this interval analytically. Also, if the same study is repeated at another time the selected values will probably not be exactly the same. In spite of this λ_{opt} is contained in the range of regularisation parameters selected.

The limitations of the L-curve method were analysed by Hanke [1996]. He showed that the L-curve method fails to find λ for very smooth solutions, i.e., when the SVD coefficients $|u_i^T b|$ decay very fast to zero, and therefore the optimal regularisation parameter occurs before the solution norm starts increasing, because a large number of SVD components have to be included in the solution to increase its norm significantly. Vogel [1996] has shown another non-convergence of λ which is related to the asymptotic behaviour of the L-curve as the dimensions of the problem increase. He pointed out that the regularised solution fails to converge to the true solution as the dimensions of the problem tend towards infinity. For the deblurring problem, an image reconstructed with 5 % noise is slightly undersmoothed, with a

Method	5%	10%	15%	20%
Optimal	0.387 ± 1.0%	0.432 ± 1.2%	0.454 ± 1.1%	0.468 ± 1.1%
Heuristic	0.393 ± 0.8%	0.457 ± 0.5%	0.463 ± 0.7%	0.485 ± 0.6%
L-Curve	0.395 ± 1.5%	0.432 ± 1.2%	0.454 ± 1.1%	0.468 ± 1.1%
GCV	0.395 ± 1.5%	0.433 ± 1.2%	0.477 ± 1.7%	0.499 ± 1.8%
UPRE	0.395 ± 1.5%	0.432 ± 1.2%	0.473 ± 1.7%	0.496 ± 1.8%
DP	0.400 ± 0.8%	0.438 ± 0.9%	0.458 ± 0.9%	0.471 ± 0.9%
NCP	0.400 ± 0.8%	0.439 ± 0.7%	0.456 ± 1.1%	0.471 ± 0.9%
f-slope	0.414 ± 0.5%	0.439 ± 0.9%	0.456 ± 0.9%	0.469 ± 1.1%
QOC	0.451 ± 0.2%	0.463 ± 0.4%	0.473 ± 0.5%	0.482 ± 0.6%
CNR	0.578 ± 2.8%	0.853 ± 2.3%	0.878 ± 2.4%	0.884 ± 2.4%
CNR · Ψ^{-1}	0.403 ± 1.5%	0.456 ± 1.1%	0.457 ± 0.9%	0.472 ± 0.9%
Method	25%	30%	35%	40%
Optimal	0.479 ± 1.0%	0.488 ± 1.0%	0.495 ± 1.2%	0.502 ± 1.0%
Heuristic	0.489 ± 0.6%	0.525 ± 0.6%	0.525 ± 0.6%	0.544 ± 0.6%
L-Curve	0.479 ± 1.0%	0.488 ± 1.0%	0.495 ± 1.2%	0.502 ± 1.2%
GCV	0.513 ± 1.9%	0.526 ± 1.9%	0.534 ± 1.2%	0.541 ± 1.9%
UPRE	0.510 ± 2.0%	0.520 ± 1.7%	0.528 ± 1.9%	0.541 ± 1.9%
DP	0.481 ± 0.8%	0.489 ± 0.8%	0.496 ± 1.0%	0.503 ± 1.0%
NCP	0.481 ± 0.8%	0.490 ± 0.8%	0.496 ± 1.0%	0.503 ± 1.0%
f-slope	0.480 ± 1.3%	0.488 ± 1.2%	0.497 ± 1.2%	0.504 ± 1.2%
QOC	0.489 ± 0.6%	0.495 ± 0.8%	0.501 ± 0.8%	0.506 ± 1.0%
CNR	0.895 ± 2.6%	0.903 ± 2.6%	0.935 ± 2.7%	0.987 ± 2.5%
CNR · Ψ^{-1}	0.481 ± 0.8%	0.489 ± 0.8%	0.498 ± 1.0%	0.502 ± 1.0%

Table 6.3: Relative error for the deblurring problem with different noise levels.

visible noise component. Nonetheless the results are fairly good, as confirmed by the small relative errors (table 6.3) and by the small regularisation parameter error (table 6.2). The latter reflects the small sensitivity of this method to perturbations of similar nature in the data.

The f-slope method has been compared against the L-curve for different test problems [Wu, 2003] and it was claimed that the f-slope method has a good performance for all the problems tested and can often perform better than the L-curve method. In fact, it seems that when little regularisation is needed the f-slope may perform better, but there is no evidence that the same happens when it is necessary to apply more regularisation to the problem. For our deblurring example it is possible to reconstruct an image that is close to the true image when the noise in the data is equal to or higher than 15%. For smaller noise levels it tends to overregularise the solution.

The UPRE and GCV methods are similar in the sense that they both are estimators of the predictive risk. However the GCV method does not require knowledge of the noise variance. Vogel [2002] used a two-dimensional deblurring problem with added white noise to compare these two methods, and observed that the minimisers of the predictive risk, the UPRE and GCV methods are identical. The GCV method has been shown to perform well in different situations. However, a drawback is that sometimes it can be difficult to locate its minimum. The function may be flat near the optimal regularisation parameter or it may display multiple minima, and consequently the method can fail to locate the correct λ [Thomson *et al.*, 1989]. The same applies to the UPRE method. In one example presented by Hansen and O’Leary [1993] of a one-dimensional deblurring test problem, the L-curve and GCV methods yielded reasonable results when the DPC is satisfied. When the DPC is marginally satisfied, the decay rates of

the data space coefficients and the singular values are approximately the same, and the GCV fails to find a regularisation parameter due to the flatness of the curve. For our test problem, the minimum is clearly identifiable for both UPRE and GCV methods, and as expected, the performances of both methods are very similar. The solution for λ obtained with these methods is generally undersmoothed, but the image relative error is still small when the data noise is lower than 10 %, meaning that it produces a reasonable reconstruction.

A study undertaken by Thompson [1991] on the selection of the regularisation parameter for the deblurring problem showed that the DP method tends to oversmooth the solutions and the GCV method generally performs well but eventually may produce underestimated regularisation parameters. The DP method oversmooths even if a good estimate of the noise variance is available [Hansen, 1998]. This observation is confirmed by our study. From Tables 6.2 and 6.3 we can see that $\lambda_{DP} > \lambda_{opt}$, but $\varepsilon_{DP} \approx \varepsilon_{opt}$ which means that it produces acceptable image reconstructions.

The NCP and the DP methods have similar performances, as expected, since they both select the regularisation parameter such that the data norm is equal to the noise variance. However, the former seems to be more sensitive to the different noise realisations, in particular for noise levels above 25%. Hansen *et al* [2006] found that the NCP gives better results than the GCV method, which is confirmed here at higher noise levels.

Hansen [1998] compared the L-curve, the GCV and the QOC methods by applying them to six test problems with different data perturbations but with the same noise level. He found that the L-curve method is most likely to undersmooth the solution, the QOC method generally oversmooths and the GCV method can have both conditions. Another complication that may arise with the QOC method is the presence of a large number of local minima. For our particular case the QOC method overestimates the regularisation parameter, and the image appears too smooth. However, as the noise of the data increases the solutions obtained using this method become more accurate.

The CNR method gives very low regularisation parameters and generated much greater relative errors than for any of the other methods, and the errors associated with λ are as large as those corresponding to the heuristic method. As expected, the weighted contrast gives a λ which is closer to λ_{opt} than the CNR alone. The method still tended to underestimate λ for low noise levels, but for higher noise levels the selection of λ was good. Note that the regularisation parameter uncertainty is particularly large for noise in the data between 15% and 25%.

We applied all the methods to the deblurring problem and we found that some methods produce better reconstructions, some favour noisy images whereas others favour smooth images. In the next section the methods are applied to optical topography data.

6.4 Image reconstruction of phantom data

Data were obtained with the optical topography system developed at UCL with light sources at 670 nm (refer to § 3.2.1.1), and using an object consisting of absorbing targets within a small tank filled with an intralipid solution. Intralipid is a fat emulsion and is suitable for phantoms because it mimics turbid tissues. The intralipid was diluted to provide a transport scattering coefficient of $\mu'_s = 1 \text{ mm}^{-1}$, and a near infrared dye was added to provide an absorption

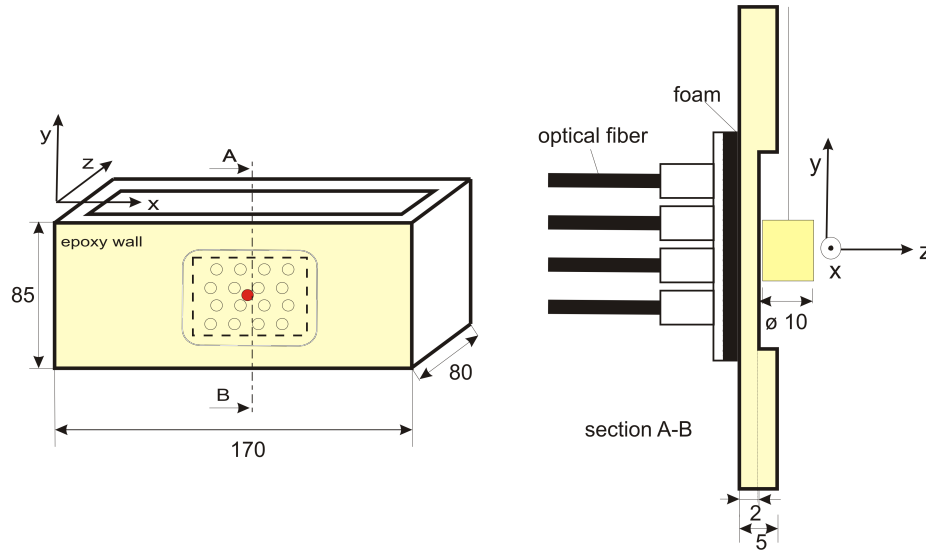


Figure 6.20: Phantom used to acquire optical topography data. The red dot represents the origin $(x, y) = (0, 0)$ relative to the tank and probe. The axis represent the x , y , and z directions. The dashed square represents the 2 mm window in the epoxy wall. Section A-B shows the target at position $z=0$, which is located close to the 2 mm window. Dimensions are in mm.

coefficient $\mu_a=0.01 \text{ mm}^{-1}$. These values are typical of biological tissues at 770 nm. The refractive index was about 1.33. The cylindrical absorbing targets, with 5 mm radius and 10 mm height, have the same μ'_s as the medium but different μ_a . From now on we consider C1 to be the target with $\mu_a=0.0125 \text{ mm}^{-1}$, C2 to be the target with $\mu_a=0.02 \text{ mm}^{-1}$ and C3 to be the target with $\mu_a=0.04 \text{ mm}^{-1}$.

The optical topography array consisted of 8 sources and 8 detectors and data were collected for all possible source-detector distances (see figure 3.9), which results in 64 measurements. The array was placed on one of the walls of the tank, which is made of epoxy resin with the same optical properties as the solution, and is 2 mm thick. Figure 6.20 shows the position $(x, y)=(0 \text{ mm}, 0 \text{ mm})$, which corresponds to the centre of both the array and the tank wall. The centre of the target is approximately 7 mm behind the wall when it is located at $z=0 \text{ mm}$.

The support frame of the probe was made of plastic and lined with a 10 mm thick near-infrared absorbing foam, to prevent light travelling from the sources directly to the detectors. The probe contained 16 holes of approximately 6 mm diameter to which the plastic connectors were attached, which in turn hold the probe optical fibers.

Each target was positioned individually at $x=0 \text{ mm}$, $y=0 \text{ mm}$, and $z=17 \text{ mm}$, and data were collected for around 20 seconds (refer to figure 6.20 for axes directions). Before each acquisition a baseline measurement was collected for calibration, without any target inserted in the solution. The optical topography system at UCL was designed to acquire data at two wavelengths, and the data is stored in a file where each row of data represents one detector, the 1st column corresponds to time, the 2nd column to the event marker and the remaining columns to the detected intensities. A program written in Matlab® separates the data into two files, which subsequently are averaged to filter the noise.

The TOAST forward model was used to calculate the Jacobian matrix using a FEM mesh with 115027 nodes, 74418 elements and dimensions $85 \text{ mm} \times 85 \text{ mm} \times 50 \text{ mm}$, which was

mapped to a pixel basis. Images of the phantom were reconstructed on a regular grid of $24 \times 24 \times 21$ pixels using linear reconstruction, which is described in §5.2.2, based on the algorithm implemented in MATLAB® by Adam Gibson [Gibson *et al.*, 2006].

6.4.1 Results

The results for the methods that satisfy the criteria defined in §6.2.2.5 and also the heuristic method, used for comparison, are described below and summarised in table 6.4. The FWHM and other important quantities that give us important information about the reconstructed images will also be considered here, to help identify the optimal regularisation parameter. The λ values used are the same as before.

Discrete Picard Condition Before selecting the regularisation parameters it was necessary to determine if the DPC is satisfied, in order to know if regularisation can lead to a stable solution. As seen in figure 6.21 the DPC is partially satisfied for all the different targets. The vertical line represents the value i_x , where for $i < i_x$ the data space coefficients $|u_i^T b|$ decay faster than the singular values σ_i and for $i > i_x$ the DPC is not satisfied; here the data space coefficients level off at a level determined by the perturbations in the data b (horizontal line), and because the singular values continue to decay, the coefficients $|u_i^T b| / \sigma_i$ are highly dominated by noise. When the i_x value is not easily identifiable from the solution coefficients $|u_i^T b| / \sigma_i$, we approximate i_x to the beginning of the noise level of the coefficients $|u_i^T b|$. For example for target C2 apparently $|u_i^T b| / \sigma_i$ increases for $i \gtrsim 40$, but the coefficients $|u_i^T b|$ level off at $i_x \approx 28$. For target C1 $i_x \approx 25$ and for target C3 $i_x \approx 26$.

Heuristic Method Using the images reconstructed for different regularisation parameters, and the knowledge that the absorbing object should be located approximately in the centre of the image, we are able to select a range of λ values that can be used to reconstruct images that are in agreement with our expectations. The intervals of regularisation parameters found for the three targets are very similar; for C1 we found $\lambda_1 \in [0.004, 0.025]$, whereas for C2 the interval was slightly smaller with $\lambda_2 \in [0.004, 0.015]$ and even smaller for C3 where $\lambda_3 \in [0.005, 0.01]$. Images for regularisation parameters outside these intervals are either too blurred, or dominated by noise, or the target appears misplaced in the image, and artifacts become dominant in the images. Figure 6.22 shows the reconstructed images (xy plane at $z=17$ mm) of target C1 for three regularisation parameters: a value lower than the minimum acceptable λ found by this method ($\lambda = 0.002$), a value which is included in the set of regularisation parameters ($\lambda = 0.007$), and a value which is above the maximum acceptable value ($\lambda = 0.04$). We can see the presence of high contrast artifacts in the images reconstructed for λ outside the interval of regularisation parameters and that the standard deviation is higher for smaller λ . As expected, the image contrast is overestimated when the image is reconstructed using a low regularisation parameter and is underestimated when a higher regularisation parameter is used, whereas when an adequate regularisation parameter is used the image contrast is close to the true value.

FWHM For this particular study the position of the targets is known approximately, hence plots of the position of the absorbing target as a function of the regularisation parameter were generated, to indicate the accuracy of the range of regularisation parameters selected. However, this approach cannot be used in applications where the true positions of features in the imaged

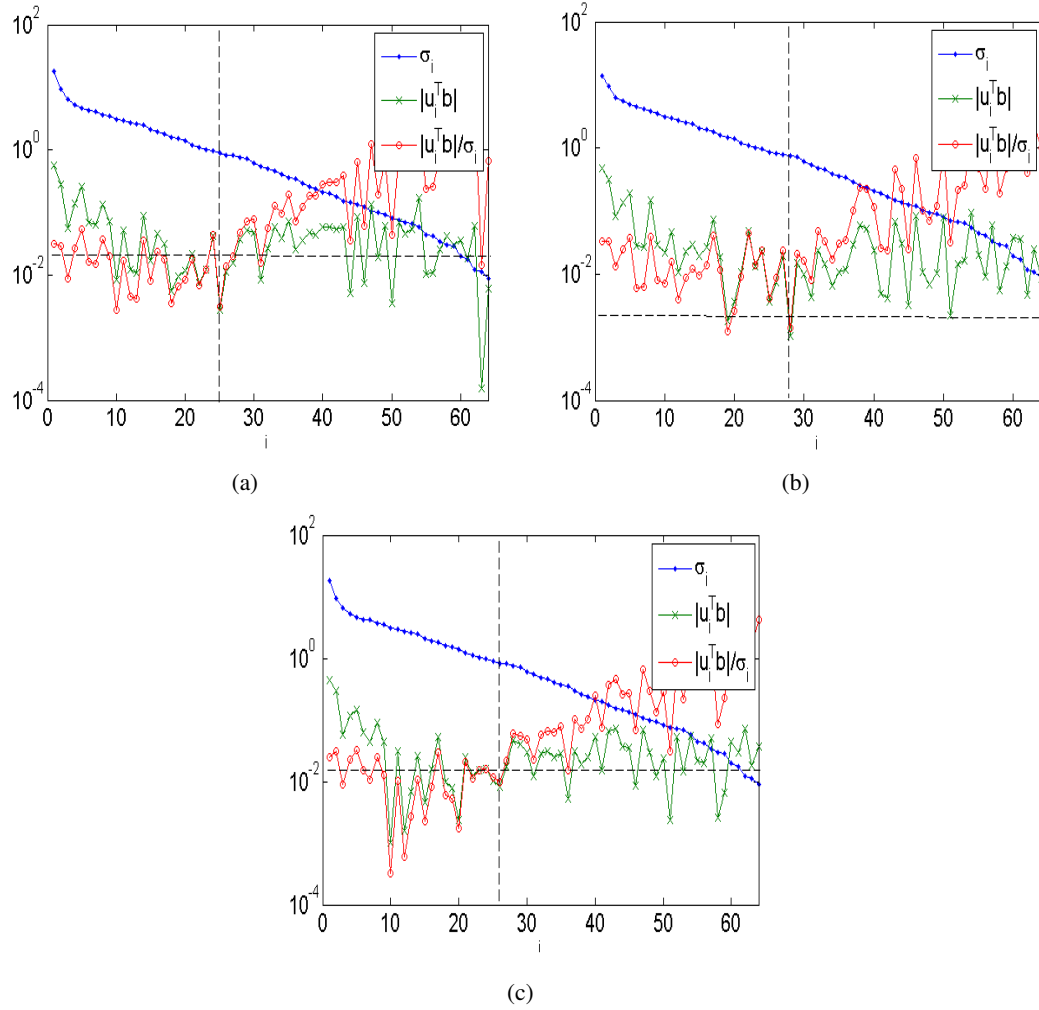


Figure 6.21: DPC for 3 different absorbing targets. There are 64 singular values σ_i . The horizontal dashed line represents the noise level and the vertical line the approximate value i_x , for which $i > i_x$ no longer satisfies the DPC: (a) C1 - $\mu_a = 0.0125\text{mm}^{-1}$ and $i_x \approx 25$, (b) C2 - $\mu_a = 0.02\text{mm}^{-1}$ and $i_x \approx 28$, (c) C3 - $\mu_a = 0.04\text{mm}^{-1}$ and $i_x \approx 26$.

volume are not known. The positions of the target along the x and y axis in pixels are shown in figure 6.23 and 6.24 respectively, the z axis is not shown since we have only considered targets at a depth of $z = 17$ mm. The arrows in the figures indicate the range of values for which the object appears in the image in the correct position, which is located ideally at pixel number 12 (refer to figure 6.23 and 6.24 again for λ values). For other values of λ the positions change significantly depending on the regularisation parameter used, which is more obvious along the x axis.

The change in the maximum absorption coefficient $\mu_{a_{max}}$ as a function of the regularisation parameter is plotted in figure 6.25. It can be seen that its value is high for small regularisation parameters and it decays rapidly to an almost constant value for $\lambda > 0.005$.

The FVHM represents the apparent volume of the object. The apparent volume of the target varies with the amount of regularisation applied, increasing relatively slowly until it reaches a point of maximum volume and then dropping drastically to a constant value when too much regularisation is applied and the image is too smooth (figure 6.26). The rectangle

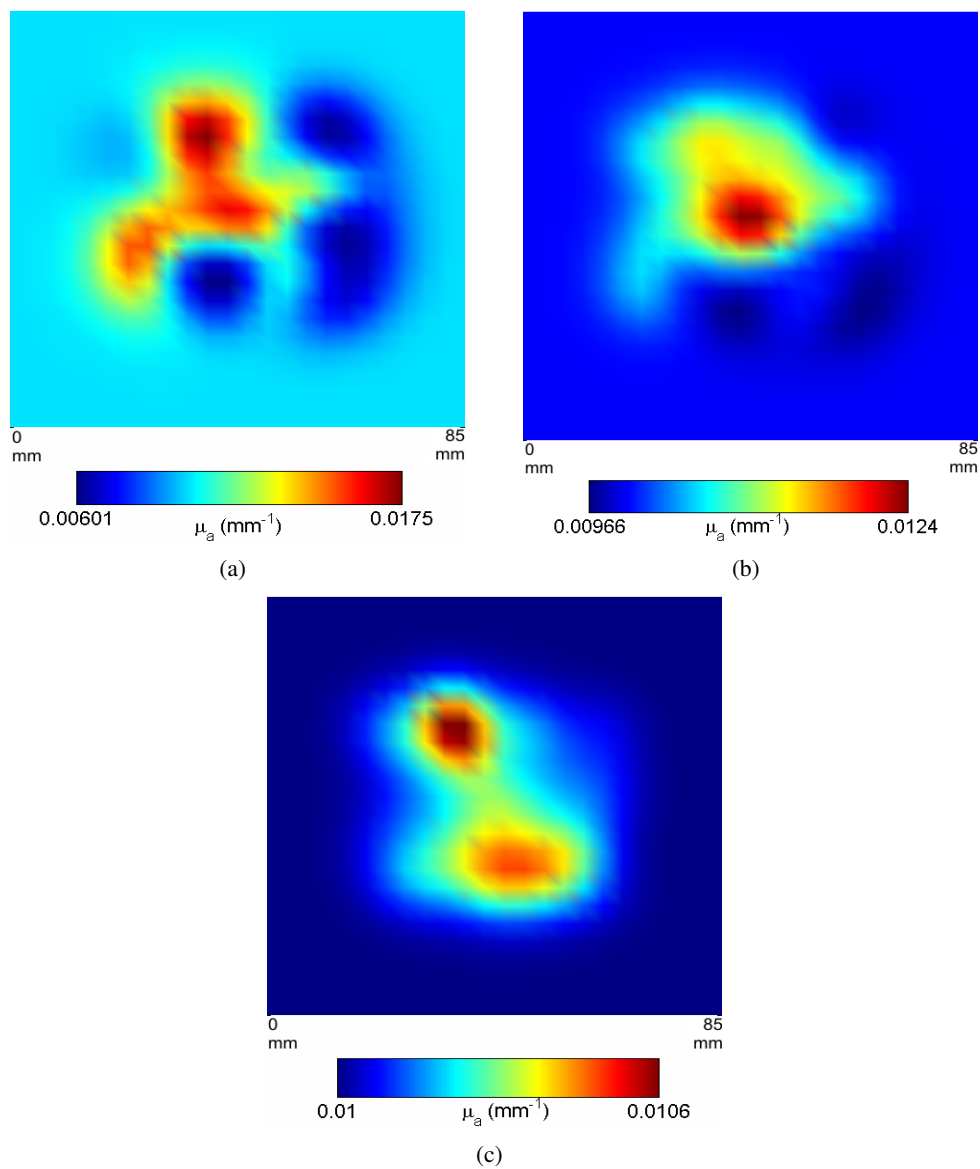


Figure 6.22: Images reconstructed (xy plane at $z=17$ mm) of target C1 for (a) $\lambda = 0.002$ (b) $\lambda = 0.007$ and (c) $\lambda = 0.04$. The true position of the target is 17 mm behind the centre of the imaging surface. The colourmap represents the changes in the absorption coefficient μ_a .

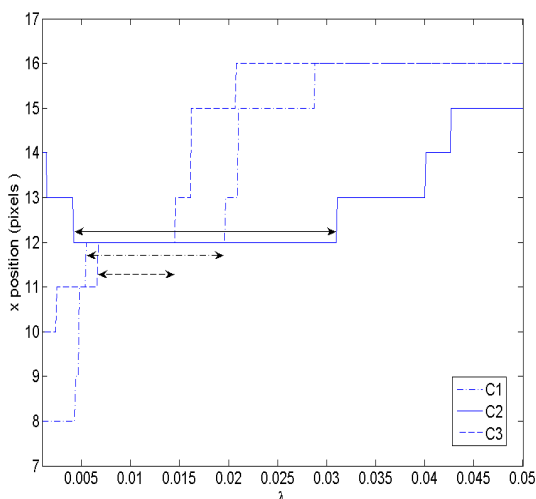


Figure 6.23: Position of the target on the x axis of the reconstructed image for different λ values. The target is in the correct position for $\lambda_1 \in [0.0055, 0.0196]$ ($-\cdot-$), $\lambda_2 \in [0.0042, 0.031]$ ($-$) and $\lambda_3 \in [0.0067, 0.0145]$ ($- -$).

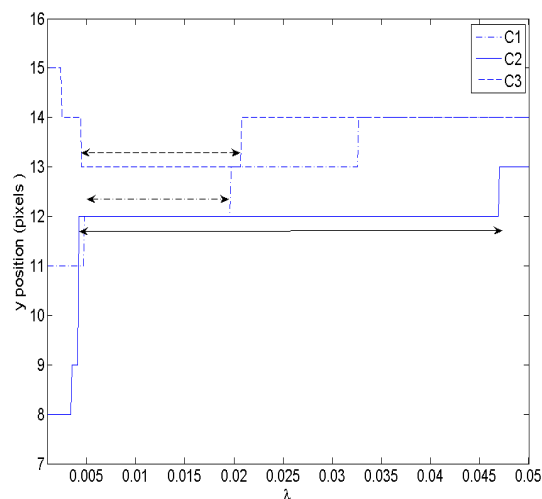


Figure 6.24: Position of the target on the y axis of the reconstructed image for different λ values. The target is at pixel 12 for $\lambda_1 \in [0.0047, 0.0196]$ ($-\cdot-$), $\lambda_2 \in [0.0044, 0.0469]$ ($-$) and $\lambda_3 \in [0.0045, 0.0207]$ ($- -$).

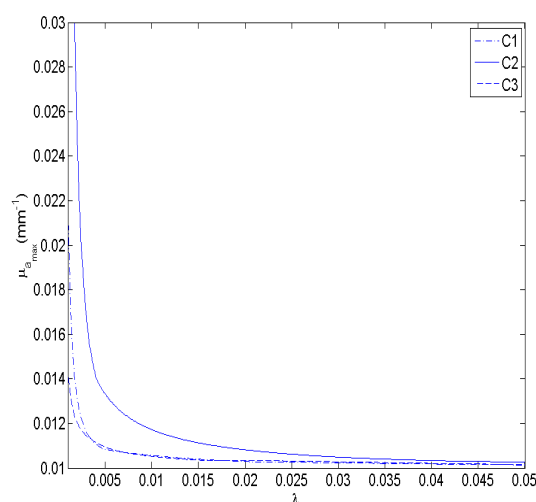


Figure 6.25: Maximum absorption coefficient for each image reconstructed.

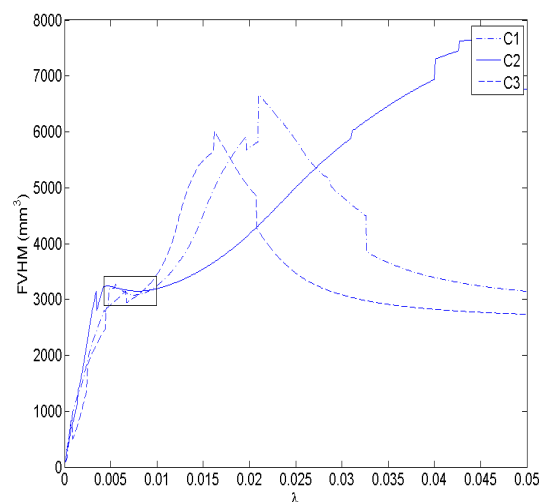
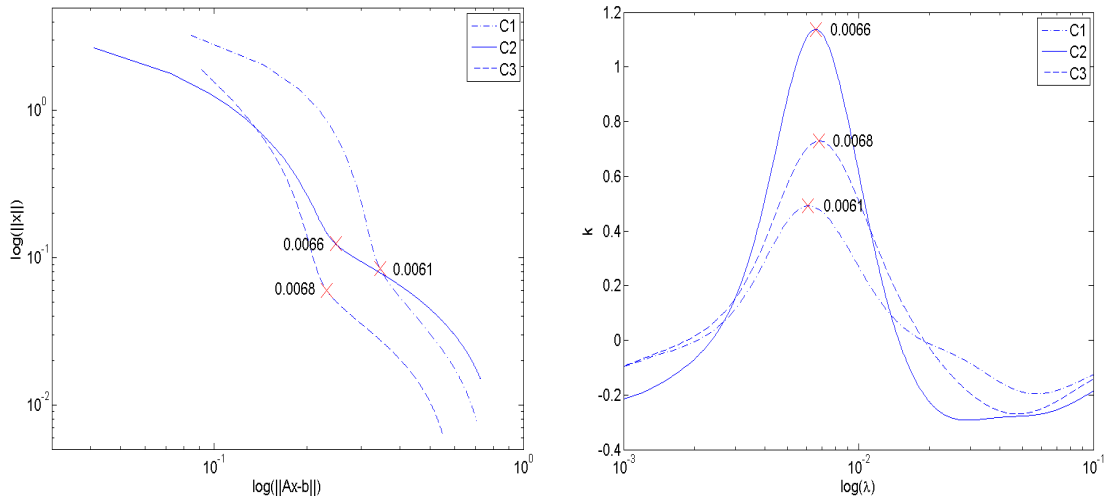


Figure 6.26: FWHM of the target in the images. The rectangle shows $\lambda \in [0.0045, 0.01]$ where the size of the target is approximately constant for C1 and shows a linear dependence for C2 and C3.

designates a region of the curves where the volumes are fairly constant, meaning that any regularisation parameter selected from this interval, $\lambda_{1,2,3} \in [0.0045, 0.01]$, gives a reasonable image resolution.

L-curve The L-curve for the optical data does not exhibit a pronounced corner, but it is still possible to calculate the point of maximum curvature. Figure 6.27 shows the L-curve method applied to the data for the three targets in this study. The regularisation parameter found with this method is consistent for all the experiments undertaken here. For the first target the point of maximum curvature was reached for $\lambda_1 = 0.0061$, for C2 λ_2 takes the value of 0.0066, while for C3 the corner of the L-curve appears at $\lambda_3 = 0.0068$.



(a) L-curves for the phantom containing C1, C2 and C3. (b) Curvatures of L-curves for the different phantoms.

Figure 6.27: L-curve method for the optical topography problem.

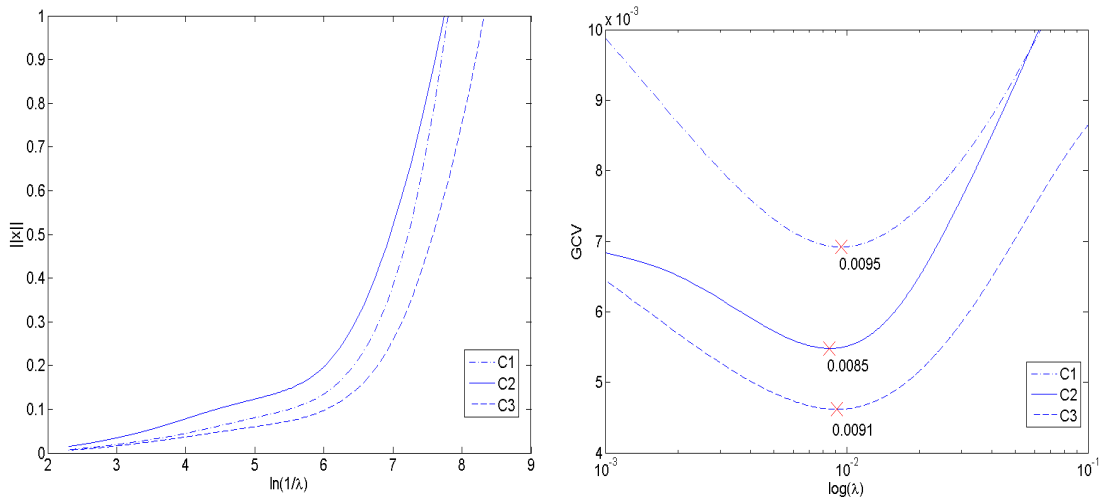


Figure 6.28: The f-slope method fails to find a regularisation parameter λ for experimental data.

Figure 6.29: The GCV method and its minima for the phantoms in study.

f-slope The solution norm exhibits a fast growth with $\ln(1/\lambda)$, which does not display a flat region characteristic of the f-slope method, and therefore it was not possible to identify the minimum slope for each curve (figure 6.28).

GCV For all three targets we were able to localise an absolute minimum for the convex GCV functions and their values are very similar (figure 6.29). This method obtained for the first target C1 a minimum value at $\lambda_1 = 0.0095$, a regularisation parameter of $\lambda_2 = 0.0085$ for C2 and finally C3 yields a minimum at $\lambda_3 = 0.0091$.

CNR To calculate C it is necessary to know $\mu_{a_{max}}$ and $\mu_{a_{bkg}}$. The former is the mean of 9 pixels, where one of them is the pixel of maximum intensity in the image and the other 8 are the pixels which surround it in the xy plane. The latter is obtained from an average of 24 pixels from the boundary of the image furthest from the peak in the image. The curves for the CNR are shown in figure 6.30(a), and their local maxima are identifiable and marked by the crosses. The

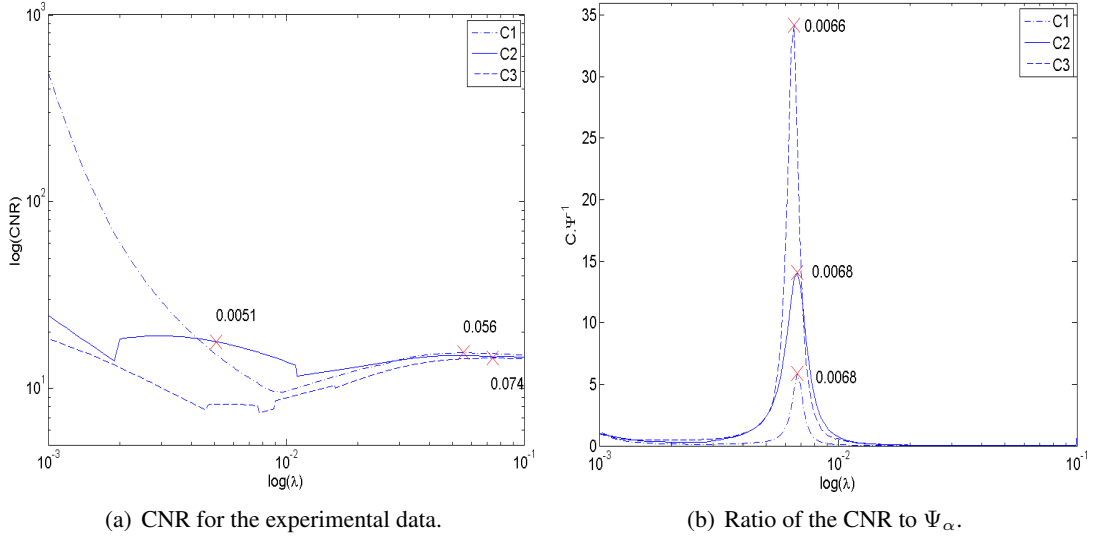


Figure 6.30: Plots of the CNR method as a function of the regularisation parameter

curve corresponding to target C1 is very similar to C3, showing a global minimum followed by a local maximum, which occurs at $\lambda_1 = 0.056$ and $\lambda_3 = 0.074$. The minimum and maximum occur for lower regularisation parameters in the case of the C2 target, and this method results in $\lambda_2 = 0.0051$. When dividing the CNR by Ψ_α the results become more consistent. The α values were $\alpha_1 = 0.27$, $\alpha_2 = 0.48$ and $\alpha_3 = 0.32$ for the three targets. Each curve shows a maximum as shown in figure 6.30(b), with the highest peak value at $\lambda_3 = 0.0066$, followed by $\lambda_2 = 0.0068$ and the smaller peak occurs for C1 at $\lambda_1 = 0.0068$.

QOC As shown in figure 6.31 the QOC criterion for the experimental data for C1 and C3 exhibits a steadily decreasing function without any local minima, and consequently this method fails. For target C2 a local minimum is visible at $\lambda_2 = 0.0097$.

DP The DP function has a zero for target C1 at $\lambda_1 = 0.0071$, for C2 at $\lambda_2 = 0.0061$ and for C3 at $\lambda_3 = 0.0051$ (figure 6.32). The noise variance was estimated from the noise vector $u_i^T \mathbf{e}$ which levels off at $i_x = \sigma$, and considering the coefficients $u_i^T \mathbf{b} = u_i^T \mathbf{b}_{exact} + u_i^T \mathbf{e}$, we find that $E(|u_i^T \mathbf{e}|^2) = \sigma^2$. Overall, for $i < i_x$ the coefficients $|u_i^T \mathbf{b}_{exact}|$ dominate and decay levelling off at the noise level at $i \geq i_x$, which corresponds to $|u_i^T \mathbf{e}|$.

UPRE Figure 6.33 shows the UPRE functions for each target. The minima were localised at $\lambda_1 = 0.0081$, $\lambda_2 = 0.0071$ and $\lambda_3 = 0.0081$.

NCP It is necessary to pre-process data so that all measurements have zero mean and unit variance. In optical topography we use the difference imaging approach, where the baseline is subtracted from the data, hence all the measurements should have zero mean, but different variance. Whitening could be accomplished by dividing each measurement by the respective standard deviation. The regularisation parameters obtained with this method were quite low: $\lambda_1 = 0.0034$, $\lambda_2 = 0.0032$ and $\lambda_3 = 0.0034$. Figure 6.34 shows the optimal normalised cumulative periodograms for each target. The residual is affected by high frequency components for small λ values, whereas for large λ contributions from small frequency oscillations are more significant.

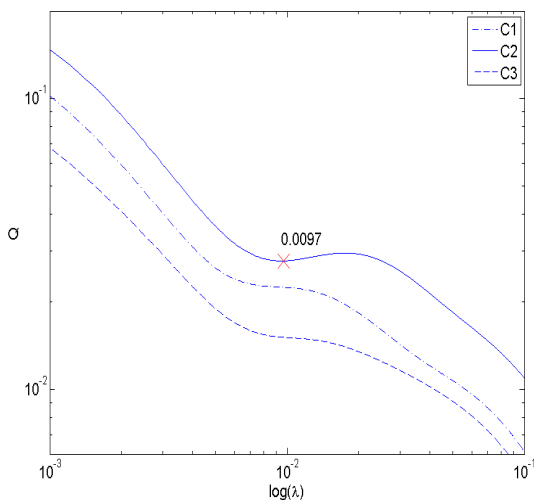


Figure 6.31: The QOC fails to exhibit a minimum for C1 and C3.

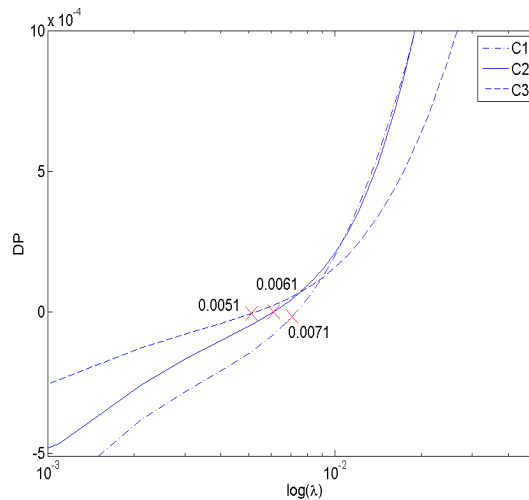


Figure 6.32: DP method and the respective zeros for each function.

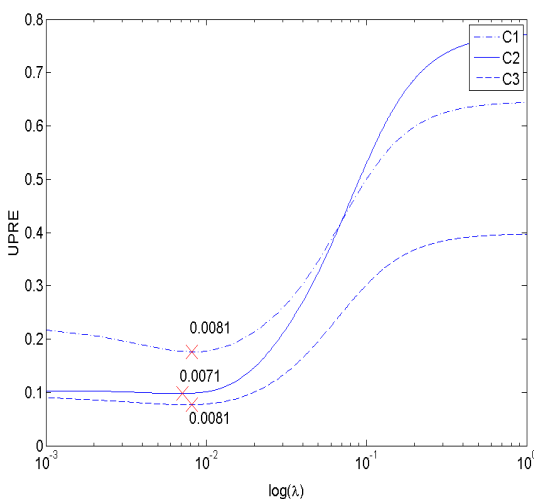


Figure 6.33: UPRE functions and its minima.

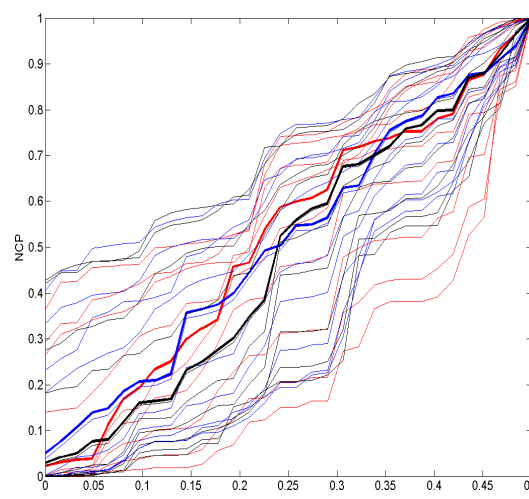


Figure 6.34: Normalised cumulative periodograms for C1 represented by the blue lines, C2 by the red and C3 the black lines. The ticker lines correspond to the optimal NCPs.

If we set the Kolmogorov-Smirnoff limits [Brockwell & Davis, 1998] to a significance level of 5%, which is equivalent to a 95% confidence level, and choose the largest regularisation parameter inside these limits, we obtain $\lambda_2 = 0.0064$, $\lambda_2 = 0.0056$ and $\lambda_3 = 0.006$, which are more reasonable values.

6.4.2 Discussion

In *in vivo* applications the solution of the optical imaging problem is unknown so the results of the Heuristic method must be considered as the optimal regularisation parameter, and the selection methods compared against this λ_{opt} value or interval. Tikhonov regularisation suppresses noise and only low frequency components of the solution are reconstructed, which results in images that are known to be smooth with a limited resolution. Since we

methods	λ_1	λ_2	λ_3
Heuristic	[0.004, 0.025]	[0.004, 0.015]	[0.004, 0.01]
FVHM	[0.0045, 0.01]	[0.0045, 0.01]	[0.0045, 0.01]
xpos	[0.0047, 0.0196]	[0.0044, 0.0469]	[0.0045, 0.0207]
ypos	[0.0055, 0.0196]	[0.0042, 0.031]	[0.0067, 0.0145]
L-curve	0.0061	0.0066	0.0068
GCV	0.0095	0.0085	0.0091
UPRE	0.0081	0.0071	0.0081
DP	0.0071	0.0061	0.0051
NCP	0.0034	0.0032	0.0034
f-slope	–	–	–
QOC	–	–	–
CNR	0.056	0.0051	0.074
$\text{CNR} \cdot \Psi^{-1}$	0.0068	0.0068	0.0066

Table 6.4: Regularisation parameters for the experimental data using different selection methods

are constrained by this factor it can be difficult to identify an image that is a reasonable representation of the true object. The images reconstructed using the Heuristic method all have in common several features such as: correct positioning of the target, relatively high contrast, low presence of noise and moderate smoothing. For easier comparison with the other methods, the apparent position and FVHM of the target in the images were calculated for the different regularisation parameters. The range of regularisation parameters found using the FVHM analysis is approximately in agreement with the heuristic selections and for this range of λ the target is correctly positioned in the image (refer to table 6.4).

Our optical imaging problem is an under-determined problem, where we have 64 measurements corresponding to 64 source-detector pairs at different separations and we aim to reconstruct a three-dimensional image consisting of 12096 pixels. The decay rate of the singular values measures the degree of ill-posedness [Hofmann, 2005] and it indicates that this is a very ill-posed problem. The accuracy of the regularised solution is highly affected by these factors, which means that the selection methods are also influenced and as a consequence they may fail to find a suitable regularisation parameter. The L-curve method reflects this, by the absence of a sharp corner. In spite of these difficulties, the images reconstructed with the regularisation parameters obtained with this method are consistent with the Heuristic method. The L-curve method has been used previously in optical imaging, for simulated and real data, and its results are considered to generate acceptable images [Gaudette *et al.*, 2000; Zhou *et al.*, 2006; Zhao *et al.*, 2006b]. To our knowledge the f-slope method has not been tested previously on real data, only for testing models where it has a good performance [Wu, 2003]. This method does not exhibit a flat region, characteristic of this method, when applied to the optical topography problem. This seems to be due to its ill-posedness which results in fast changes of the solution norm as mentioned before. Another method that failed to converge was the QOC. No minimum was found for C1 and C3. For C2 the regularisation is in the interval of acceptable λ values defined by the Heuristic method.

Hansen [1992] tested the L-curve against the GCV method for a test problem perturbed by errors that are highly correlated, for example sampling errors or approximation errors in matrix A . He found that the GCV method fails to compute a useful solution; a very small λ is selected because this method treats errors as being part of the signal. The L-curve seems

to find a solution that is a good approximation to the exact solution, as long as the correlated errors do not satisfy the DPC. In the presence of uncorrelated errors both methods choose good regularisation parameters, although the GCV method gives a slightly overregularised λ . The GCV method for the optical topography phantom data chooses regularisation parameters which produce good image reconstructions. The minimum value of the GCV function occurs for λ values higher than those found with the L-curve method. This result could be explained by a low presence of correlated errors.

The L-curve, GCV, DP and UPRE methods were compared for simulated, phantom and neonatal EIT data [Abascal *et al.*, 2008]. All the methods were successful for simulated and clinical data. Overall, the preferred methods were the GCV and L-curve. For optical topography data both DP and UPRE methods converged and the predicted regularised parameters λ are in agreement with the values found heuristically. Note also the similarity between predictors from the GCV and UPRE methods.

The NCP method returned a very low regularisation parameter, which is not included in the values found heuristically. However, if we set significance levels of 5% the value obtained is more reasonable.

An artifact with high contrast leads to an erroneous selection of $\lambda_{1,3}$ from the CNR curve. A good reconstruction is achieved for target C2 because although the artifact is still present its contrast is smaller compared to that produced by the target.

If Ψ is applied as a weighting factor to CNR, the regularised solutions are similar to the ones generated from the L-curve method. This is probably due to the similarity between Ψ and the L-curve method [Regińska, 1996]. The regularisation parameters selected with these methods are contained in the set of λ_{opt} from the Heuristic method.

The quality of the reconstructed images is highly dependent on the quality of the acquired baseline; an inadequate baseline can produce meaningless results. Intralipid solutions are not completely stable over time, and even if the solution is regularly stirred its properties are not exactly the same during the whole experiment, which can explain the difference between the results for the targets and the similarity between results for targets C1 and C3.

6.5 Conclusions

Optimising image reconstruction for diffuse optical imaging requires special attention to the amount of regularisation applied, which controls the smoothness of the regularised solution and balances the influence of the noise present in the image against its accuracy. We have examined a number of methods for selecting the regularisation parameter, as described in §6.2. Diffuse optical imaging is a challenging ill-posed and under-determined problem, and for *in vivo* applications the true distribution of optical properties is unknown, making it difficult to predict the optimal solution. We therefore initially tested the methods by applying them to a simpler ill-posed problem, the deblurring problem, where we can control the amount of blur and noise applied to the data, where the exact solution is known (and by using this information it is easy to know the optimal regularisation parameter), and where the size of the data array is the same as the solution. We believe that if a method does not produce a good regularised solution for the deblurring problem, we cannot rely on it for optical topography. However, if a good

solution is found with a certain method for the deblurring problem the same is not necessarily true for the more demanding optical problem.

From this study we hoped to identify a single method that selects a regularisation parameter, which produces a meaningful regularised solution that works for diverse experimental conditions and does not rely on human judgement. It is also desirable that the chosen method is fast and efficient, particularly if we have to manipulate a large amount of data or perform real-time imaging. With this in mind, in §6.2.2.5 we list three criteria that a selection method must meet. On the basis of these criteria, we reject the heuristic method, the fixed noise figure method and methods related to optimising the full width at half maximum of the image. All the remaining methods performed acceptably well for the deblurring problem. However, the f-slope and QOC failed to converge for the more demanding optical topography problem.

The CNR method gave poor results, but it was much more successful when Ψ was minimised simultaneously. The use of Ψ in this way, as proposed by Regińska [1996] and developed further here, could be applied to other methods and may be worth further examination.

The DP and UPRE methods use statistical information in order to predict the regularisation parameter. If the noise variance or an accurate estimative is available, then these methods are able to provide good predictions of the regularisation parameter. The NCP method has the advantage of selecting λ automatically, hence it does not directly require an estimate of the noise variance. Nevertheless, this method is sensitive to error fluctuations and only gives reasonable results for the optical topography problem under certain assumptions.

Another method that obtained a good regularisation parameter was the GCV method. This method can be useful when we search for smooth solutions since it tends to overregularise the solution. The UPRE and GCV methods both estimate the predictive risk, although the GCV method does not require any prior knowledge of noise variance. The GCV method is also known to have a high theoretical convergence [Vogel, 2002]. DP, UPRE, NCP and GCV only consider the data norm.

The L-curve consistently demonstrated the lowest error for the deblurring problem (see table 6.2 and 6.3). It is easy to implement and simultaneously minimises both residual norm and solution norm. This method requires the computation of both the solution norm and residual norm for a large number of regularisation parameters, which is a disadvantage common to all the tested methods.

The purpose of this study is to select a single method that can be used for optical imaging problems, and the L-curve was found to be the optimal selection method, since it performs well for both deblurring and optical topography problems.

So far we have studied ideal or almost ideal cases, whereas *in vivo* studies suffer further sources of error. These include motion artifacts, changes in the contact between the optodes and the skin which can result in intensity fluctuations in the collected data, and detection of light which has not passed through the investigated medium. All these effects produce correlated errors, which are not necessarily apparent in the raw data. The L-curve has been shown to perform well in the presence of these type of errors [Hansen & O'Leary, 1993; Hansen, 1992].

As mentioned previously, Hanke [1996] showed that one of the disadvantages of the L-curve method is related to the reconstruction of very smooth solutions (when the coefficients $|u_i^T b|$ decay fast to zero) where this method fails to find a good regularisation parameter. Under these circumstances, GCV should be investigated as an alternative to the L-curve method.

In conclusion, the quality of the images does not depend uniquely on the regularisation parameter, but for the reconstruction methods available at present to solve ill-posed inverse problems the choice of the regularisation parameter is important, since it controls the degree of accuracy and smoothing of the solution. The L-curve seems to find reasonable regularisation parameters for the optical topography imaging problem, but it is still necessary to test the validity of this method for *in vivo* situations, which is done in §9.

Quantitative assessment of the depth sensitivity using tissue-like phantoms

Images produced from optical topography measurements have limited depth information, and are restricted to the superficial layers of the tissue underneath the array of sources and detectors. The use of arrays with several source-detector pairs at different separations increases the ability to discriminate between signals at different depths, and overlapping measurements improve the lateral resolution and accuracy [Zeff *et al.*, 2007; Dehghani *et al.*, 2009; Heiskala *et al.*, 2009]. Since light penetration is limited by the highly scattering nature of tissue, the sensitivity is higher at the surface than for deeper tissues. Larger source-detector separations are required to sample at greater depths, and large separation implies more attenuation and less detected signal. The volume of tissue sampled also increases with larger separation, and consequently spatial resolution and contrast also decrease with depth. Image reconstruction using zero order Tikhonov regularisation assigns high sensitivity measurements to high sensitivity regions in the reconstructed image. Therefore, perturbations in optical properties tend to be reconstructed in shallow areas and deeper tissues are overshadowed by these. Also, there is often no PMDF (refer to §5.2.1 for definition) overlap immediately below the optodes, hence the reconstruction tends to place perturbations below the optodes. It is therefore necessary to develop means of modifying the algorithm or the images in order to locate the perturbations in their correct positions.

One possible correction to compensate for this effect is to use a spatially variant regularisation (SVR) in the Tikhonov regularisation process instead of a constant value [Arridge & Schweiger, 1993; Pogue *et al.*, 1999]. In this chapter, different spatially variant regularisation parameter types are used to reconstruct images from simulated and experimental data, in order to analyse their influence in the image quality. Some of the main spatially variant regularisation parameter types found in the literature are described below.

Pogue *et al* [1999] used a spatially variant regularisation parameter on tomography data. In their theoretical and experimental study the sources and detectors were placed in a

ring surrounding the object. The regularisation parameter varied exponentially in the radial direction, applying less regularisation to the region of lower sensitivity, nearest the centre of the object, and more regularisation towards the surface, to reduce noise sensitivity near the optodes. Their results show a relatively even contrast and resolution across the image, and reduction in the effects of noise near the optodes. The regularisation parameter can be expressed as

$$\lambda(r_j) = \lambda_e \exp(r_j/R) + \lambda_c \quad (7.1)$$

where r_j is the radial position of pixel j , R is the radius of the image, and λ_e and λ_c are free variables determined empirically .

Culver *et al* [2003] performed phantom and *in vivo* measurements and used a variable regularisation parameter to reconstruct the images. They adapted the variable regularisation parameter to three dimensions, allowing the regularisation parameter to vary for a whole volume. Images of multiple phantom targets at different spatial locations were successfully reconstructed. Kepshire *et al* [2007] investigated imaging of absorbing and fluorescent targets at different depths using both simulations and experiments. A SVR similar to equation (7.1) was applied to the image reconstruction, but they replaced R by an empirically determined decay coefficient α that they found to be $\alpha=25$ mm. The location of the targets in the image can be accurately measured within 10 mm below the surface.

Recently Zhao *et al* [2007] have introduced the layer-based sigmoid adjustment (LSA). They correct sensitivity variations by adjusting directly the forward matrix according to a sigmoidal function. Simulations showed that without any depth adjustment the absorbing region in the image appears to shift towards the surface. When SVR is used the object appears to be deeper than it actually is. The LSA method achieves better results both for superficial and deep objects. They showed that depth contrast is improved and position errors of less than 3 mm could be obtained for targets located at depths from the surface between 1 cm and 3 cm. A similar study, using a similar adjustment of the forward matrix, where they multiply the elements in the sensitivity matrix by the maximum singular value for each depth layer in inverse order, yielded identical results [Niu *et al.*, 2008].

Endoh *et al* [2008] assigned a different regularisation parameter to each voxel by setting the regularisation parameter proportional to the square-root of the diagonal elements of $J^T J$, hence compensating for both depth and lateral sensitivity decrease. They successfully reconstructed simultaneously two simulated targets located at different depths. This type of regularisation had been previously used by Culver *et al* [2005] to generate images of haemodynamic activity in rat brain, but without the intention of obtaining depth discrimination. This type of regularisation is well known in EIT image reconstruction as NOSER (Newton's One Step Reconstructor) [Cheney *et al.*, 1990; Adler *et al.*, 2007].

Another approach is to use a covariance matrix as the penalty term [Guyen *et al.*, 2005; Boas & Dale, 2005]. Each entry of this diagonal matrix is the variance σ_y^2 of the measurements, allowing adaptive regularisation according to the accuracy at each position derived from the measurements. Here the regularisation parameter λ appears as a scaling factor of the covariance matrix. One can use the covariance matrix of the measurements C and the covariance matrix of

the unknowns R to derive an inverse operator W

$$W = RJ^T(JRJ^T + \lambda C)^{-1} \quad (7.2)$$

The perturbations in the optical properties x can be extracted from the measurement data b using $x = Wb$ following an iterative procedure [Gibson *et al.*, 2005; Yalavarthy *et al.*, 2007]. The matrix R can also be given by the NOSER method (refer to §7.2) [Adler *et al.*, 2007]. Adler *et al.* [2007] has proposed a method that incorporates spatial and temporal correlations in the image in the regularisation prior, where R in the previous equation is the regularisation matrix that includes both spatial and temporal priors. This method could be applied to optical topography imaging of brain functional activity, where the aim is to produce spatial and temporal images of changes in the concentrations of oxy- and deoxy-haemoglobin.

It has been shown that a regularisation matrix based on space-uniformisation of the variance of reconstructed EIT images yields better images of regions distant from the surface than the zero order Tikhonov regularisation [Cohen-Bacrie *et al.*, 1997]. This method eliminates the differences in the image variance, which is a direct consequence of the differences in the variance of the measurements.

A very simple spatially variant regularisation parameter can be defined by dividing the image into two separate regions, containing voxels with two different regularisation parameters: the voxels of the region of interest are assigned a smaller regularisation parameter and the background voxels are assigned a higher regularisation parameter [Srinivasan *et al.*, 2004; Li *et al.*, 2003].

Alternative approaches to assigning different regularisation parameters to different regions have been tested. Endoh *et al.* [2005] tried to compensate for the lack of sensitivity in deeper regions by increasing the corresponding voxel size in the image.

In theory, time-domain (TD) optical systems offer better depth resolution than CW systems [Selb *et al.*, 2005]. TD systems send a pulse of light into the tissue and record the time the photons take to travel through the tissue back to the detectors on the surface. The TPSF provides a distinction between photons that have travelled shorter distances, the so-called early photons, which means they have only reached superficial layers, and the late photons that have probed deeper tissues. This approach provides higher sensitivity to deeper tissues of the brain (cortex) and can reject early photons, with information from outer layers of the head where activation is not expected to occur (scalp and skull). Selb *et al.* [2006] tested the depth resolution of a TD system using a phantom and *in vivo* measurements. They were able to detect the superficial and deep inclusions in the phantom using the early and late photons respectively, and observed a higher contrast with small variance for the superficial inclusion. They were also able to image activation of the motor cortex of an adult brain when performing a finger-tapping task, and verified that the contrast was higher when information from early photons was not used. Activation of the brain was measured with a TD system during a motor stimulation and a Valsalva manoeuvre and it was possible to distinguish between intracranial absorption and scalp activation respectively [Steinbrink *et al.*, 2001; Montcel *et al.*, 2005].

Frequency-domain (FD) systems have also been used to provide depth information in

the brain [Kohl-Bareis *et al.*, 2002]. FD systems measure amplitude and phase at a single frequency. Amplitude is more sensitive to absorption changes near the surface whereas phase is more sensitive to perturbations in deeper layers of the brain.

The use of a variable regularisation parameter has proven to increase depth resolution. Different types of regularisation were incorporated separately in our linear reconstruction algorithm. The performance of each method was evaluated and, finally, the one that generated images with greater accuracy was incorporated in the linear image reconstruction algorithm.

To determine which spatially variant regularisation parameter is more appropriate for our data, we tested the different types of regularisation on both simulated and phantom data. Simulations, if correctly implemented, are good approximations to reality, and are inexpensive, relatively easy and fast to perform. However, they usually reproduce an ideal and simplified system performance where, for example, all sources have identical noise levels. To fully characterise a specific imaging system, appropriate phantom measurements are essential. This study is of high importance since it will provide an indicator of how trustworthy the results are when images are reconstructed from *in vivo* measurements.

7.1 Methods

7.1.1 Simulations

A TOAST forward model was used to calculate both simulated data and a Jacobian matrix, using a FEM mesh with 115027 nodes, 74418 elements and dimensions $85 \text{ mm} \times 85 \text{ mm} \times 50 \text{ mm}$. Initially the medium is homogeneous with optical properties $\mu'_s = 1 \text{ mm}^{-1}$, $\mu_a = 0.01 \text{ mm}^{-1}$ and $\eta = 1.4$, and data were generated to be used as reference data. Regions, with dimensions $10 \text{ mm} \times 10 \text{ mm} \times 10 \text{ mm}$ and higher absorption coefficient ($\mu_a = 0.0175 \text{ mm}^{-1}$), were then inserted in turn in the centre of the mesh at depths from 10 mm to 30 mm in steps of 5 mm. Five data sets were produced, to which 0.01% Gaussian noise was added, and images were reconstructed on a regular grid of $24 \times 24 \times 15$ pixels using a linear reconstruction algorithm described in §5.2.2.

7.1.2 Optical topography phantom experiments

Data were obtained using the UCL optical topography system and probe described in §3, operating at a wavelength of 670 nm.

7.1.2.1 Liquid phantom

A tank filled with an intralipid solution was used to perform measurements similar to the experiment described in § 6.4. As before, the intralipid solution has optical properties $\mu_a = 0.01 \text{ mm}^{-1}$, $\mu'_s = 1 \text{ mm}^{-1}$ and refractive index $\eta = 1.33$. The probe was placed on the centre of the epoxy resin wall.

We inserted individually four different cylindrical absorbing targets in the liquid at different known positions along the x , y and z axis. In order to place the target in the desired position we used three translation stages assembled on an optical table. Figure 7.1 shows the experimental setup with one of the targets inserted in the liquid phantom. As described previously, we considered the x and y directions to be parallel to the array (x is the horizontal axis and y is the vertical axis), and z axis to be perpendicular.

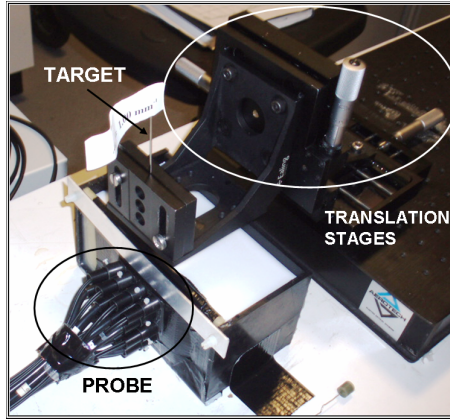


Figure 7.1: The experimental setup.

target	μ_a (mm^{-1})
T1	0.0175
T2	0.04
T3	0.02
T4	0.005

Table 7.1: Targets used in the experiment. The transport scattering coefficient is $\mu'_s=1 \text{ mm}^{-1}$ for all targets.

The absorbing targets used for this experiment had a diameter and height of 10 mm and the same μ'_s as the medium but different μ_a . All targets and their respective absorption coefficient μ_a are listed in table 7.1. Since one of the main applications of optical topography is brain imaging, targets should ideally represent a small absorption change typical of the evoked response in the brain. Another reason to use targets with properties close to the surrounding medium is that the linear reconstruction approach is only valid for small changes. Nevertheless, a highly absorbing target is also used for comparison.

Experiment 1 - variable depth. Target T1 was used and positioned at $x = 0$ mm and $y = 0$ mm, which is the centre of the array and where sensitivity to changes in the medium should be highest. The target was moved in depth from $z = 12$ mm to $z = 32$ mm in steps of 5 mm (figure 7.2(a)).

Experiment 2 - fixed depth. For this experiment target T1 was fixed at depth $z = 12$ mm and moved in the xy plane, as illustrated in figure 7.2(b). The target was moved from $x = -10$ mm to $x = 10$ mm in steps of 10 mm and the same for the y direction.

Experiment 3 - two targets. The purpose of this experiment is to determine the ability of the system to separate two targets 2 cm apart. In this study we have used two pairs of targets, the first pair represented increases in the absorption and the second pair corresponds to an increase and a decrease in absorption. Both targets were centrally placed at depth $z = 12$ mm (see figure 7.2(c)). We chose T3 for both measurements which was paired with target T2 for the first measurement and with T4 for the second measurement.

Data processing. For all the experiments data were collected at each target position for around 20 seconds and subsequently averaged to reduce the noise. Previous to each acquisition a baseline measurement was collected for calibration, without any target inserted in the solution. Differences between reference measurements and target measurements were used in the linear image reconstruction, to generate three-dimensional images representing absorption changes occurring within the phantom. A TOAST forward model was used to calculate the Jacobian matrix, as described previously.

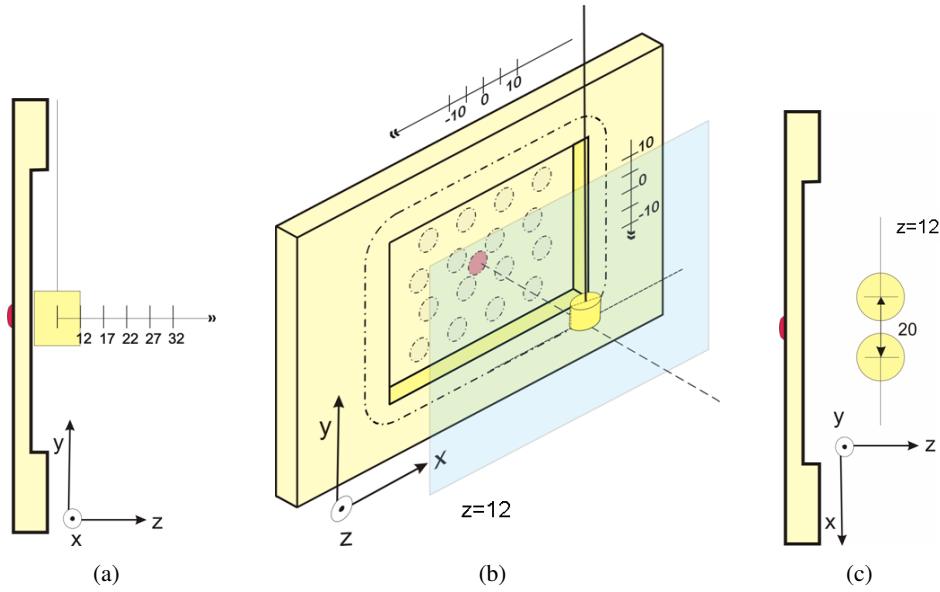


Figure 7.2: Phantom experiments. The centre $(x, y)=(0, 0)$ is marked in red. Dimensions are in mm. (a) Experiment 1 - Target at $(0, 0, z)$ moved from $z = 12$ mm to $z = 32$ mm in steps of 5 mm. (b) Experiment 2 - Target at $(x, y, 0)$ moved from $x = -10$ mm to $x = 10$ mm and in steps of 10 mm and for each x moved from $y = -10$ mm to $y = 10$ mm. (c) Experiment 3 - Two targets inserted in the phantom.

7.1.2.2 Solid phantoms

Experiment 1 - variable depth. A solid dynamic phantom with tissue-like optical properties was used, which contained seven discrete targets impregnated with thermochromic pigment located at different depths from the surface. The targets consist of a mixture of thermochromic dye and polyester resin, and contain a 12Ω surface mount resistor together with a $4.7 k\Omega$ bead thermistor (to monitor the temperature) [Correia *et al.*, 2009a]. When heated, the pigment changes from black to white, with a nominal activation temperature of $47^\circ C$. The targets produced a change in absorption of approximately 0.6 mm^{-1} when heated above the activation temperature, which occurs after approximately 70 seconds. The phantom consisted of polyester resin mixed with titanium dioxide (TiO_2) powder, which produced a transport scattering coefficient $\mu'_s = 0.8 \pm 0.05 \text{ mm}^{-1}$, and an absorption coefficient $\mu_a = 0.001 \pm 0.0002 \text{ mm}^{-1}$. The background absorption is deliberately relatively low (representing that of fatty tissue at near-infrared wavelengths), enabling depth sensitivity to be evaluated at large depths. It had dimensions of $115 \text{ mm} \times 100 \text{ mm} \times 73 \text{ mm}$ and the targets were located 23 mm above the bottom surface of the block. The targets were 4.5 mm diameter cylinders with a length of 41 mm and were arranged in the phantom as illustrated in figure 7.3(a). This target arrangement minimises the imaging array dimensions necessary to cover all targets simultaneously and confines the targets to a central position, which takes advantage of the higher sensitivity in the centre of the imaging array. The smallest spacing between targets is 11.2 mm, which is sufficient to ensure that heating of a target does not cause a colour change in its neighbours. The target nearest the surface was at depth 5 mm and the deepest at 35 mm, as shown in figure 7.3(b). The target closest to the surface will be referred to as target 1, the second target closest to the surface as target 2 and so on.

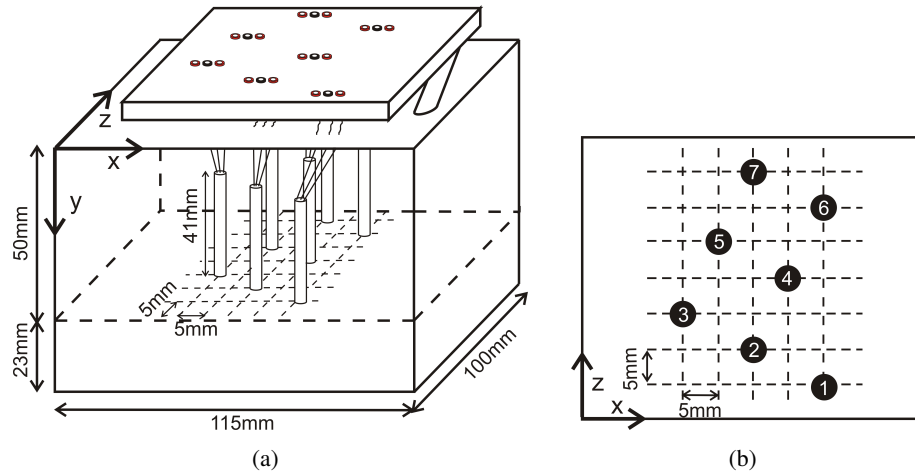


Figure 7.3: (a) Schematic of the solid phantom containing seven targets impregnated with thermochromic dye located at different depths. The targets are electrically activated through the sockets on top of the phantom; (b) Top view of the phantom showing the target arrangement. Targets are at depths from 5 mm to 35 mm in intervals of 5 mm.

The phantom construction was designed and produced at UCL. A detailed description of its construction can be found in [Correia *et al.*, 2009a].

Experiment 2 - fixed depth. Another electrically-activated phantom has been designed and produced at UCL based on an array of semiconductor diodes, which are used to heat targets containing thermochromic dye embedded in a solidified polyester resin with tissue-like optical properties [Hebden *et al.*, 2008a]. Sixteen targets were arranged in four rows of four, with the centre of each target 20 mm apart and at a depth below the top surface of 15 mm (figure 7.4). Each target can be activated individually. The targets are contained within a block of polyester resin mixed with TiO_2 powder, with the same optical properties as the phantom described in experiment 1. Each target was an 8 mm diameter cylinder with a height of 8 mm. The targets produced a change in absorption of approximately 0.6 mm^{-1} , when heated above the activation temperature.

Experiment 3 - two targets. The first electrically-activated dynamic tissue-equivalent phantom produced at UCL contained two targets impregnated with thermochromic pigment. A detailed description of its construction and results from transmission measurements performed across the phantom can be found in [Hebden *et al.*, 2008b]. The resin block surrounding the two targets has a transport scattering coefficient $\mu'_s = 1.0 \pm 0.1 \text{ mm}^{-1}$ and an absorption coefficient $\mu_a = 0.001 \pm 0.0002 \text{ mm}^{-1}$. The targets were placed 10 mm below the top surface and 30 mm apart. The targets contain different pigment concentrations, target A produced an absorption change of approximately 0.64 mm^{-1} and target B of 0.26 mm^{-1} , when heated above the activation temperature.

Data processing. In the first experiment, the probe was placed on the centre of the xy surface at $z=0$ mm (using the coordinate system shown in figure 7.3(a)), with the centre of the array immediately above the centre of target 2. In the second experiment, the probe was placed on the centre of the top surface. Data were acquired for the four central targets of the array. In the third experiment, the probe was placed directly above the two targets.

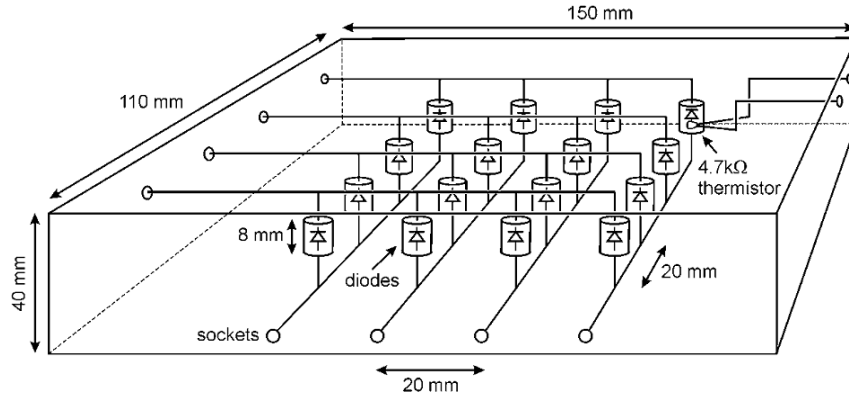


Figure 7.4: Solid phantom where each diode is embedded within a cylindrical target containing thermochromic pigment [Hebden *et al.*, 2008a].

Data were acquired for 20 s with the targets at room temperature, and then each target was heated in turn (except for experiment 3, where targets were activated simultaneously) for approximately 70 s, followed by a cooling period of 120 s to allow the targets to return to room temperature. For each target and each source-detector channel, a reference measurement was calculated from the mean signal of the 20 s recorded at room temperature. Similarly, the activation measurement was obtained by averaging over the 20 s segment of data centred on the time at which the maximum temperature was reached. Difference between the reference measurements and the corresponding activation measurements were used to generate three-dimensional images representing absorption changes occurring within the phantom. The Jacobian was calculated for a homogeneous medium with optical properties $\mu'_s = 0.8 \text{ mm}^{-1}$ and $\mu_a = 0.001 \text{ mm}^{-1}$.

7.2 Variable regularisation parameter

The L-curve method was employed to find the regularisation parameters for the simulation and liquid phantom with target T1 at different depths. This should provide a variation of λ with depth for these discrete positions, and it should be possible to find an expression that best fits the data to represent our depth-dependent regularisation parameter. Figure 7.5 shows the normalised regularisation parameters for the liquid phantom with target T1 at different depths, and exponential and sigmoidal fits to the data. The fit results are similar for the simulated data.

For the exponentially variant regularisation (signified by the label E), the expression found empirically was

$$\lambda_E = \exp\left(-\frac{z}{9}\right), \quad (7.3)$$

which is very similar to that found using the values obtained from the L-curve method. This expression was found by testing different exponential decays and observing the effect on the image reconstruction.

The image can be reconstructed using the following expression:

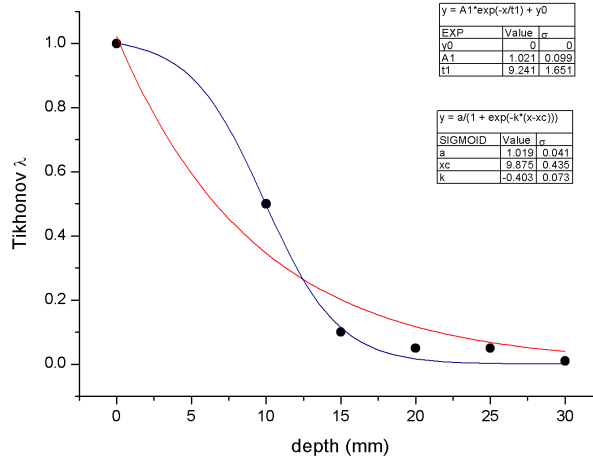


Figure 7.5: Exponential and sigmoidal fit to the normalised regularisation parameters found using the L-curve method, for the phantom with target T1 at different depths.

$$x_{\lambda_v} = \left(J^T J + (\lambda_c \lambda_v)^2 \right)^{-1} J^T b \quad (7.4)$$

$$= \lambda_v^{-1} \left(\lambda_v^{-1} J^T J \lambda_v^{-1} + \lambda_c^2 I \right)^{-1} \lambda_v^{-1} J^T b \quad (7.5)$$

$$= \lambda_v^{-2} J^T \left(J \lambda_v^{-2} J^T + \lambda_c^2 I \right)^{-1} b, \quad (7.6)$$

where λ_v is the spatially variant regularisation parameter, $\lambda_c = \lambda \sigma_{max}$ and σ_{max} is the maximum singular value of $J J^T$. The value λ is found using the L-curve method.

The sigmoid (S) function, obtained directly by fitting the data obtained from the L-curve for the target T1 at different depths, when used as a regularisation parameter did not produce images with improved spatial accuracy. Empirically, and for all the data, we found the optimal function to be:

$$\lambda_S = 1 - \frac{1}{1 + \exp\left(\frac{(z-20)}{6.5}\right)}. \quad (7.7)$$

Another method tested was the NOSER type regularisation, where the regularisation parameter takes the form

$$\lambda_{NOSER} = \text{diag} \left(J^T J + \gamma \right)^{1/2}, \quad (7.8)$$

where the constant $\gamma = \alpha \sigma_{max}^{-1}$ ensures that the regularisation parameter is never too small, and is approximately equal to the noise present in the data. As shown in §6, the discrete Picard condition can help find a reasonable estimate of the noise level. For experimental data $\alpha = 1$ most of the time, but this parameter can be adjusted to improve contrast and resolution, specially for perturbations located in deeper regions. The NOSER type regularisation can be used in equation (7.6), which we rewrite as

$$x_{\lambda_{NOSE R_1}} = \lambda_{NOSE R}^{-1} \tilde{J}^T (\tilde{J} \tilde{J}^T + \lambda I)^{-1} b, \quad (7.9)$$

where $\tilde{J} = J \lambda_{NOSE R}^{-1}$. The alternative expression where $\lambda_{NOSE R}$ is not squared can also be considered:

$$x_{\lambda_{NOSE R_2}} = (J^T J + \lambda_c^2 \lambda_{NOSE R})^{-1} J^T b \quad (7.10)$$

As shown by Endoh *et al* [2008], using this type of regularisation to reconstruct images gives uniform sensitivity and equal weights to the unknowns. They give an example of a simple under-determined problem with one equation and two unknowns, $Jx = b$, where $J = [j_1, j_2]$, $b = [b]$ and $x = [x_1, x_2]^T$. For a zero order Tikhonov with a constant regularisation parameter the solution is

$$x_{\lambda_c} = (J^T J + \lambda_c^2 I)^{-1} J^T b \quad (7.11)$$

$$= \left(\begin{bmatrix} j_1^2 & j_1 j_2 \\ j_1 j_2 & j_2^2 \end{bmatrix} + \begin{bmatrix} \lambda_c^2 & 0 \\ 0 & \lambda_c^2 \end{bmatrix} \right)^{-1} \begin{bmatrix} j_1 \\ j_2 \end{bmatrix} [b] \quad (7.12)$$

$$= \frac{1}{j_1^2 \lambda_c^2 + j_2^2 \lambda_c^2 + \lambda_c^4} \begin{bmatrix} j_2^2 + \lambda_c^2 & -j_1 j_2 \\ -j_1 j_2 & j_1^2 + \lambda_c^2 \end{bmatrix} \begin{bmatrix} j_1 \\ j_2 \end{bmatrix} [b] \quad (7.13)$$

$$= \begin{bmatrix} \frac{j_1 b}{j_1^2 + j_2^2 + \lambda_c^2} & \frac{j_2 b}{j_1^2 + j_2^2 + \lambda_c^2} \end{bmatrix}. \quad (7.14)$$

From this solution it is clear that if j_1 is larger than j_2 , then x_1 will have larger weighting than x_2 . Therefore this type of regularisation favours high sensitivity regions. On the other hand, if a variable regularisation parameter is used (equation 7.10), $\lambda_v = [\lambda_1, \lambda_2]$ and the solution becomes

$$x_{\lambda_v} = \begin{bmatrix} \frac{j_1 \lambda_2 b}{j_1^2 \lambda_2 + j_2^2 \lambda_1 + \lambda_1 \lambda_2} & \frac{j_2 \lambda_1 b}{j_1^2 \lambda_2 + j_2^2 \lambda_1 + \lambda_1 \lambda_2} \end{bmatrix}. \quad (7.15)$$

If λ_v is chosen to be proportional to $diag(A^T A)^{(1/2)}$, i.e., $\lambda_1 = \lambda_c^2 \sqrt{j_1^2}$ and $\lambda_2 = \lambda_c^2 \sqrt{j_2^2}$, then the solution has equal weighting

$$x_{\lambda_v} = \begin{bmatrix} \frac{b}{j_1 + j_2 + \lambda_c^2} & \frac{b}{j_1 + j_2 + \lambda_c^2} \end{bmatrix}. \quad (7.16)$$

If equation (7.9) is used the solution becomes

$$x_{\lambda_v} = \begin{bmatrix} \frac{b}{2j_1 + j_1 \lambda_c^2} & \frac{b}{2j_2 + j_2 \lambda_c^2} \end{bmatrix}, \quad (7.17)$$

which has larger weighting for regions with lower sensitivity and should prevent the reconstructed features from appearing artificially close to the surface. One can include the measurement covariance matrix, like in equation (7.2), which penalises noisier measurements [Adler *et al.*, 2007]. The measurement covariance matrix C is a diagonal matrix defined as $C = diag(\sigma_b^2) / max(\sigma_b^2)$, where σ_b^2 is the vector of measurement covariances, derived from the standard deviation of the measurements σ_b . The method that combines both NOSER and covariance matrix C in the image reconstruction is signified by the label NOSER&C.

Instead of using R in eq. (7.2) as a NOSER type matrix, the regularisation matrix can be defined as the image covariance, which is the noise sensitivity: $R = P/\max(P)$, where $P = \text{diag}(\text{inv}J C \text{inv}J^T)$ and $\text{inv}J = J^T (J J^T + \lambda_c^2 C)^{-1}$. This method is signified by the label R&C. The simpler case where $C=I$ can also be considered, which means that the regularisation is equal to the normalised image covariance for measurements with equal variance (signified by the label R).

The variance uniformisation (VU) constraint [Cohen-Bacrie *et al.*, 1997] is another statistical approach. Recall that the singular value decomposition of the matrix $J = USV^T$, hence the variable regularisation parameter is defined as

$$\lambda_{VU} = V D V^T, \quad (7.18)$$

where D is a diagonal matrix given by

$$D = \frac{S}{\sqrt{c}} - S^T S, \quad (7.19)$$

where c is a parameter that controls the level of variance of the reconstruction.

7.3 Image analysis

An analysis of the image quality was performed, to determine how the depth-dependent regularisation influences the resolution, spatial accuracy and contrast. It is anticipated that depth discrimination will be improved significantly.

Spatial resolution is often characterised by the Full Width at Half Maximum (FWHM) of the point spread function (PSF). The true PSF represents the image of an object of infinitesimal size. The image cross-sectional plane for each target is, in fact, the convolution of the true target structure with the PSF. Therefore the squared FWHM of the target image is approximately equal to the squared width of the true target added to the squared width of the PSF, i.e., $FWHM^2 = FWHM_{true}^2 + FWHM_{PSF}^2$. Profiles containing the highest intensity pixel in the image were generated and the FWHM was calculated for each direction. The estimates of the PSF width for each direction were calculated using $FWHM_{PSF} = \sqrt{FWHM^2 - FWHM_{true}^2}$ and then averaged. The error is given by the pixel size, which is 3.3 mm for the z direction and 3.5 mm for the x and y directions. For a three dimensional comparison between the images, we consider the normalised FVHM to be the image volume that contains voxels which have corresponding values greater than or equal to 50% of the maximum value divided by the total volume. The error associated with FVHM corresponds to the propagation of the uncertainty associated with the target volume, i.e., $|\frac{\partial FVHM}{\partial V_t}| \sigma_{V_t}$, where σ_{V_t} is the standard deviation of the target volume V_t .

Contrast is a measure of the ability to distinguish a feature based on its brightness in the image relative to the background. Image contrast in terms of μ_a can be expressed as $C = (\mu_{a_{max}} - \mu_{a_{bkg}}) / \mu_{a_{bkg}}$, where $\mu_{a_{max}}$ refers to the maximum absorption coefficient value in the ROI and $\mu_{a_{bkg}}$ is the mean background absorption coefficient. The error associated with derived contrast values corresponds to the propagation of the uncertainty associated with the calculation of the background absorption coefficient, i.e., $|\frac{\partial C}{\partial \mu_{a_{bkg}}}| \sigma_{\mu_{a_{bkg}}}$, where $\sigma_{\mu_{a_{bkg}}}$ is the

standard deviation of μ_{abkg} .

We define the spatial accuracy (SA) as a measure of the error between the true and the apparent positions of the target. The apparent position x_m is derived from the weighted mean position along the one dimensional profiles, which is calculated as $x_m = \sum_i^n \mu_{a_i} x_i / \sum_i^n \mu_{a_i}$. Again, the uncertainty in position corresponds to the intrinsic pixel size.

The total error for resolution, contrast and spatial accuracy, for each experiment, is calculated using the root squared error (RSE). For example, for the SA for i target positions the RSE is $\sqrt{\sum_i (x_{m,i} - x_{true,i})^2}$, where x_{true} is the true position of the target.

7.4 Results and discussion

Different regularisation types were used in the reconstruction of images from the simulated and liquid phantom data. An extensive analysis of the images focused on the depth recovery accuracy of the inclusions was performed to help us identify a suitable regularisation type. Once it was identified, the same was used to reconstruct images from the remaining measurements.

7.4.1 Simulations and liquid phantom experiments

Experiment 1. Images of the simulated and phantom data were reconstructed using the different types of regularisation. For the purpose of this analysis only the xz cross-sectional images of each target were used. Contrast, resolution and spatial accuracy were calculated for each image reconstructed using the different variable regularisation parameters. The results are summarised in tables B.1-B.4 in appendix B, and images are displayed from the simulations and phantom data with target T1, for the all regularisation types.

Figure 7.6 shows the weighted mean depth for the simulation versus the true depth of the target, for the different variable regularisation parameters. Those that use a statistical approach are shown in figure 7.6(a) and the remaining types in figure 7.6(b). The uncertainty in depth corresponds to the intrinsic uncertainty due to the pixel size, which is 3.3 mm. The straight line represents the expected depth. When the target is at 15 mm the image reconstruction using a constant regularisation performs better than any of the other methods. All the methods fail to accurately locate the target at 30 mm. Nevertheless, at 30 mm they all show a significant improvement compared to the reconstruction using a constant regularisation λ_c .

For the liquid phantom experiment, the targets were placed at different depth from 12 mm to 32 mm in intervals of 5 mm. For the phantom with target T1, it is clear from figure 7.7 that the reconstructed images using a spatially variant regularisation parameter provide better depth localisation than when only λ_c is used. However, all regularisation types fail to accurately locate the target at 32 mm depth.

Figure 7.8 shows a stacked bar chart, where each bar represents the global depth accuracy calculated using the RSE (z_{RSE}), for each regularisation type for both data sets. The spatially variant regularisation types analysed show a visible improvement in the depth localisation of the target, in comparison with the image reconstruction using a constant regularisation parameter. Overall, the differences between the different regularisation types are not very significant. Nevertheless, the NOSER type gives slightly higher depth accuracy.

The PSF width error was calculated in a similar manner as the depth accuracy error, i.e., using the RSE. Figure 7.9 displays a stacked bar chart of the PSF_{RSE} for all the regularisation

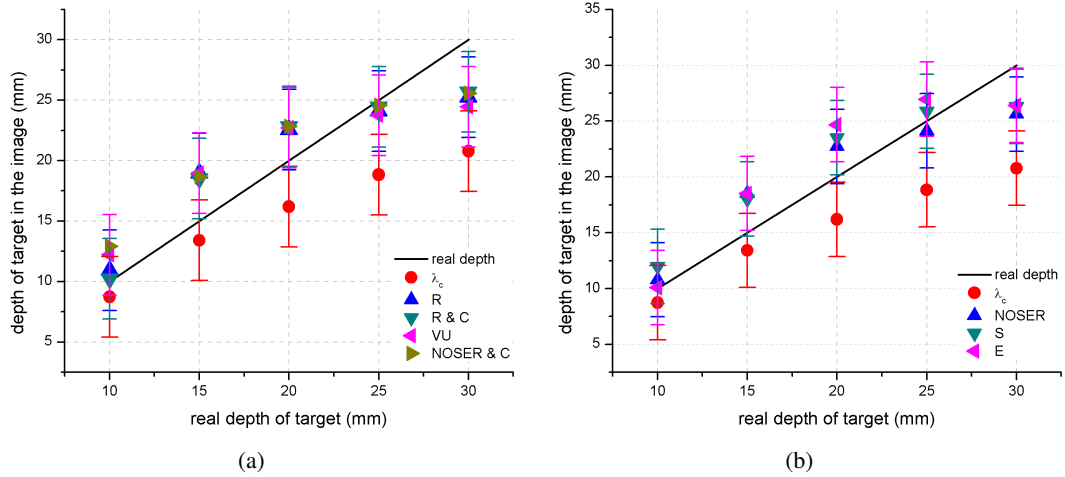


Figure 7.6: Graph showing how the apparent depth z_m of target in the image varies with true depth. Images reconstructed from simulated data using (a) regularisation parameters that use noise statistics and (b) the remaining types.

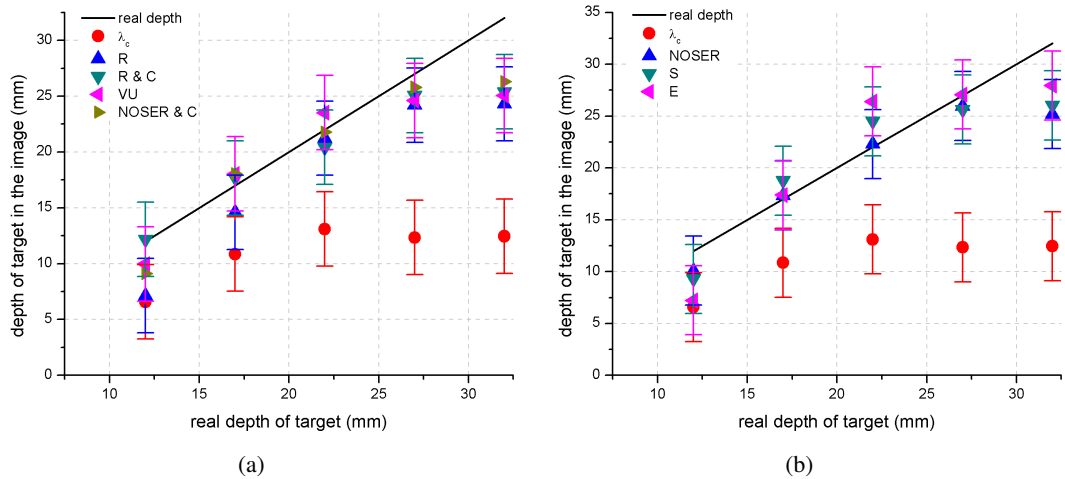


Figure 7.7: Graph showing how the apparent depth z_m of target in the image varies with true depth. Images reconstructed from measurements performed on the liquid phantom with target T1, using (a) regularisation parameters that use noise statistics and (b) the remaining types.

types. The VU regularisation yields the narrowest PSF. It also has the narrowest PSF for phantom images, whereas the sigmoidal regularisation yields the narrowest PSF for images reconstructed from simulated data. Images reconstructed using a constant regularisation parameter yield a narrow PSF, because the target is always located near the surface due to the higher sensitivity in this region, hence the perturbation appears confined to a smaller volume. Furthermore, the L-curve regularisation parameter selection method finds a λ_c which produces images reconstructed with a "good" resolution. Since the reconstructed images are three dimensional, it is necessary to analyse the FVHM error ($FVHM_{RSE}$) whose bar chart is displayed in figure 7.10. The apparent volume of the target in the images is larger for the R and R&C methods. The image resolution is highest for the VU or sigmoidal regularisation types.

The bar chart in figure 7.11, shows that the global contrast error, C_{RSE} , is higher for

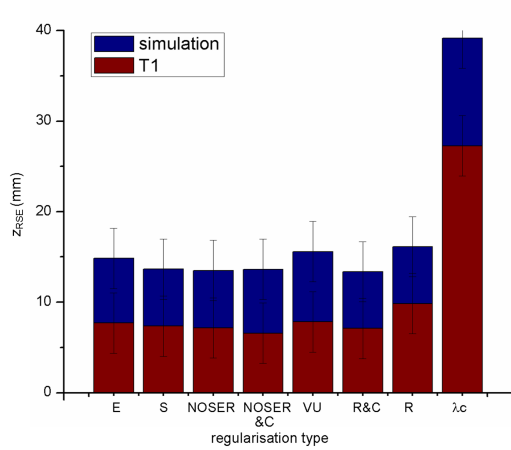


Figure 7.8: Each bar represents the depth root squared error (z_{RSE}) of the target in the images obtained using each regularisation type.

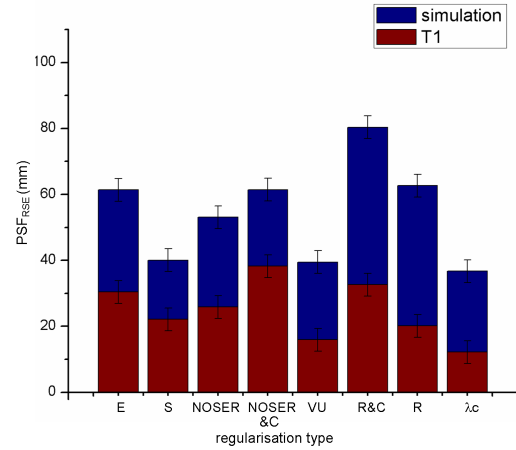


Figure 7.9: PSF width root squared error (PSF_{RSE}) of the reconstructed images using different regularisation types.

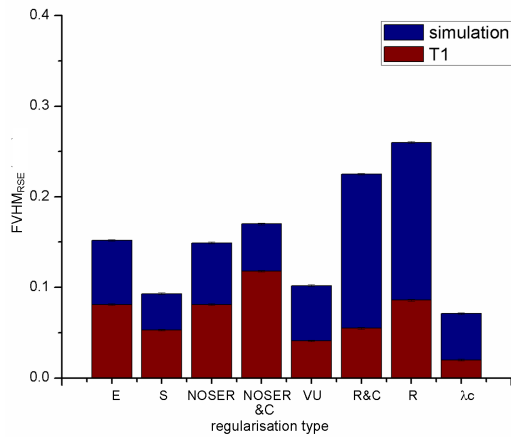


Figure 7.10: FVHM root squared error ($FVHM_{RSE}$) of the target in the images obtained using each regularisation type.

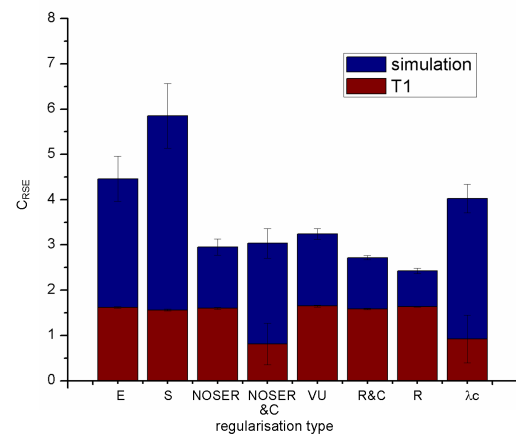


Figure 7.11: Contrast root squared error (C_{RSE}) of the reconstructed images using different regularisation types.

the regularisation types S and E. The other methods exhibit similar contrast. In general, the absorption change was underestimated in the phantom reconstructions and overestimated for the simulations when the target is at 10 mm depth. Contrast decreases with depth, and for this reason absorption changes occurring in deeper regions are underestimated (refer to appendix B for contrast values).

All the spatially variant regularisation types clearly improve the depth localisation of the target in the image. If only λ_c is used, one is not able to accurately locate targets placed deep below the surface. From the bar charts we find that the NOSER regularisation produces images with all the desired qualities: a reasonably good resolution, contrast and depth localisation. Using the NOSER type regularisation, all the apparent depths of the target in the image for the simulation and liquid phantom agree with the true depth within the associated uncertainty, except when the target is at 30mm and deeper. At 15 mm depth, the simulated target appears deeper in the reconstructed image. However, the point of maximum intensity is at 16.7 mm. For most regularisation types, the amount of blur in the z direction is higher when the simulated

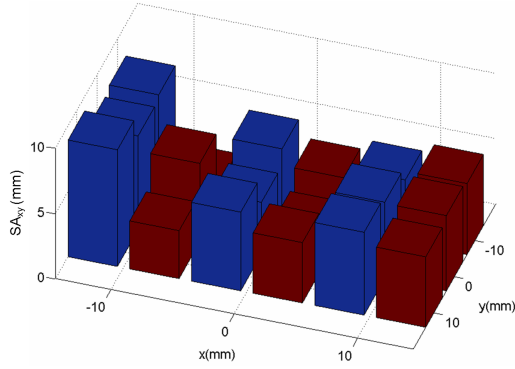


Figure 7.12: Lateral spatial accuracy of the target in the images obtained using constant regularisation (blue) and NOSER type regularisation (red).

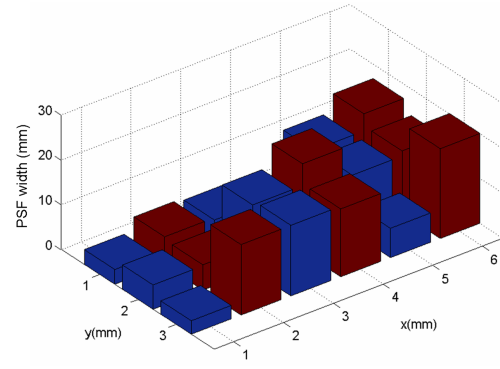


Figure 7.13: PSF width of images obtained using constant regularisation (blue) and NOSER type regularisation (red).

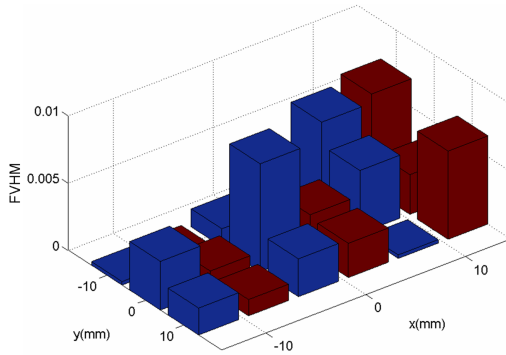


Figure 7.14: FWHM of the target in the images obtained using constant regularisation (blue) and NOSER type regularisation (red).

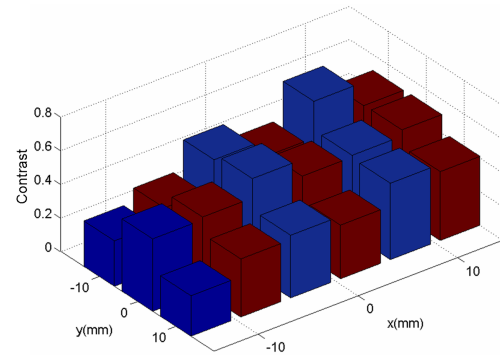


Figure 7.15: Contrast accuracy of the reconstructed images using constant regularisation (blue) and NOSER type regularisation (red).

target is at 15 mm depth (not shown) which leads to a decrease in the depth accuracy.

The NOSER penalty is applied to the whole three dimensional volume, instead of varying regularisation only in depth, which is an advantage. We used equation (7.10), which gave better depth accuracy at the expense of resolution and contrast. The VU method could be used as an alternative method. The exponentially variant regularisation could be adapted to vary in three dimensions, but that would require determining the decay rate for both x and y directions. When dealing with very large matrices, and if only correction in the depth direction is required, it is computationally more efficient to use methods such as S or E. These also produce images with good quality, but require less matrix manipulation.

Experiment 2. The effect of the NOSER regularisation on the lateral position of the target was analysed. The target was placed at 9 different (x, y) positions at 12 mm depth. The contrast, resolution and spatial accuracy for images reconstructed using a constant and a variable λ are summarised in appendix B. Figure 7.12 shows the lateral spatial accuracy SA_{xy} (calculated using the RSE) for each of the target positions, using a constant regularisation in the image reconstruction (blue bars) and the NOSER type regularisation (red bars). The SA_{xy} improves with NOSER. The PSF width increases with the use of NOSER (see figure 7.13). However, when we analyse the three dimensional image we find that the resolution increases for most

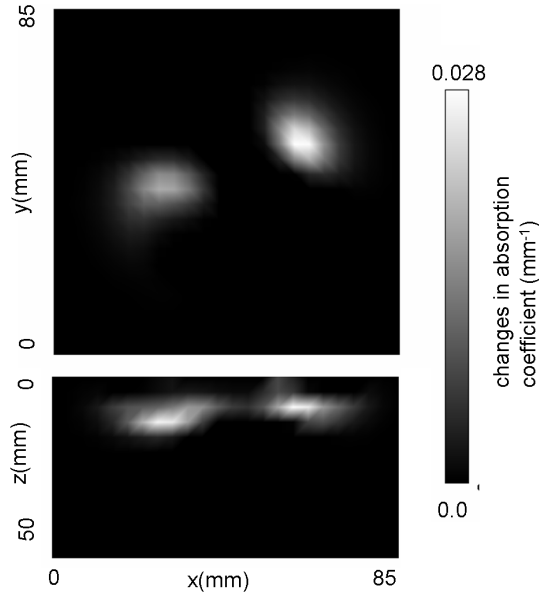


Figure 7.16: Image reconstructed using NOSER regularisation of targets T3 (left) and T2 (right).

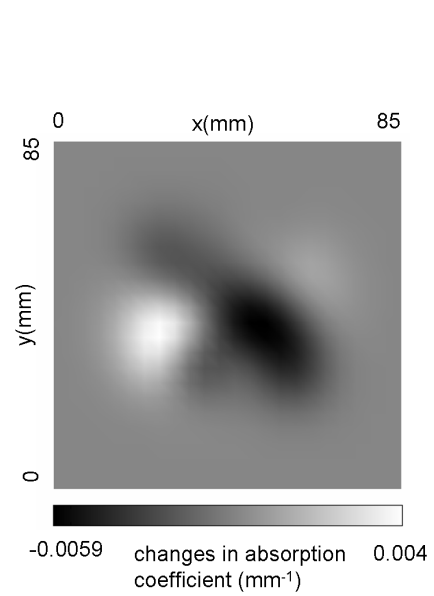


Figure 7.17: Image reconstructed using NOSER regularisation of targets T3 (left) and T4 (right).

of the target positions (see figure 7.14). The image contrast becomes more uniform when NOSER is used, as shown in figure 7.15. Thus, not only the accuracy of the reconstructed images improves when NOSER regularisation is used, but images are produced with a quality independent of the location of the target.

Experiment 3. The reconstructed image, using the NOSER regularisation, of targets T2 and T3 at 12 mm depth is shown in figure 7.16. Target T2 is located on the right and T3 on the left. The apparent separation of the targets is 31.9 ± 3.5 mm, which is much larger than the true separation. Furthermore, target T3 is misplaced along the y direction by 7.1 ± 3.5 mm. Figure 7.16 shows a xz plane that crosses the two targets, and we see that they appear at slightly different depths. However, it is possible that the targets were not exactly at the same depth and y position. The apparent separation of the targets in the image reconstructed using λ_c is the same. For this image, the contrast of target T2 is 0.2 ± 0.01 and T3 is 0.1 ± 0.02 , and the PSF widths are 4.7 ± 3.4 mm and 6.8 ± 3.4 mm, respectively. Contrast is better and the PSF width is narrower when the NOSER regularisation is used: the contrast of the target T2 is 2.5 ± 0.8 mm and PSF width is 1.1 ± 3.4 mm. For target T3 these values are 1.7 ± 0.2 mm and 5.4 ± 3.4 mm, respectively.

The image of targets T3 and T4 can be seen in figure 7.17. Target T3 is on the left side. The apparent separation of the targets in the NOSER reconstructed images is 24.3 ± 3.5 mm. Whereas, when a constant λ is used to reconstruct the images the apparent separation is 28.8 ± 3.5 mm. The reconstruction using the variant regularisation produces higher contrast, which is closer to the true value, 0.8 ± 0.2 for target T3 and -0.5 ± 0.06 for target T4, against 0.3 ± 0.06 and -0.21 ± 0.04 for the reconstructed images using a constant regularisation parameter. The PSF width is larger for target T4, 9.1 ± 3.4 mm (with λ_c is 10.3 ± 3.4 mm). For target T3 the value is 4.1 ± 3.4 mm (with λ_c is 6.8 ± 3.4 mm).

The apparent separation accuracy using λ_c is already quite high. Including NOSER in the

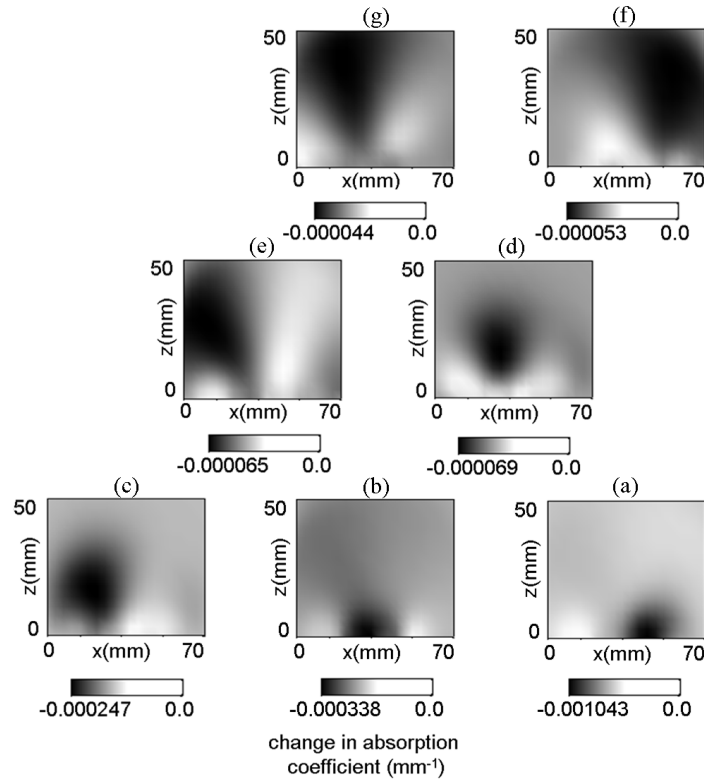


Figure 7.18: Optical topography images of each target in the xz plane. Bottom images from right to left correspond to (a) target 1 located at true depth $z_{true}=5$ mm, (b) target 2 at $z_{true}=10$ mm and (c) target 3 at $z_{true}=15$ mm. Middle images from right to left show (d) target 4 at $z_{true}=20$ mm and (e) target 5 at $z_{true}=25$ mm. Top images from right to left show (f) target 6 at $z_{true}=30$ mm and (g) target 7 at $z_{true}=35$ mm.

reconstruction algorithm improved the resolution and contrast.

7.4.2 Solid phantom

Experiment 1. For the purpose of this analysis only the xz cross-sectional images of each target were used. This gives us information regarding apparent depth and lateral positions of the targets. Figure 7.18 shows the images for each target in the xz plane at $y = 33$ mm. Note that the targets produce a localised decrease in absorption when they are activated electrically. The targets in the image have different relative locations, and visually they seem to be approximately accurate. It is also clear that the spatial resolution and contrast decreases with depth.

Profiles containing the highest intensity pixel in the image were generated for both the x and z directions and the FWHM was calculated in each case. Then, the estimates of the PSF width were calculated. Figure 7.19 shows the mean values of the PSF width for each target plotted as a function of depth, and as expected the resolution decreases with depth, except for the discrepancy in the general trend between targets 3 and 4. The lower resolution of target 3 is almost certainly due to the lower sensitivity of the imaging array to perturbations located near the edges, as opposed to higher sensitivity to perturbations in central locations.

As shown in figure 7.20, contrast decreases with increasing depth up to 20 mm depth, and beyond this point contrast values are very low. This means that the reconstructed absorption changes become less accurate with depth. The error associated with derived contrast values

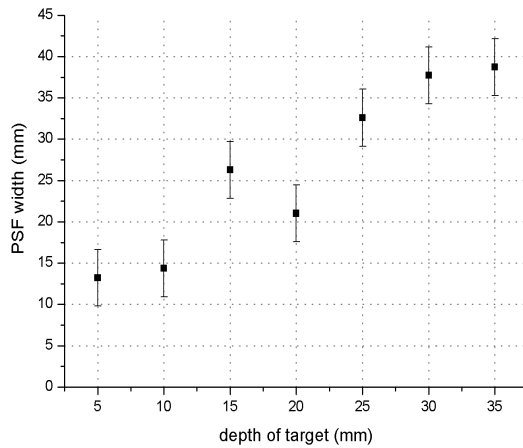


Figure 7.19: Graph showing how resolution varies with depth.

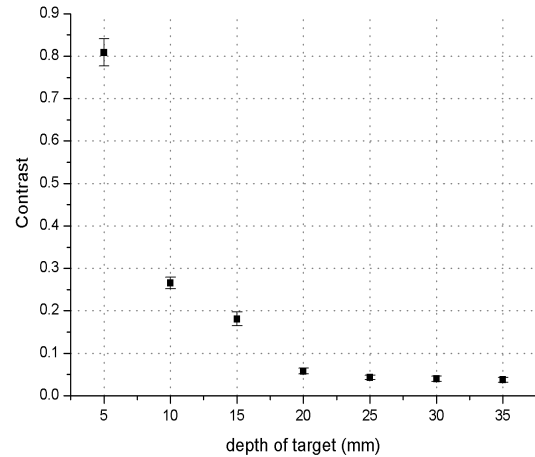


Figure 7.20: Graph showing how contrast varies with depth.

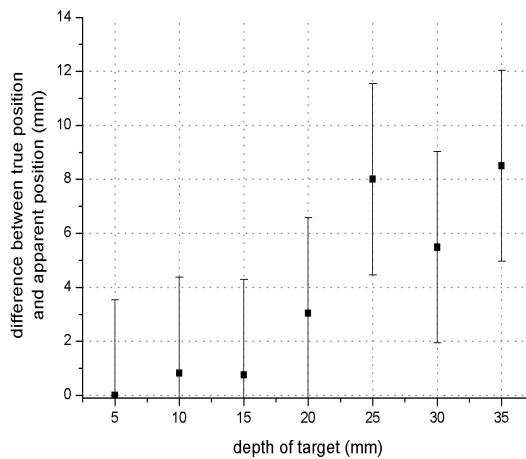


Figure 7.21: Graph showing the difference between the real target position and apparent position in the image, assuming that target closest to the surface is in the correct position.

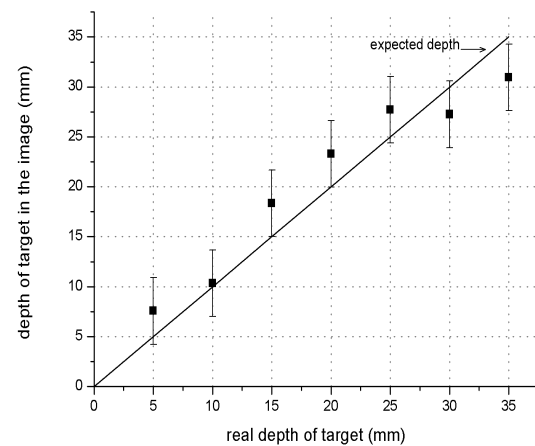


Figure 7.22: Graph showing how the depth of the target in the image varies with true depth. The expected depth is represented by the straight line.

increases from 4% of the background value for the first target to 15% for the deepest target.

It is assumed that the target closest to the surface was in the correct location, and therefore the positions of the other targets were calculated relative to this. As shown in figure 7.21, the lateral spatial accuracy reduces with depth, due to the lower sensitivity of the measurements and the smaller number of overlapping measurements at deeper regions. Meanwhile in figure 7.22, the calculated apparent position in the z direction reveals a very good agreement with the true depth of the target. The apparent depth agrees with the true depth within the associated error, at all depths up to at least 30 mm.

The results shown were obtained using the NOSER regularisation in equation (7.9). The results obtained using the alternative equation (7.10) were similar, although the image contrast was not quite as good. These images and results have been published in [Correia *et al.*, 2009a].

Summarising, the study presented here indicates that while spatial resolution degrades roughly linearly with depth, contrast falls exponentially. It is also shown that, with an appropriate choice of regularisation parameter, the apparent depth of a feature can match the

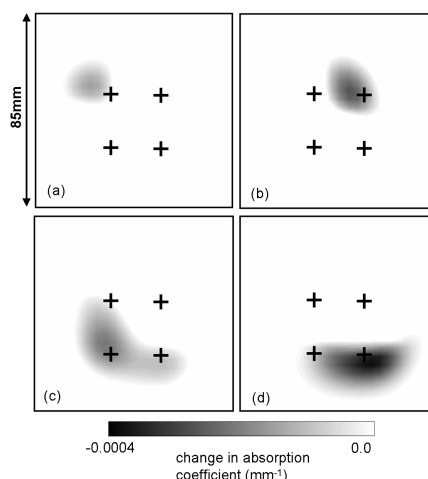


Figure 7.23: Images of four perturbations embedded in a solid phantom, reconstructed using a spatially variant regularisation parameter, NOSER. Targets were activated in the following order: (a) upper left; (b) upper right; (c) lower left; and (d) lower right. The crosses represent the approximate expected positions of the targets.

true depth within the statistical uncertainty of the measurement. Relatively good resolution, contrast and spatial accuracy were obtained for images reconstructed of targets up to 20 mm depth. For targets located deeper, our system was still able to reconstruct sensible images but with lower quality.

There are three factors which contribute towards the poor recovery of true contrast values shown in figure 7.20. First, the linear reconstruction algorithm inherently assumes that all changes in optical properties are small (infinitesimal). Second, the Jacobian matrix used in the reconstruction algorithm assumes homogeneous optical properties (as is common practice in scientific applications of optical topography), while the phantom is actually highly heterogeneous. The image reconstruction is also sensitive to the initial estimate of the optical properties used to calculate J . Although non-linear methods (where the Jacobian is updated iteratively [Arridge, 1999]) have been shown to provide more quantitatively accurate results, they are computationally complex and slow, and not suitable for real-time imaging of haemodynamic activity. Third, the decrease in spatial resolution with depth (shown in figure 7.19) implies that the contrast is spread over a greater volume at larger depths, and thus the maximum absorption value in the image is lower (hence the decrease observed in figure 7.20). The latter factor also applies to the previous studies as well as the sensitivity of the reconstruction to the initial values used to calculate the matrix J .

Experiment 2. The reconstructed images, using the NOSER regularisation, of the four targets embedded in the solid phantom are shown in figure 7.23. The images correspond to a depth of 15 mm below the surface. The targets were activated sequentially, from (a) to (d). The crosses indicate the approximate position of the targets. The absorption changes caused by heating the target roughly occur in the expected region. Figure 7.23(a) shows the lowest contrast, possibly because it was the first target to be activated. Since the phantom is initially at room temperature, the target cools more quickly and produces lower contrast. The reconstruction underestimated the true change, which is expected since linear reconstruction can only provide realistic values

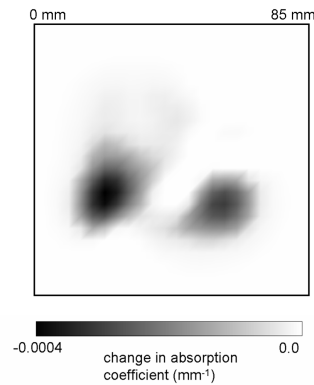


Figure 7.24: Images of two targets embedded in a solid phantom, reconstructed using NOSER types regularisation.

in the presence of small perturbation.

The images reconstructed using a constant regularisation parameter do not differ significantly. However, the first activation is reconstructed closer to the expected position and resolution is slightly better. The images reconstructed using λ_c have been published in [Hebden *et al.*, 2008a], along with images from another phantom. This other phantom consisted of a block of polyester resin, TiO_2 powder mixed together with thermochromic pigment through the volume. For this phantom design the absorption changes spread through an unknown volume. Therefore, it is not as suitable as the phantom studied to quantify spatial accuracy.

Experiment 3. Figure 7.24 shows a decrease in absorption produced by the two targets embedded in the solid phantom. This image was obtained using the NOSER regularisation and corresponds to a depth 10 mm below the surface. The apparent separation of the targets is 35 mm, which is slightly larger than the true separation of 30mm. Target A is on the left, showing a larger absorption change, and target B on the right. As before, the absorption change is underestimated. Including the NOSER regularisation in the reconstruction increased contrast and resolution compared to images reconstructed using λ_c (published in [Hebden *et al.*, 2008b]).

7.5 Conclusions

One of the limitations of optical tomography as a brain imaging tool is the low sensitivity to haemodynamic activity occurring within deeper tissues compared to that occurring nearer the surface. Nevertheless, considerable effort has been devoted to the development of image reconstruction algorithms capable of generating three-dimensional images from diffuse reflectance data, and schemes have been devised which aim to improve the ability of optical tomography to discriminate along the depth direction.

Simulated data were generated and measurements on phantoms were performed, for perturbations in the absorption coefficient placed at different spatial locations. Several spatially variant regularisation types were evaluated in the image reconstruction algorithm, and their performance was analysed in terms of contrast, resolution and spatial accuracy. They all had a significant effect on the depth accuracy, and performed much better than the simple

reconstruction using a constant regularisation parameter. The influence is more apparent for perturbations deeper than 15 mm; down to that depth the images reconstructed using λ_c have an accuracy similar to that obtained when a variable regularisation parameter is used. The NOSER type regularisation improves the quality of the images, in all directions, by applying a different regularisation to each voxel. When imaging two perturbations placed side by side at the same depth, the effect of the NOSER regularisation is more obvious on the contrast and resolution, although the spatial accuracy slightly improves. However, this study was performed at a very superficial depth, and if tests had been performed using targets deeper in the phantom, the benefit of the NOSER type regularisation would be more obvious. Contrast and resolution decrease with depth, and an increase in the number of overlapping measurements at deeper regions could potentially improve the quality of the images of deeper located targets.

This study to quantify the quality of the reconstructed images was valuable in order to understand the potential and limitations of the UCL optical topography system, and the results can inform the interpretation of results of clinical studies. It also allows a quantitative comparison between optical topography systems and other medical imaging modalities.

In conclusion, using the NOSER type variable regularisation parameter in the image reconstruction algorithm, enables good depth discrimination to be obtained from optical topography measurements.

Identification of the optimal wavelengths using photon density measurement functions

The choice of wavelengths for NIRS is limited by the "transparency" window for biological tissue, which is approximately from 600 nm to 900 nm. In this region the absorption by the main chromophores (oxy-haemoglobin, deoxy-haemoglobin and water) is relatively low, which allows the light to travel relatively deep into the tissue. The wavelength pair 780 nm and 830 nm has been widely used because each wavelength is located symmetrically on either side of the isosbestic point at 805 nm, where the HbO₂ and HHb specific extinction coefficients ϵ cross over. These wavelengths are also widely available as laser diode sources. However, this wavelength pair may not necessarily be the optimal choice for optical topography imaging. In this chapter a method to select the optimal wavelengths for optical topography is proposed based on the similarity between volumes interrogated by NIR light.

A few NIRS studies have been published which explore the selection of an optimal wavelength pair that minimises cross-talk, where a change in [HbO₂] may replicate a change in [HHb] and vice versa, and reduces the noise level of the calculated chromophore concentration changes caused by noise in the measurements [Uludağ *et al.*, 2004; Yamashita *et al.*, 2001]. Most NIRS approaches use the MBLL (mentioned previously in §3.4) to calculate the changes in the chromophore concentrations. The MBLL is an empirical description of light propagation through scattering and absorbing media, where a change in the logarithm of the detected light intensity is proportional to a change in the chromophore concentrations: $-\ln \Delta I = \epsilon \Delta c dDPF$, where d is the distance between the points where the light enters and leaves the medium and DPF is the differential pathlength factor [Boas *et al.*, 2001a; Strangman *et al.*, 2002]. The DPF accounts for the increase in the distance travelled by light due to scattering. It depends on tissue absorption, scattering and optode separation, and therefore is wavelength dependent [Boas *et al.*, 2001a].

Yamashita *et al* [2001] have shown that the most accurate measurements of the concentration changes in the brain (motor cortex) using NIRS are obtained for wavelengths

below 700 nm paired with 830 nm. Meanwhile, Strangman *et al* [2003] showed experimentally and theoretically that the wavelength pair 691 nm or 752 nm with 830 nm produces more accurate results than the pair 780 nm and 830 nm, which exhibits a high sensitivity to cross-talk and to all sources of noise. The 780/830 nm pair is more sensitive to errors in the DPF than other pairs. For the 780/830 nm pair Strangman *et al* [2003] also verified that a small error in the optical properties results in a large error in the calculated oxygen saturation. They state that the sensitivity to measurement errors is so large that they gain very little benefit by using wavelengths with overlapping spatial sensitivities, which theoretically should reduce cross-talk. They proved analytically that the cross-talk is proportional to the difference between the partial pathlength factor (PPF). For an inhomogeneous change in μ_a the medium is divided into partial volumes of uniform absorption change, and the DPF is replaced by the PPF through a region of uniform absorption change at two wavelengths. They concluded that any wavelength in the range of roughly 770 nm - 800 nm paired with 830 nm will result in poor chromophore concentration information.

Another NIRS study of the adult head that yielded similar results was performed by Uludağ *et al* [2004]. They evaluated theoretically, for the full wavelength range, and experimentally different wavelength pairs in terms of cross-talk and noise present in the calculated chromophore concentration changes. They concluded that one of the wavelengths has to be between roughly in the range 650 nm - 720 nm and the other above 820 nm, in order to achieve both low cross-talk and noise. Okui *et al* [2005] performed a theoretical study where they analysed the cross-talk between $\Delta[\text{HbO}_2]$ and $\Delta[\text{HHb}]$ and vice-versa, since these have different magnitude and sign. They found that the optimal wavelength to pair with one at 830 nm should be in the interval 690 nm - 750 nm. They also showed that the change in the thickness of the superficial layers influences the PPF and consequently the cross-talk. Umeyama *et al* [2009] proposed a new cross-talk measure, based on the mean squared error of the calculated concentration changes, which deals with both $\Delta[\text{HbO}_2]$ and $\Delta[\text{HHb}]$ cross-talk.

Sato *et al* [2006] investigated the dependence of the DPF on 4 wavelength pairs and for 3 different source-detector distances. They analysed the hemodynamic activation signals from the visual cortex and concluded that for 692 nm, 759 nm and 782 nm paired with 830 nm the DPF was approximately the same, suggesting that the volume investigated is the same. However, for the pair 678/830 nm there was possibly a slightly difference in the spatial sensitivity. In a previous study Sato *et al* [2004] found that using a 692 nm wavelength paired with 830nm increases the signal to noise ratio compared to the 780/830 nm pair.

Corlu *et al* [2005] introduced the first method for choosing the optimal wavelengths for diffuse optical imaging. The non-uniqueness problem associated with CW imaging is manifested as cross-talk between the images of absorption and scattering. A multispectral reconstruction method can help overcome the non-uniqueness problem. Instead of converting the absorption and scattering coefficients at each wavelength into the chromophore concentrations and scattering components, using images reconstructed with the conventional method, all the measurements at all wavelengths are used to reconstruct directly the chromophore concentrations, which are wavelength independent (refer to §5.2.2.2 for more detail). Corlu *et al* [2005] introduced two conditions for choosing the measurement wavelengths

that give optimum absorption and scattering separation and also optimum separation between the main breast tissue chromophores (HbO₂, HHb, water and Lipid). Based on this method, Eames *et al* [2008] explored the influence of using data for the full spectrum on the reconstructions. They concluded that by using optimal spectral windows instead of the full spectrum, not only the reconstruction is computationally faster but cross-talk is reduced, which suggests that not all wavelengths add useful information.

One cannot guarantee that the volumes over which measurements are sensitive to changes in absorption and scattering are identical for the wavelengths selected by the two conditions defined by Corlu *et al* [2005]. This issue is of importance in order to accurately locate optical property changes occurring within the object of study, since different wavelengths may have different sensitivities to the same changes in the medium and the same measured change could be associated with changes in different regions within the reconstructed volume. In this study the optimal wavelength set is estimated by incorporating a third condition which takes into account the interrogated volume. The spatial similarities between sensitivity profiles for different wavelengths are analysed and quantified.

8.1 Unique and simultaneous reconstruction

The non-uniqueness problem associated with using CW measurements results in cross-talk between absorption and scattering. Minimising the non-uniqueness of the multispectral method by selecting the optimal wavelengths represents one of the conditions introduced by Corlu *et al* [2005]. Rewriting the non-uniqueness conditions introduced by Arridge and Lionheart [1998] in terms of the wavelength independent variables in the multispectral method and considering b in equation 2.7 to be constant for simplicity, it follows that the inverse problem has a non-unique solution when the residual norm R [Corlu *et al.*, 2005]

$$R = \left\| \mathbf{1} - E (E^T E)^{-1} E^T \mathbf{1} \right\| \quad (8.1)$$

is zero, hence the *maximisation* of R is required. E is a wavelength dependent matrix of the form

$$E(\lambda) = \begin{pmatrix} \frac{\varepsilon_1(\lambda_1)}{\lambda_1^b} & \frac{\varepsilon_2(\lambda_1)}{\lambda_1^b} & \dots & \frac{\varepsilon_i(\lambda_1)}{\lambda_1^b} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\varepsilon_1(\lambda_n)}{\lambda_n^b} & \frac{\varepsilon_2(\lambda_n)}{\lambda_n^b} & \dots & \frac{\varepsilon_i(\lambda_n)}{\lambda_n^b} \end{pmatrix}. \quad (8.2)$$

The ability to separate chromophore concentrations can be evaluated from the matrix form of equation (2.3) [Corlu *et al.*, 2005]

$$\begin{pmatrix} \Delta\mu_a\lambda_1 \\ \Delta\mu_a\lambda_2 \\ \dots \\ \Delta\mu_a\lambda_n \end{pmatrix} = \begin{pmatrix} \varepsilon_1(\lambda_1) & \varepsilon_2(\lambda_1) & \dots & \varepsilon_i(\lambda_1) \\ \vdots & \vdots & \vdots & \vdots \\ \varepsilon_1(\lambda_n) & \varepsilon_2(\lambda_n) & \dots & \varepsilon_i(\lambda_n) \end{pmatrix} \begin{pmatrix} \Delta c_1 \\ \Delta c_2 \\ \dots \\ \Delta c_i \end{pmatrix}, \quad (8.3)$$

using the condition number \mathbf{k} of the extinction coefficient matrix $\varepsilon_i(\lambda)$, which is the ratio of the largest to the smallest singular value of the matrix. A small $\mathbf{k}(\varepsilon_i(\lambda))$ indicates a smooth singular value decay resulting in a reduced sensitivity of the system to errors in the measurements,

and therefore chromophores contribute evenly to absorption. In order to distinguish between chromophore concentrations the wavelength sets chosen should *minimise* the condition number $\mathbf{k}(\varepsilon_i(\lambda))$, which defines the second condition. This condition number also describes the linear independence of the system of equations. When $\mathbf{k}(\varepsilon_i(\lambda))$ is small the matrix has full rank, and therefore the wavelength set provides distinct and independent information.

8.2 Interrogated volume

In optical topography brain activation studies using multiple wavelengths it is invariably assumed that measurements at each wavelength contain information from the same cerebral region. If not, data at some wavelength could be measuring a region where the activation occurs whereas the data at others could be more sensitive to surrounding areas, and the study may return erroneous results. Each row of the Jacobian matrix J represents a PMDF, which is the measurement sensitivity of a source-detector pair to a perturbation in the optical properties of the object of study [Arridge, 1995]. In terms of image reconstruction, if the forward model is calculated for the same optical properties but, however, the measurement wavelengths have different sensitivities to changes in absorption and scattering, then for a perturbation in the optical properties of the tissue, the reconstructed perturbation will occur in different spatial locations. In practice, the optical properties of tissues beneath the probe are unknown, and in order to reconstruct changes in tissue properties as accurately as possible, it is necessary to acquire data at wavelengths where the sensitivities at each wavelength coincide.

The aim of this study is to find which combinations of wavelengths have overlapping sensitivities, which is an indicator of the similarity between interrogated volumes, by comparing PMDFs for a certain source-detector pair using different measurement wavelengths.

8.3 Methods

8.3.1 Similarity Measures

It is essential to find a reliable method of comparing the sensitivity profiles generated for the different wavelengths. Several similarity measures commonly used in image registration are described below and their adequacy for comparing PMDFs is evaluated. These techniques are explained in more detail in [Hill *et al.*, 2001].

8.3.1.1 Sum of squared differences

The sum of squared differences (SSD) method calculates the sum of differences between the squares of the values in the two PMDFs. For PMDFs X and Y with voxels i , the SSD is defined as

$$SSD = \frac{1}{N} \sum_i^N |X(i) - Y(i)|^2 \quad \forall i \in X \cap Y, \quad (8.4)$$

where N is the number of voxels. The SSD will be zero when the distributions are identical and will increase with increasing difference. This similarity measure is optimum when the images differ only by Gaussian noise [Hill *et al.*, 2001]. One of its disadvantages is the high sensitivity to a small number of voxels that have large differences between the PMDFs [Hill *et al.*, 2001].

8.3.1.2 Ratio-image uniformity

An alternative similarity measure is the ratio-image uniformity (RIU), which has been employed to register PET images [Woods *et al.*, 1992]. It is based on the assumption that if two images are aligned, then the voxels in the images are related by a scaling factor. If the images are misaligned, then this scaling factor is no longer constant. The RIU is found by *minimising* the normalised standard deviation of the ratio image r calculated from images X and Y as follows [Hill *et al.*, 2001]

$$\begin{aligned} r(i) &= X(i)/Y(i) \\ RIU &= \frac{\sigma_r}{\mu_r}, \end{aligned} \quad (8.5)$$

where σ_r is the standard deviation of r and μ_r is the mean of r , both calculated over the number of voxels N in the overlap domain. Because RIU may be biased to the image used as reference, the roles of the two images can be switched, and the two normalised standard deviations are then averaged, hence the ratio image becomes

$$r(i) = \left(\frac{X(i)}{Y(i)} + \frac{Y(i)}{X(i)} \right) / 2. \quad (8.6)$$

A disadvantage of this method is the high sensitivity to noise for low value voxels. For example, comparing two images that only differ by noise, if for the noise-free image one of the background values is 1 and the corresponding voxel in the noisy image is 5, then the average ratio is 2.6. This value should be close to 1, which happens for images with larger voxel values and the same noise level.

8.3.1.3 Correlation coefficient

The correlation coefficient (CC) is the optimum measure if the images X and Y are linearly dependent. This involves *maximising* the function

$$CC = \frac{\sum_i^N (X(i) - \bar{X})(Y(i) - \bar{Y})}{\left[\sum_i^N (X(i) - \bar{X})^2 (Y(i) - \bar{Y})^2 \right]^{1/2}}, \quad (8.7)$$

where \bar{X} and \bar{Y} are the means of images X and Y , respectively. Hence, when the correlation coefficient is equal to 1 the images are linearly related. If the variables are independent they are uncorrelated, and hence the CC value is zero. However, if two variables are (linearly) uncorrelated it does not necessary follow that they are independent. If the CC is negative it means that the variables are inversely correlated; the values of one variable increase and the values of other variable decrease. CC fails when either or both images have uniform values.

8.3.1.4 Joint histogram and joint entropy

The joint histogram is a plot of the value of each voxel in image X against the value of the corresponding voxel in image Y . The values in the histogram represent the number of voxels for a certain intensity value pair. If two images are aligned and identical the joint histogram has tight clusters, and the joint entropy is minimised. These clusters disperse as the images become less similar, resulting in an increase in entropy. Therefore, we want to *minimise* the join entropy

H, which is defined as

$$H(X, Y) = - \sum_{i,j} p_{X,Y}(x_i, y_j) \log p_{X,Y}(x_i, y_j), \quad (8.8)$$

where $p_{X,Y}(x_i, y_j)$ is the joint probability distribution function, which is the normalised joint histogram. It represents the probability of value x_i occurring in image X simultaneously with value y_j in image Y . Entropy is never negative. If all voxels in the images have the same intensity, then the histogram will only have a single element with probability 1, which gives $H=0$. This method can fail when images are misaligned and only overlap in regions that contain noise.

8.3.1.5 Mutual information

Mutual information (MI) measures the statistical dependence between images, and it requires *maximising* the function

$$MI(X, Y) = H(X) + H(Y) - H(X, Y), \quad (8.9)$$

where $H(X, Y)$ is the joint entropy, and $H(X)$ and $H(Y)$ are the marginal entropies for image X and Y respectively, which can be calculated using the following equations

$$H(X) = \sum_i \left(\sum_j p_X(x_i, y_j) \log \sum_j p_X(x_i, y_j) \right), \quad (8.10)$$

$$H(Y) = \sum_j \left(\sum_i p_Y(x_i, y_j) \log \sum_i p_Y(x_i, y_j) \right). \quad (8.11)$$

The marginal probability distribution p_X is a projection of the normalised joint histogram onto the axes corresponding to values occurring in image X , which is basically the sum of the columns of the normalised joint histogram. The marginal probability distribution p_Y represents the equivalent for image Y , and is the sum of the rows of the normalised joint histogram. To maximise mutual information it is necessary to minimise joint entropy and maximise the marginal entropies. Rewriting equation (8.9) in terms of the conditional entropy $H(X|Y)$ ($H(Y|X)$), which represents the uncertainty in X (Y) when Y (X) is known, then MI becomes:

$$MI(X, Y) = H(X) - H(X|Y) = H(Y) - H(Y|X). \quad (8.12)$$

This method measures the amount of information in one image that can represent the other image. If we can exactly predict the values in one of the images using information from the other image, then the conditional entropy will be zero, and the mutual information reduces to $H(X)$ ($H(Y)$) alone.

Another approach is the normalised mutual information (NMI), which in some cases has been shown to perform better than MI [Hill *et al.*, 2001]

$$NMI = \frac{H(X) + H(Y)}{H(X, Y)}. \quad (8.13)$$

8.3.1.6 Test photon density measurement functions

To ensure that the similarity measures return sensible results when comparing PMDFs we generated PMDFs for different optical properties, chosen to be different enough to return distinct sensitivity profiles, and compared their SSD, RIU, CC, H, MI and NMI. A single source-detector pair 40 mm apart was used in the forward model, because it has been suggested theoretically that this was sufficient to measure haemodynamic changes in adult grey matter [Okada *et al.*, 1997]. We normalised the PMDFs to their maxima for absorption and diffusion separately. Sensitivity values in the PMDF which are below 1% of the maximum value are set to zero, since it does not compromise the quality of the reconstructed images [Eames *et al.*, 2007] and eliminates undesirable contributions from the corresponding nodes to the similarity measures. SSD are calculated for N equal to the number of non-zero differences, i.e., for the number of voxels with sensitivity above 1% from the PMDF with largest sensitivity volume. This returns SSD values independent of the PMDF volume. On the other hand, for RIU, N is chosen as the number of overlapping pixels with sensitivity higher or equal to 1%, in order to avoid dividing by small background values. For CC, the image mean is calculated for values above the threshold. For each PMDF, the depth centroid $z_{centroid}$ was calculated along the z axis at position $(x, y) = (42.5 \text{ mm}, 42.5 \text{ mm})$

$$z_{centroid} = \frac{\sum_i^{N_z} A_i z_i}{\sum_i^{N_z} A_i}, \quad (8.14)$$

where A is the one-dimensional depth sensitivity along this axis, z is the distance from the xy plane and N_z is the number of z pixels. The centroid can be considered to be the average penetration depth. For the purpose of this analysis it is sufficient to analyse the absorption sensitivity profiles.

The first test consisted of comparing a reference PMDF with PMDFS whose input optical properties both change, μ_a is fixed and μ'_s is varied and vice-versa. The second test compares PMDFs with similar sensitivity volumes, for both large and small volumes.

8.3.2 Adult head model

The software package TOAST was used to model the propagation of light in tissue at wavelengths from 650 nm to 990 nm spaced in 10 nm intervals. To generate PMDFs for each wavelength one needs to insert the corresponding optical properties of the medium in the forward problem. Initially the medium is considered to be homogeneous with optical properties of grey matter. Subsequently a three layer model is used to generate PMDFs, which consists of two extracerebral layers with the optical properties of skin and skull and an intracerebral layer with grey matter optical properties.

8.3.2.1 Homogeneous model

The first model consisted of a slab with dimensions 85 mm \times 85 mm \times 30 mm with the optical properties of grey matter, whose main chromophores in the near infrared region are hemoglobin, water, lipid, and cytochromes. The extinction coefficients of water, cytochrome, HHb and HbO₂ are described in §2.2, along with the typical chromophore concentrations. The optical properties are $\mu_{a,brain}$ and $\mu'_{s',brain}$, given by equation (2.15) and (2.16), respectively.

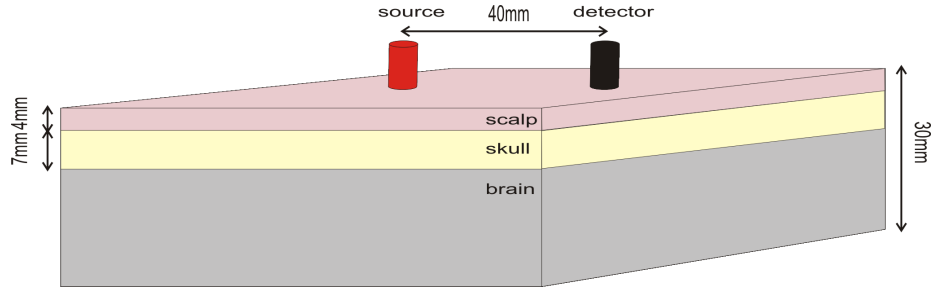


Figure 8.1: Three layer adult head model and probe placement.

8.3.2.2 Three layer model

Figure 8.1 shows the second adult head model, which is a slab with the same dimensions as the homogeneous model, consisting of three layers that imitate the scalp, skull and intracerebral tissue. Scalp is represented by a 4 mm thick layer and skull is 7 mm thick. Absorption and scattering spectra for each layer are described in §2.2.

8.3.3 Optimal wavelengths

Sensitivity profile maps (PMDFs) were obtained for the wavelength range 650-990 nm in steps of 10 nm, making a total of 35 wavelengths. Initially the PMDFs were combined in sets of 3 wavelengths (C_3^{35}), making a total of 6545 wavelength sets and 19635 PMDF pairs to be compared (3×6545 , since for each wavelength set 3 PMDF pairs are compared, considering repetitions). Thereafter, wavelength sets of 4 wavelengths are used, which results in 52360 sets and 209440 PMDF pairs. For a medium with i absorption chromophores at least $n = i + 1$ wavelengths are required when the scattering parameter a is allowed to vary [Corlu *et al.*, 2005]. Therefore, in order to find the solution for both absorption and scattering coefficients, for a medium with two chromophores, it is necessary to employ at least 3 wavelengths. Several optical topography systems acquire data at two wavelengths. Therefore, all possible wavelength pairs are compared (total of 595), to provide an insight into what can be achieved using two measurement wavelengths only.

The aim of this study is to find which wavelength set gives the best combination of: i) good separation between absorption and scattering, ii) good separation between $[\text{HbO}_2]$ and $[\text{HHb}]$ and iii) excellent overlap between sensitivity profiles. Quantitatively we aim to find the wavelength sets which maximise the residual norm R (equation (8.1)), minimise the condition number \mathbf{k} of the matrix $\varepsilon_i(\lambda)$ (equation (8.3)) and have small sensitivity profiles SSD (equation (8.4)). The matrix E in equation 8.2 for three wavelengths is reduced to:

$$E(\lambda) = \begin{pmatrix} \frac{\varepsilon_{\text{HbO}_2}(\lambda_1)}{\lambda_1^b} & \frac{\varepsilon_{\text{HHb}}(\lambda_1)}{\lambda_1^b} \\ \frac{\varepsilon_{\text{HbO}_2}(\lambda_2)}{\lambda_2^b} & \frac{\varepsilon_{\text{HHb}}(\lambda_2)}{\lambda_2^b} \\ \frac{\varepsilon_{\text{HbO}_2}(\lambda_3)}{\lambda_3^b} & \frac{\varepsilon_{\text{HHb}}(\lambda_3)}{\lambda_3^b} \end{pmatrix}, \quad (8.15)$$

where the scatter power b is considered to be 0.53 [Matcher *et al.*, 1997].

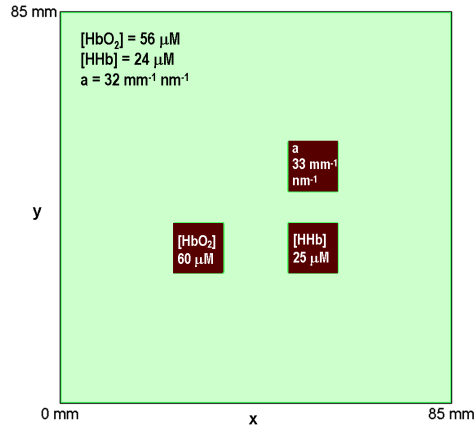


Figure 8.2: Cross-section xy at $z=11.7$ mm of the medium used in the simulation.

Finally, we quantify the differences between PMDFs in terms of penetration depth, volume and sensitivity.

8.3.4 Image reconstruction

Simulations were performed to compare the cross-talk effects in images reconstructed from measurements at wavelengths selected using our method against those from measurements at wavelengths selected using Corlu's method [Corlu *et al.*, 2005]. The simulated medium consisted of a slab with 59202 nodes, resulting in 38572 elements, and dimensions $85 \text{ mm} \times 85 \text{ mm} \times 30 \text{ mm}$. The UCL array with 8 detectors and 8 sources is centrally placed at $z = 0 \text{ mm}$ (refer to §3.3.1). However, here the maximum source-detector separation was considered to be 41 mm. The software TOAST was used to generate CW model data at each wavelength, to which 5% random Gaussian noise was added. The background had brain optical properties, which were described in §2.2 and shown in appendix A. Hence, deoxy-haemoglobin concentration [HHb] is $24 \mu\text{M}$, oxy-haemoglobin concentration [HbO₂] is $56 \mu\text{M}$ and scattering amplitude a is $32 \text{ mm}^{-1} \text{ nm}^{-1}$. Three perturbations with dimensions $10.6 \text{ mm} \times 10.6 \text{ mm} \times 10.6 \text{ mm}$ were placed within the medium, as shown in figure 8.2. The centre of each perturbation was at 11.7 mm below the imaging surface. A single parameter was changed for each perturbation, which consisted of either an increase in [HHb] of $1 \mu\text{M}$, in [HbO₂] of $4 \mu\text{M}$ or in a of $1 \text{ mm}^{-1} \text{ nm}^{-1}$. These perturbations are kept small to guarantee that the concentrations and scattering amplitude can be retrieved by linear reconstruction. Furthermore, the cerebral haemodynamic changes are roughly of the same order of magnitude. Images were reconstructed using the multispectral method on a regular pixel basis of dimensions $24 \times 24 \times 9$.

8.4 Results and Discussion

8.4.1 Similarity measures test on photon density measurement functions

The first test consisted of comparing the reference PMDF (PMDF_r) for a homogeneous medium with $\mu_a = 0.01 \text{ mm}^{-1}$ and $\mu'_s = 1 \text{ mm}^{-1}$ and dimensions $85 \text{ mm} \times 85 \text{ mm} \times 30 \text{ mm}$, with PMDFs whose input optical properties both change, μ_a is fixed and μ'_s is varied and vice-versa (refer to table 8.1 for optical properties).

PMDF	μ_a	μ'_s	SSD	CC	RIU	H	MI	NMI	$z_{centroid}$
reference	0.01	1	-	-	-	-	-	-	11.11
1	0.02	1	0.0022	0.98	0.18	2.39	1.87	1.78	9.92
2	0.05	1	0.014	0.88	0.73	2.41	1.43	1.59	8.44
3	0.01	1.5	0.00082	0.99	0.060	2.44	2.03	1.83	10.83
4	0.05	1.5	0.017	0.84	0.83	2.41	1.24	1.51	8.27

Table 8.1: Optical properties (in mm^{-1}) used to generate PMDFs which were compared against the reference PMDF_r to find the sum of squared differences (SSD), Correlation Coefficient (CC), the Ratio-Image Uniformity (RIU), the joint entropy (H), the Mutual Information (MI), the Normalised Mutual Information (NMI) and the depth centroid $z_{centroid}$.

Figure 8.3 shows five absorption PMDFs sampled in the xz plane at $y=42.5$ mm, which illustrates the effects of using different optical properties to generate the sensitivity maps. The absolute differences between the PMDF_r and the other four others were calculated; $|\text{PMDF}_r(i) - \text{PMDF}_x(i)|$, for i pixels. The results are shown in figure 8.4, which provides a visual indicator of the similarity between the sensitivity maps.

The results for the different similarity measures are summarised in table 8.1. The smallest SSD, RIU, $z_{centroid}$ difference, and largest CC, MI, NMI occurred when the PMDF_r was compared with the PMDF₃, because a change in the scattering coefficient of the homogeneous media has little influence on the absorption PMDF, as seen in figure 8.3(d). This is also confirmed by the small absolute difference between these PMDFs, which is displayed in figure 8.4(c). Changes in the homogeneous media absorption coefficient results in visible changes in the corresponding PMDFs, as seen in Figures 8.3(b) – (c) for PMDF₁ and PMDF₂ and the corresponding absolute differences in figure 8.4(a) – (b). Changing both absorption and scattering, PMDF₄, resulted in the largest difference between PMDFs (refer to Figures 8.3(e), 8.4(d)). All the similarity measures (refer to table 8.1) reflect this tendency, except the joint entropy H test.

The results are logical, hence the similarity measures, except H, seem to be adequate for the comparison in question. However, we still need to test if these measures are independent of the volume occupied by sensitivities above 1%. The second test compares PMDFs with similar sensitivity volume, for both large and small volumes. For the large sensitivity volume we have chosen $\mu_a=0.01$ mm^{-1} and varied μ'_s . The selected values are $\mu'_s=1$ mm^{-1} for PMDF_{L1} and $\mu'_s=1.5$ mm^{-1} for PMDF_{L2}, which is the same as PMDF_r and PMDF₃. The small sensitivity volume was attained using $\mu_a=0.05$ mm^{-1} and varying μ'_s as previously for PMDF_{S1} and PMDF_{S2}, respectively. The results are displayed in table 8.2 for comparison.

The similarity tests H and MI fail. This is likely to be due to the strong dependence on the low sensitivity regions. For example, the joint histogram calculated from the PMDFs with smaller volume has more low intensity bins occupied than the joint histogram from the PMDFs with larger volume. Consequently the joint probability will increase, reducing the joint entropy H. For the MI test, the marginal entropies are lower for the PMDFs with smaller volume, therefore the MI value is higher.

The similarity measures SSD, CC, RIU, and NMI return similar values for this second test, so we conclude that they are appropriate to compare sensitivity profiles from homogeneous media with different optical properties irrespective of PMDF volume. We chose to use SSD

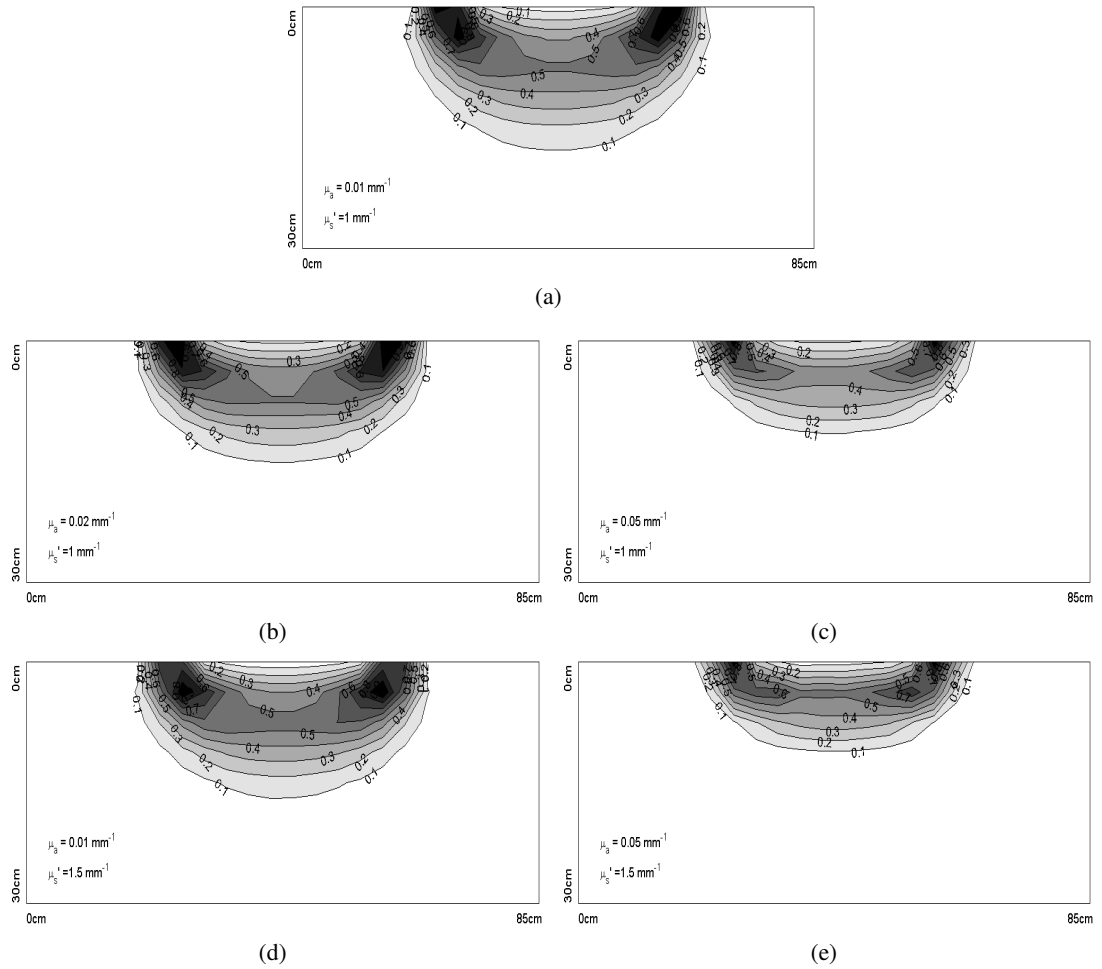


Figure 8.3: 2D photon density measurement functions (PMDFs) calculated for different optical properties in the plane xz at $y=42.5$ mm. PMDFs were compared against the (a) reference PMDF (PMDF_r) with $\mu_a = 0.01 \text{ mm}^{-1}$ and $\mu'_s = 1 \text{ mm}^{-1}$ to test the similarity measures. The contour numbers indicate the sensitivity level. (b) PMDF₁, represents an increase in μ_a from the reference value. (c) PMDF₂, represents an increase in μ_a . (d) PMDF₃, represents an increase in μ'_s . (e) PMDF₄, represents an increase in μ_a and μ'_s .

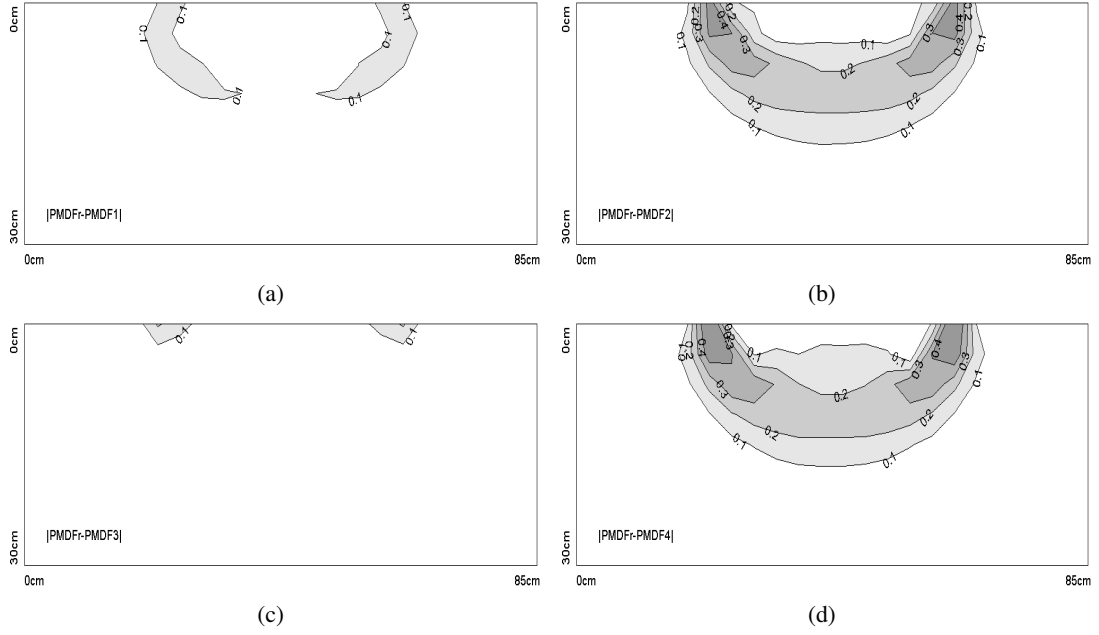


Figure 8.4: 2D absolute differences between PMDF_r and each PMDF. The numbers indicate the sensitivity level. (a) $|\text{PMDF}_r - \text{PMDF}_1|$. (b) $|\text{PMDF}_r - \text{PMDF}_2|$. (c) $|\text{PMDF}_r - \text{PMDF}_3|$. (d) $|\text{PMDF}_r - \text{PMDF}_4|$.

PMDF	μ_a	μ'_s	SSD	CC	RIU	H	MI	NMI	$z_{centroid}$
L1	0.01	1	0.00082	0.99	0.060	2.44	2.03	1.83	11.11
L2	0.01	1.5							10.83
S1	0.05	1	0.00089	0.99	0.059	1.54	1.3	1.84	8.44
S2	0.05	1.5							8.27

Table 8.2: PMDFs with large and small sensitivity volumes, PMDF_L and PMDF_S respectively, were compared to find the sum of squared differences (SSD), Correlation Coefficient (CC), the Ratio-Image Uniformity (RIU), the joint entropy (H), the Mutual Information (MI), the Normalised Mutual Information (NMI) and the depth centroid $z_{centroid}$. The optical properties are represented in units of mm^{-1} .

in the following study because it is intuitive, simple and from the beginning we expected this similarity measure to be suitable for the task.

8.4.2 Optimal wavelengths

For simplicity, the absorption and diffusion sensitivity profiles are dealt with separately. The first results are shown in figure 8.5 for the homogeneous model, considering absorption sensitivity profiles only and 3 wavelength sets. Note that the SSD values in this figure represent the average for each wavelength set. Due to the large number of points the results were limited to $R \geq 0.2$, $\mathbf{k} \leq 100$ and because the SSD for two similar PMDFs is ~ 0.001 the limit for the third criteria was set to $\text{SSD} \leq 0.002$ (dark circles in figure 8.5). Each criterion has a different order of magnitude, hence the results contained in this new range were normalised. We also calculated the distance between the criteria values of a certain set of wavelengths in this range and the point $(\text{SSD}, \mathbf{k}, R) = (0, 0, 1)$, which is the point with the smallest SSD, smallest condition number criteria and highest residual. The wavelength distribution of the ten sets of 3 wavelengths with the shortest distance are displayed in the histogram in figure

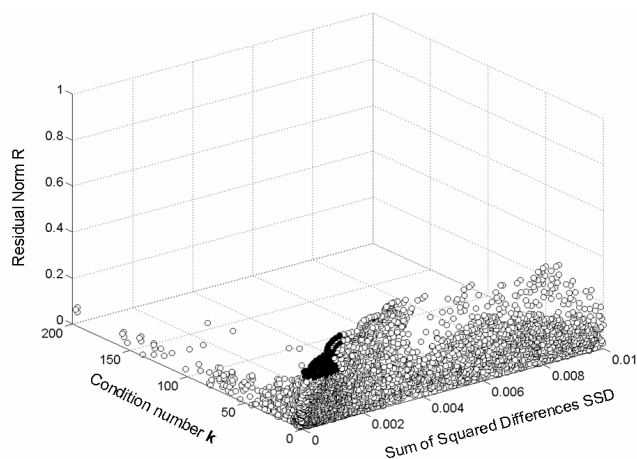


Figure 8.5: Plot of the sum of squared differences SSD, condition number \mathbf{k} and residual norm R , where each point represents a set of three wavelengths. The dark circles represent the wavelength sets which fall within the range $SSD \leq 0.002$, $\mathbf{k} \leq 100$ and $R \geq 0.2$

8.6. Histograms are obtained by counting the number of times a certain wavelength occurs, which is done separately for the 3 wavelengths, and normalised to the most frequently occurring wavelength. For comparison, figure 8.7 shows the histogram for the ten wavelength sets within the limit $R \geq 0.2$ and $\mathbf{k} \leq 100$, whose distance is closest to the point $(\mathbf{k}, R) = (0, 1)$. The means and standard deviations σ of each wavelength distribution were calculated in order to find the optimal wavelengths and measure their spread. The values are rounded up and a 5 nm uncertainty error is combined with σ , due to the bin size. Using the three wavelength selection criteria the wavelengths are 680 ± 5 nm, 726 ± 7 nm and 835 ± 17 nm. When only the high residual and low condition number criteria are used the optimal wavelengths are 650 ± 5 nm, 724 ± 7 nm and 915 ± 17 nm. The results suggest that when the new PMDF overlap condition is included the smallest wavelength becomes larger, its mean value shifts from 650 nm to 680 nm, and the largest wavelength becomes smaller, shifting from 915 nm to 835 nm.

When four wavelengths are used the histogram in figure 8.8 is obtained. The optimal wavelengths which satisfy the three criteria are 680 ± 5 nm, 715 ± 14 nm, 733 ± 7 nm and 828 ± 9 nm. Adding a fourth wavelength broadened the central wavelength range and shifted the highest wavelength to a smaller value. The histogram of wavelength distributions without including the third condition can be seen in figure 8.9, which shows that the optimum set is 650 ± 5 nm, 715 ± 7 nm, 727 ± 7 nm and 919 ± 13 nm. The differences between these results and those presented by Corlu *et al* [2005] are inevitably due to differences in the wavelength range and spacing, chromophore extinction coefficients and criteria limits imposed. As for the previous results for three wavelengths, when the similarity between sensitivity profiles is taken into account, the mean value for the smallest wavelength shifts from 650 nm to 680 nm and the highest wavelength from 919 nm to 828 nm. The central wavelength range shows some broadening. Therefore, including the SSD between PMDFs as a wavelength selection criterion reduces the separation between the optimal wavelengths.

Defining empirically the conditions for small SSD, small condition number and high residual, returns similar wavelength distributions to those found using the shortest distance

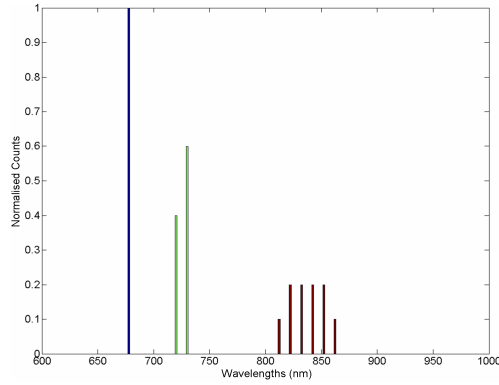


Figure 8.6: Histogram of the optimal wavelength distributions that satisfy the three wavelength selection conditions, for the homogeneous model, three wavelengths and two chromophores (HbO_2 and HHb).

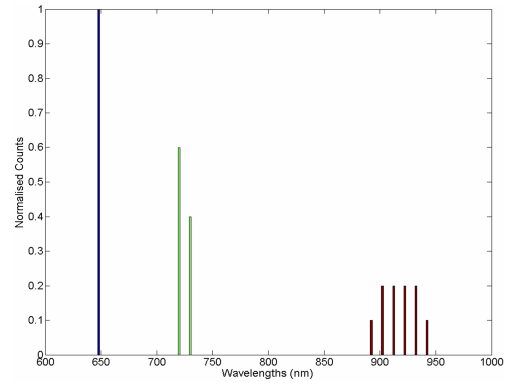


Figure 8.7: Histogram of the optimal wavelength distributions that satisfy the high residual and low condition number criteria, for the homogeneous model, three wavelengths and two chromophores (HbO_2 and HHb).

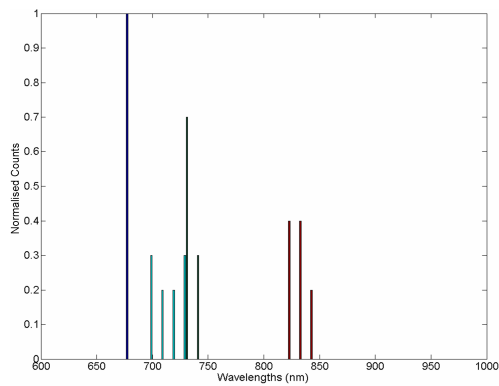


Figure 8.8: Histogram of the optimal wavelength distributions that satisfy the three wavelength selection conditions, for four wavelengths and two chromophores (HbO_2 and HHb).

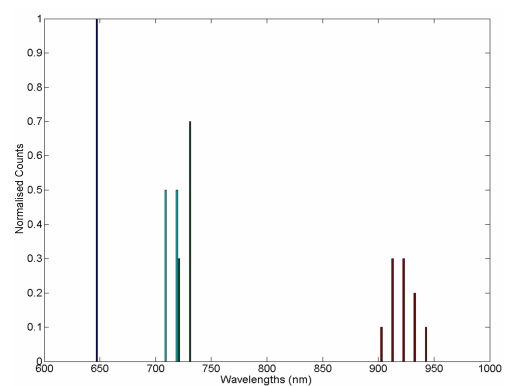


Figure 8.9: Histogram of the optimal wavelength distributions that satisfy the high residual and low condition number criteria, for four wavelengths and two chromophores (HbO_2 and HHb).

method. For example, for 3 wavelengths we defined the condition $R \geq 0.25$, $\mathbf{k} \leq 5$ and $\text{SSD} \leq 0.001$ and the optimal wavelengths are 680 ± 5 nm, 727 ± 8 nm and 836 ± 15 nm, which are very similar to the wavelengths found previously. Therefore, the shortest distance method proves to be adequate to select optimal wavelengths.

The sum of squared differences SSD for diffusion PMDFs is always slightly smaller than for absorption PMDFs. This is probably because scattering does not have a strong wavelength dependence. Thus the optimal wavelength sets found using the absorption SSDs also return small diffusion SSDs.

The histograms for the three layer adult head model for three and four wavelengths are shown in figure 8.10 and figure 8.11, respectively. For the three wavelength distributions, the optimal wavelength values are 680 ± 5 nm, 725 ± 10 nm and 876 ± 12 nm, and for four wavelengths these values are 685 ± 7 nm, 719 ± 9 nm, 731 ± 8 nm and 873 ± 9 nm. The main difference between the optimal wavelengths found for the homogeneous model and the three layer model occurs for the largest wavelength, which becomes larger for the latter. The other wavelengths are approximately the same, and similarly to the homogeneous case adding

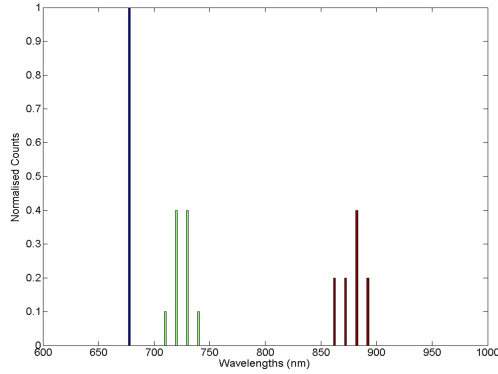


Figure 8.10: Histogram of the optimal wavelength distributions that satisfy the three wavelength selection conditions, for the three layer model, three wavelengths and two chromophores (HbO₂ and HHb).

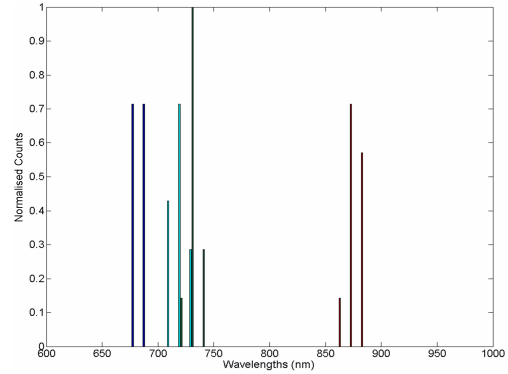


Figure 8.11: Histogram of the optimal wavelength distributions that satisfy the three wavelength selection conditions, for the three layer model, four wavelengths and two chromophores (HbO₂ and HHb).

a fourth wavelength resulted in the merge of the two central wavelengths, hence the mean value is approximately the same but the range is larger. Also for this model the diffusion SSDs are smaller than the corresponding values for absorption, hence the previous results also apply for diffusion sensitivity profiles.

Including the sensitivity profiles similarity criteria in the selection of the optimal wavelengths has a clear effect on the resultant wavelength sets. The insertion of two superficial layers in the adult head model had an influence on the results, in particular for the largest wavelength mean value, which shows that selection depends on the accuracy of the head model.

2 wavelengths. It is common to consider scattering homogeneous and constant over time. This assumption is convenient when data were obtained using only two wavelengths, and therefore solutions can only be obtained for two chromophores. The following analysis is aimed at identifying the two optimal wavelengths which give good separation between chromophores and good spatial overlap between sensitivity profiles.

First, it is shown that the solution is non-unique when only two wavelengths are used. Figure 8.12 shows a 3D plot where each point represents a wavelength pair and the respective R , \mathbf{k} and SSD values. All the points cluster at small SSD, \mathbf{k} and R and it becomes difficult to define a wavelength selection range. The red points fall in the range defined by the condition 1, $SSD \leq 0.005$, $\mathbf{k} \leq 8$ and $R \geq 1.5 \times 10^{-15}$. However, although the sum of squared differences SSD and the condition numbers \mathbf{k} can be very small, the residual norm R is never larger than 1×10^{-7} . The wavelength pairs that fall in this range are very distinct from each other, which can be seen in the histogram shown in figure 8.13. Histograms only return useful results if the points cluster around a group of wavelengths, and so, in order to show this, we display the optimal wavelength pairs for 6 different conditions in figure 8.14. It is clear that for the previous condition and condition 2, $k \leq 14$ and $R \geq 4 \times 10^{-15}$, the points do not cluster. Nevertheless, for the latter condition the corresponding histogram would have peaks at 800 nm and 960 nm. The mean and standard deviation σ of each wavelength distribution were calculated in order to find the optimal wavelengths and measure their spread. The wavelength ranges found are $835 \pm$

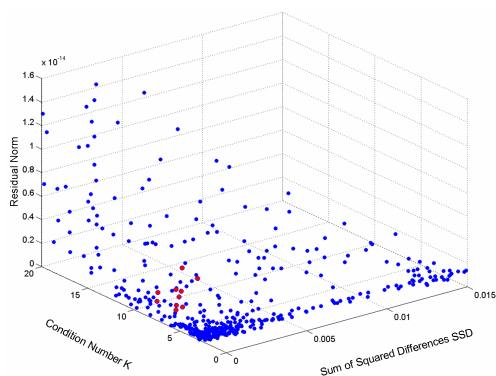


Figure 8.12: Sum of squared differences SSD, condition number k and residual norm R for different wavelength pairs. The red points fall in the range defined by the condition $SSD \leq 0.005$, $k \leq 8$ and $R \geq 1.5 \times 10^{-15}$.

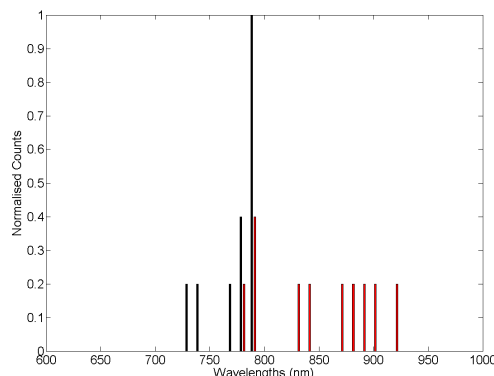


Figure 8.13: Histogram for the wavelength pairs obtained for condition $SSD \leq 0.005$, $k \leq 8$ and $R \geq 1.5 \times 10^{-15}$.

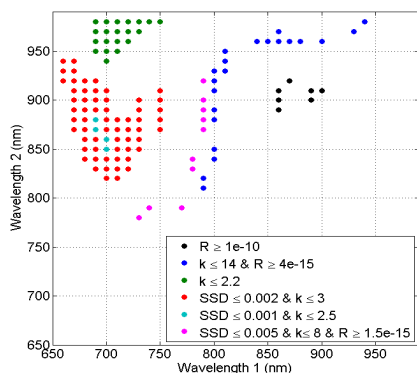


Figure 8.14: Wavelengths pairs for 6 conditions.

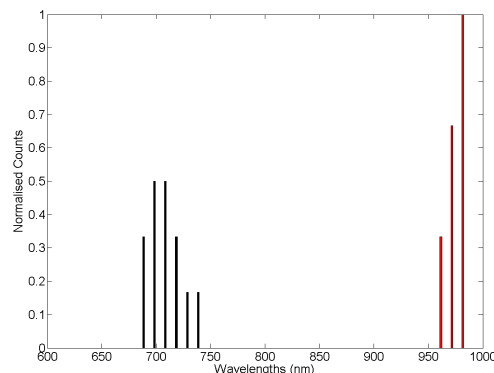


Figure 8.15: Histogram of the wavelengths that satisfy the low condition number criteria.

47 nm and 921 ± 53 nm. The large standard deviation indicates that because the wavelengths do not cluster around a particular wavelength one cannot find an optimal wavelength pair, hence σ is an important indicator of the amount of clustering. Similarly, for condition 1 the wavelength ranges are 775 ± 23 nm and 849 ± 50 nm, where σ is particularly large for long wavelengths. The wavelengths with higher residual norm, using condition 3, $R \geq 1 \times 10^{-7}$, fall in the range 878 ± 18 nm and 906 ± 11 nm. Nevertheless, the small R values confirm that one cannot separate absorption from scattering using only two measurement wavelengths.

Chromophore concentrations can be distinguished if the condition number k is small. The wavelengths whose distance is shortest to the point with the smallest condition number cluster at 710 ± 16 nm and 973 ± 9 nm (see figure 8.15). On the other hand if we set the condition $SSD \leq 0.002$ and $k \leq 3$ and find the wavelengths with shortest distance to the point $(SSD, k) = (0, 0)$, they cluster at 694 ± 7 nm and 883 ± 23 nm (see figure 8.16). A 2D plot of SSD versus k is seen in figure 8.17 where the red points satisfy the condition $SSD \leq 0.002$ and $k \leq 3$. The inclusion of the PMDF overlap condition has a clear effect on the largest wavelength.

For the three layer model, the wavelength pair with the shortest distance to the point $(SSD, k) = (0, 0)$ is 704 ± 7 nm and 887 ± 12 nm, which is very similar to the optimal wavelength set found for the homogeneous case.

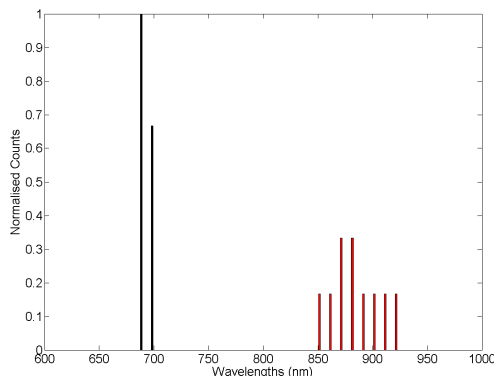


Figure 8.16: Histogram for wavelengths with the shortest distance to the point $(SSD, k) = (0, 0)$.

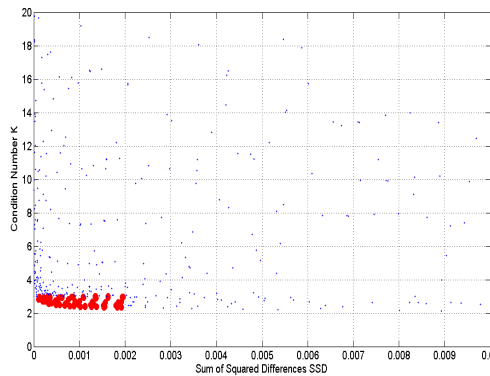


Figure 8.17: Sum of squared differences SSD and condition number k for different wavelength pairs. The red points represent the wavelength pairs that have $SSD \leq 0.002$ and $k \leq 3$.

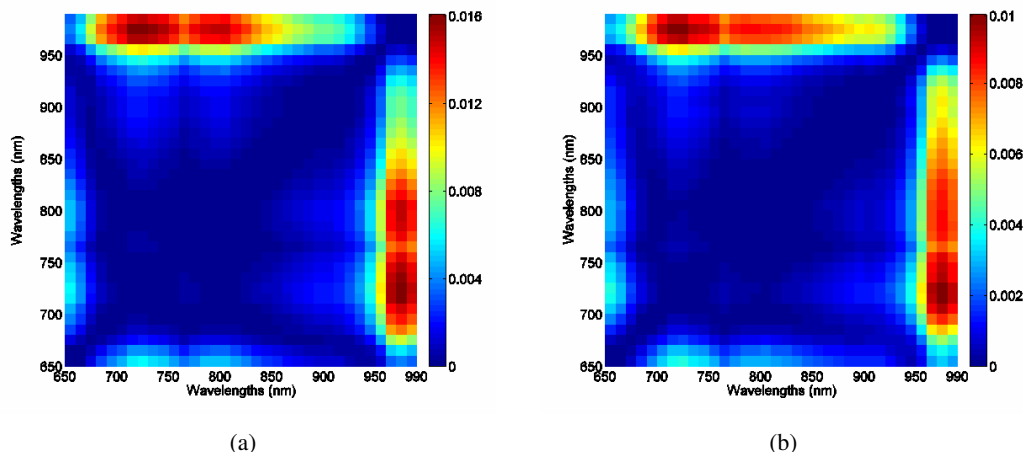


Figure 8.18: PMDFs Sum of squared differences SSD for wavelength pairs in the range 650-990nm for (a) absorption and (b) diffusion.

The influence of μ_a on the absorption and diffusion PMDFs is identical; high absorption coefficients lead to narrower PMDFs. A similar tendency is observed when μ'_s has a high value for the absorption PMDFs, although the diffusion PMDFs become broader. Nevertheless, the SSD for both absorption and diffusion PMDFs is similar, as can be seen in figure 8.18. The wavelength sets obtained using the diffusion PMDFs are identical to those found using the absorption PMDFs. The largest SSD occurs for wavelengths paired with 970 nm, which corresponds to the absorption peak of water. As mentioned previously, the sum of squared differences SSD for diffusion is slightly smaller than for absorption.

In summary, it is possible to separate $[HbO_2]$ and $[HHb]$ from measurements performed at two wavelengths. However, one cannot distinguish between absorption and scattering. When the sensitivity profiles similarity condition is included in the wavelength selection method, the largest wavelength mean value changes in comparison to that obtained using the condition k alone.

8.4.3 PMDF comparison

In the previous study, our main concern was to ensure that the sensitivities within the volumes of tissue interrogated were similar for a set of wavelengths. Therefore, the PMDFs were normalised and all the values below a certain threshold were set to zero. Here, the aim is to compare the PMDFs for wavelengths from 650 nm to 990 nm and quantify the differences in their sensitivity in terms of penetration depth, volume and sensitivity, in order to determine if these differences are significant. We use the absolute sensitivity values (because the absorption PMDF has negative sensitivity) and set the threshold to 0.01. An analysis is performed on how these differences affect the measurements and image reconstructions. In order to understand the effect that the optical properties have on the sensitivity profiles, the PMDF profiles (plane xz at $y = 42.5$ mm) for 4 different wavelengths (650 nm, 720 nm, 850 nm and 970 nm) are shown in figure 8.19, for the homogeneous medium, and in figure 8.20, for the three layer model.

8.4.3.1 Volume

The volume (V) is the number of voxels (N) with sensitivity above the threshold times the voxel volume. The relative volume compares the PMDF volume at 650 nm (V_{650}) with all the other PMDF volumes (V_λ):

$$relative\ volume = \frac{V_\lambda}{V_{650}} = \frac{N_\lambda}{N_{650}}, \quad (8.16)$$

where N_{650} is the number of voxels above the threshold in the PMDF at 650 nm and N_λ is the number of voxels above the threshold for the remaining wavelengths.

The relative volume was calculated for the absorption PMDFs (PMDF $_{\mu_a}$) and diffusion PMDFs (PMDF $_{\kappa}$), for both head models. Figure 8.21 shows the relative volumes for the different wavelengths and the value of N_{650} for each case. The volume of the PMDF $_{\mu_a}$ is larger than that of the PMDF $_{\kappa}$. For the homogeneous model, the PMDF $_{\mu_a}$ volume resembles an inverted μ_a spectrum of brain tissue, since for large μ_a the PMDF volume becomes smaller and for small μ_a the volume becomes larger. For the PMDF $_{\kappa}$ the volume decreases as expected, given that the volume decreases with decreasing μ'_s . However, the volume for PMDF $_{\kappa}$ does not vary as much as for PMDF $_{\mu_a}$. The relative volumes for the layered model are similar to those obtained for the homogeneous case.

8.4.3.2 Full Width Half Maximum

The FWHM was calculated for the PMDF cross-section at $x=42.5$ mm (half way between the source and detector), for both lateral and depth directions. The PMDF profiles for each direction include the point of maximum sensitivity. The FWHM gives us information about the amount of spread of the PMDFs. Figure 8.22 shows the difference between the depth FWHMs and the depth FWHM corresponding to the wavelength 650 nm, FWHM $_{650}$. The red lines indicate the error limits, which represent an uncertainty of 0.6 mm associated with the pixel size. In general, the difference is less than 1.2 mm, which indicates that the spread of regions of higher sensitivity does not vary much with wavelength. The FWHM differences are smaller for PMDF $_{\kappa}$, in particular for the homogeneous model, which shows a relatively flat spectrum for wavelengths below 970 nm. Figure 8.23 shows the difference between the lateral FWHMs

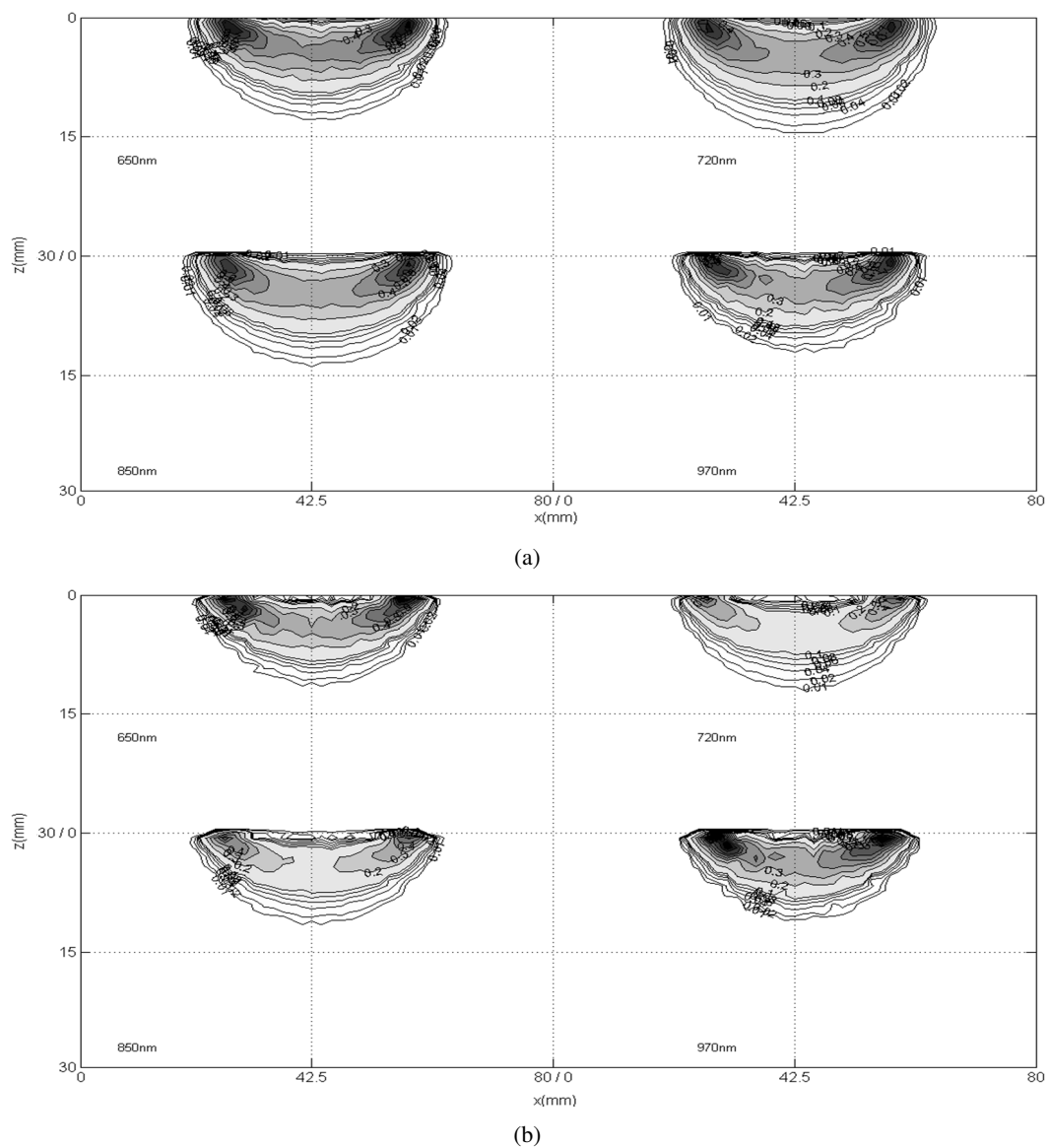


Figure 8.19: PMDF profiles for 4 different wavelengths, for the homogeneous model. (a) Absorption PMDF at 650 nm (top left), 720 nm (top right), 850 nm (bottom left) and 970 nm (bottom right). (b) Diffusion PMDF for the same wavelengths.

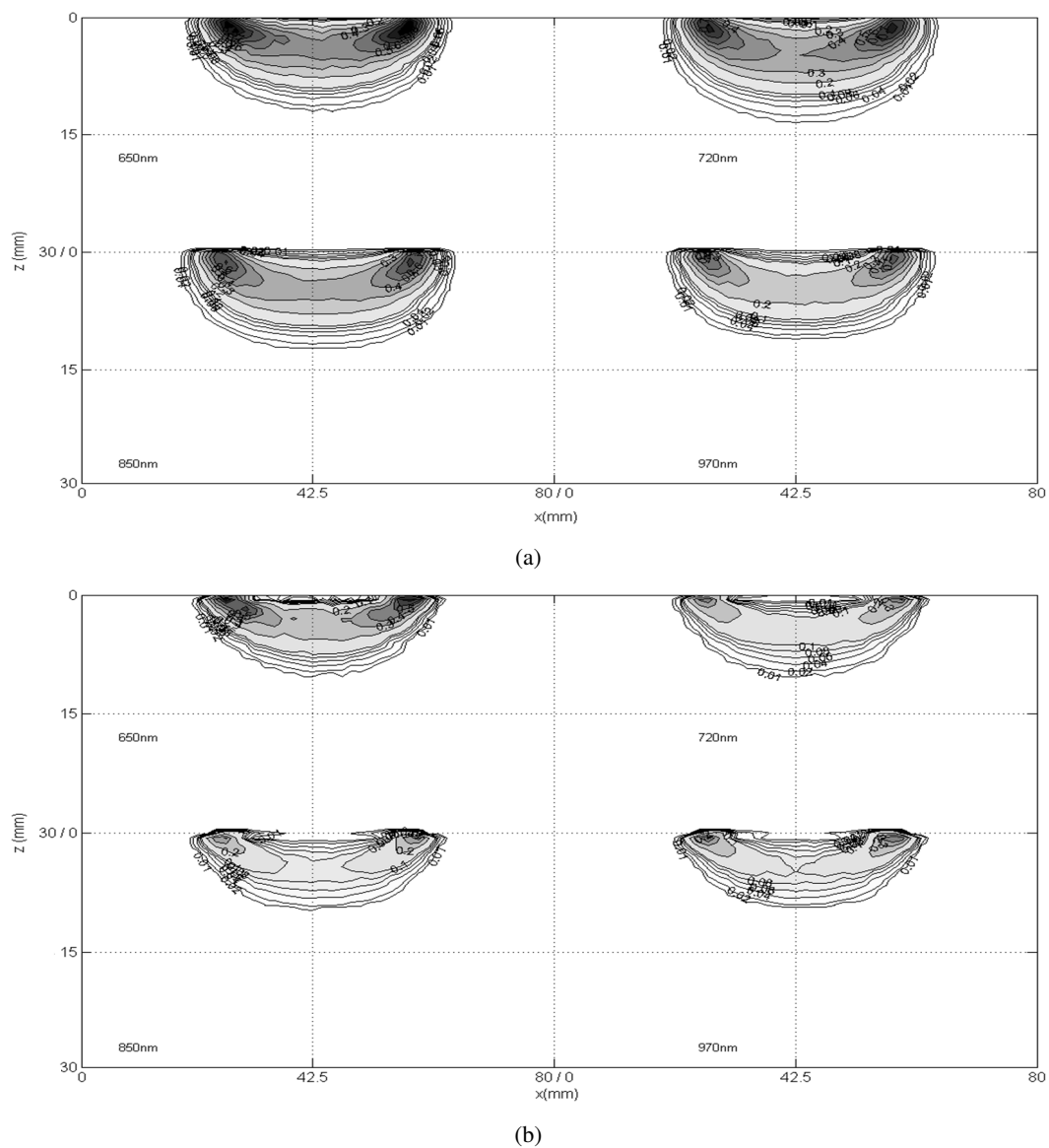


Figure 8.20: PMDF profiles for 4 different wavelengths, for the 3 layer model. (a) Absorption PMDF at 650 nm (top left), 720 nm (top right), 850 nm (bottom left) and 970 nm (bottom right). (b) Diffusion PMDF for the same wavelengths.

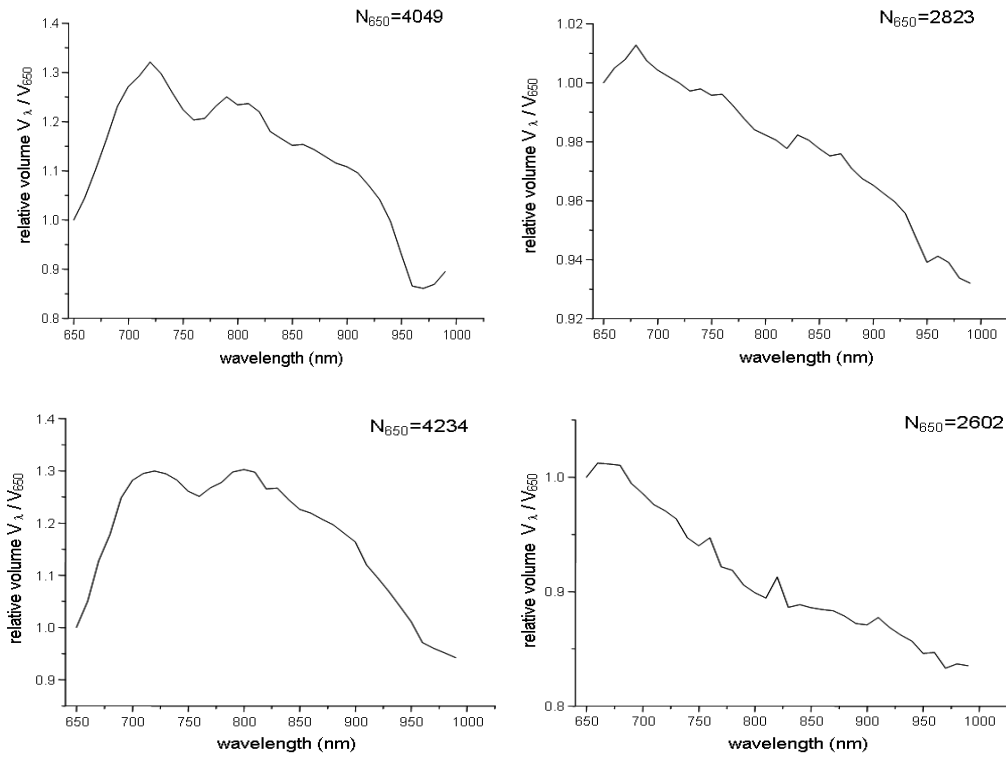


Figure 8.21: Relative volume of the PMDF $_{\mu_a}$ (left) and PMDF $_{\kappa}$ (right) for the homogeneous head model (top) and layered model (bottom).

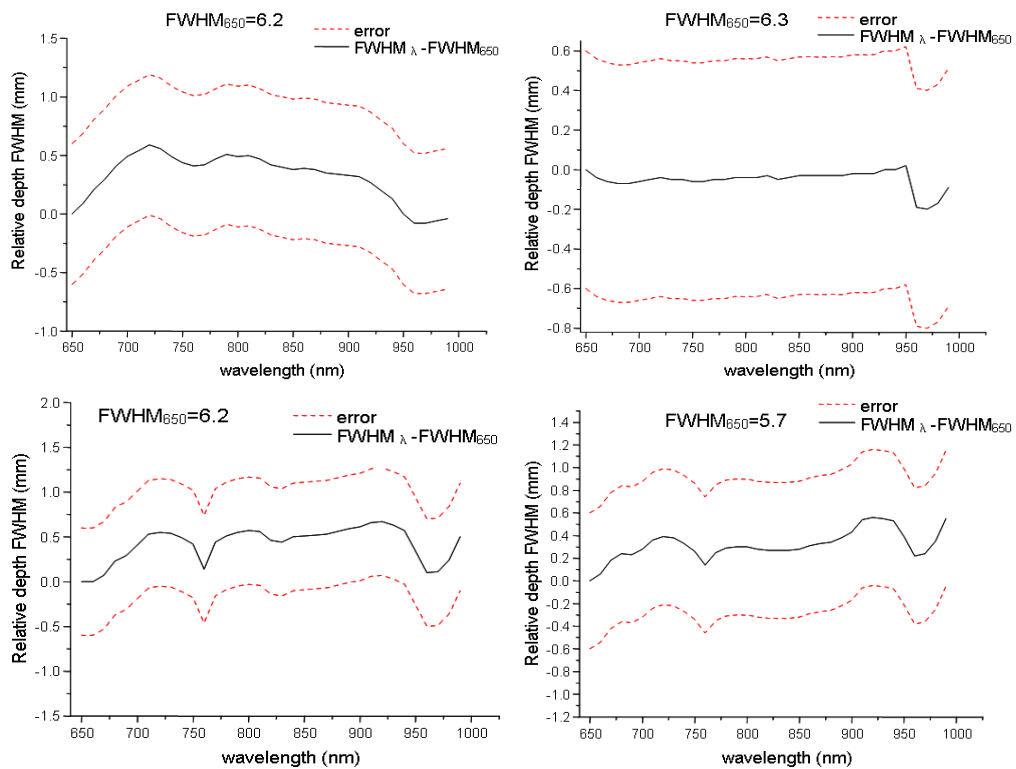


Figure 8.22: FWHM for the depth direction of the PMDF $_{\mu_a}$ (left) and PMDF $_{\kappa}$ (right) profiles for the homogeneous head model (top) and layered model (bottom).

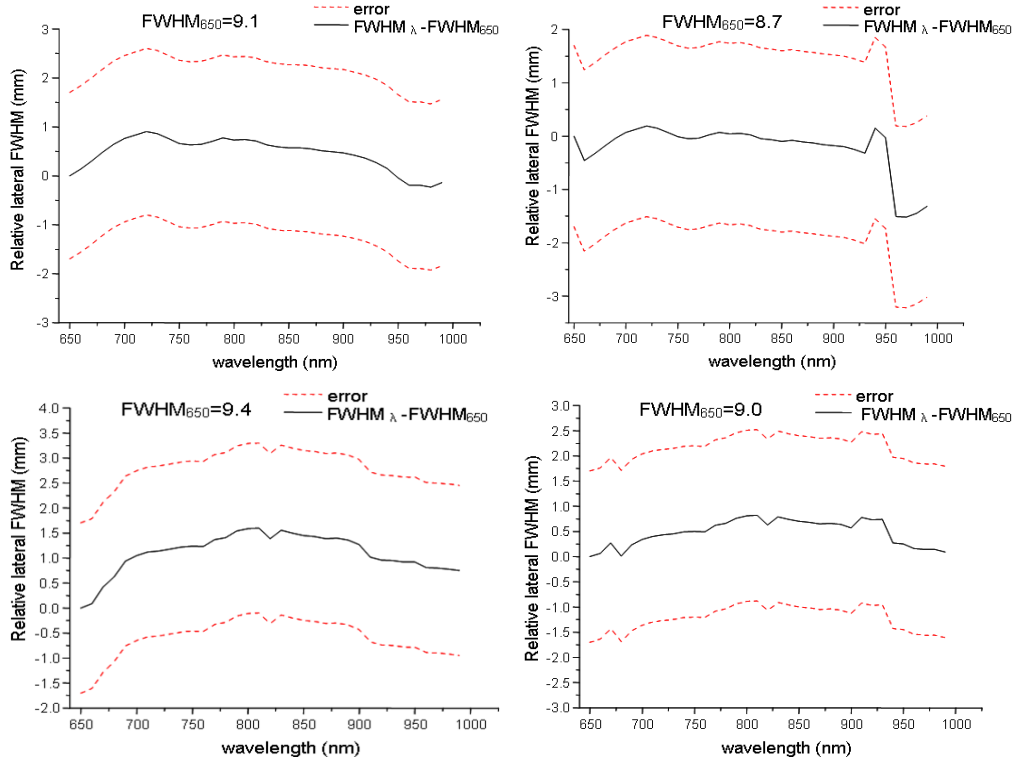


Figure 8.23: Lateral FWHM of the PMDF $_{\mu_a}$ (left) and PMDF $_{\kappa}$ (right) profiles for the homogeneous head model (top) and layered model (bottom).

and the lateral FWHM corresponding to the wavelength 650nm. The dispersion of the PMDFs is larger laterally than for the depth direction, which is confirmed by the larger FWHM values. Differences in the lateral FWHM are less than 3.5 mm. The uncertainty in the lateral direction is 1.7 mm, which explains the slightly larger difference.

8.4.3.3 Penetration depth

The penetration depth (PD) is defined as the maximum distance into the medium where perturbations in the optical properties can still be measured. The central region of the PMDF penetrates deeper into the tissue, which corresponds to the PMDF profile (z direction) that passes through the midpoint between the source and detector ($(x, y)=(42.5 \text{ mm}, 42.5 \text{ mm})$). Hence, for these profiles, we consider the distance where the sensitivity drops below the threshold to be our penetration depth.

Penetration depth is larger for the PMDF $_{\mu_a}$, as shown in figure 8.24. The graphs are plotted as differences with reference to the PMDF penetration depth at 650 nm, PD_{650} . For the homogeneous model, the maximum depth is $15.6 \pm 0.6 \text{ mm}$ and occurs for wavelength 720 nm. The lowest PD is $13.2 \pm 0.6 \text{ mm}$ and occurs for wavelength 970 nm. This means that the maximum PD difference is $2.4 \pm 1.2 \text{ mm}$. For the layered model, the maximum penetration depth is $14.4 \pm 0.6 \text{ mm}$ for wavelength 720 nm. For wavelength 970 nm the PD attains a minimum value of $12.0 \pm 0.6 \text{ mm}$. The largest difference between penetration depths is $2.4 \pm 1.2 \text{ mm}$.

The PMDF $_{\kappa}$ exhibits a small variation in the penetration depth for the different

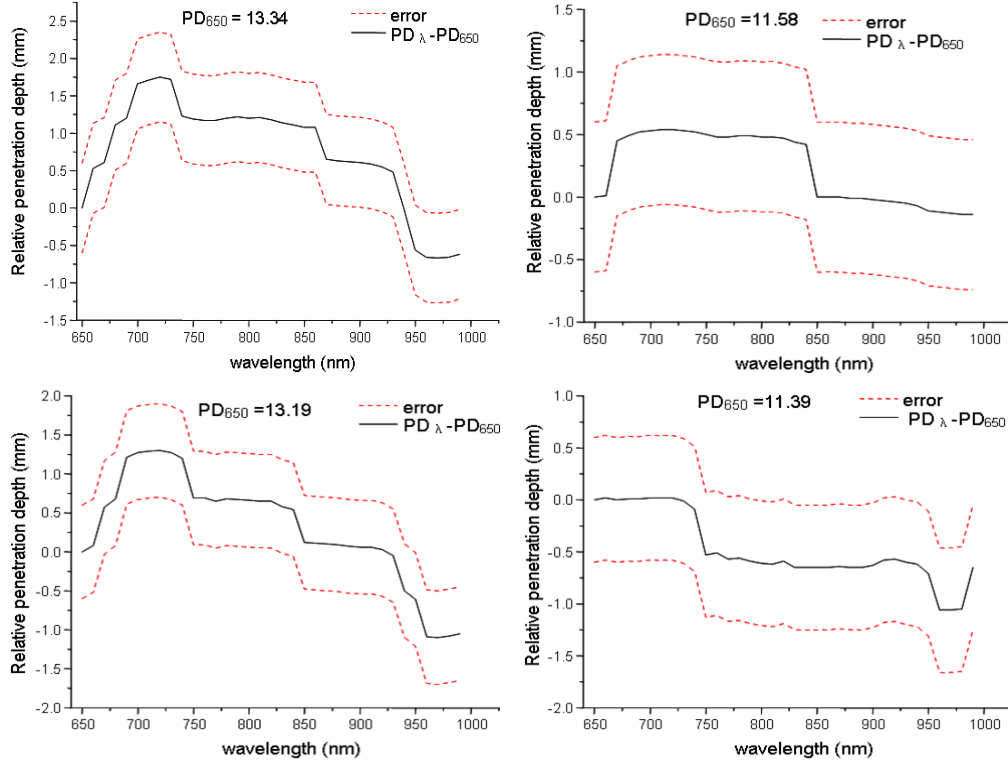


Figure 8.24: Penetration depth of the PMDF $_{\mu_a}$ (left) and PMDF $_{\kappa}$ (right) profiles for the homogeneous head model (top) and layered model (bottom).

wavelengths (figure 8.24). For the homogeneous model, the combination of low μ_a and high μ'_s for wavelengths from 670 nm to 840 nm, results in a broader PMDF $_{\kappa}$ with higher penetration depth. The maximum PD is 12.6 ± 0.6 mm and the largest difference between PD is 0.6 ± 1.2 mm. The μ'_s for skull decays faster than for brain tissue, and at wavelengths above 750 nm their μ'_s values diverge significantly. When the external layers have lower μ'_s and μ_a than the inner layer, the sensitivity profiles shift towards the surface. The high absorption of the brain tissue at 970 nm reduces the interrogation depth to 9 ± 0.6 mm, compared to the maximum at 11.4 ± 0.6 mm.

8.4.3.4 Sensitivity depth

The FWHM was used as an indicator of the spread of the regions with high sensitivity. Another two parameters, which can give similar information, are the mean sensitivity and maximum sensitivity depths. The mean sensitivity depth is calculated from the depth sensitivity profile using the following equation

$$z_m = \frac{\sum_i^n PMDF_i z_i}{\sum_i^n PMDF_i}, \quad (8.17)$$

where z_i is the depth of pixel i . The maximum sensitivity depth z_{max} is simply the depth at which the maximum sensitivity occurs in the PMDF profile. From figure 8.25 we see that for the homogeneous model, the maximum z_m occurs at 720 nm, where the absorption coefficient is lower. The graphs are plotted as differences with respect to the values at 650 nm. In general,

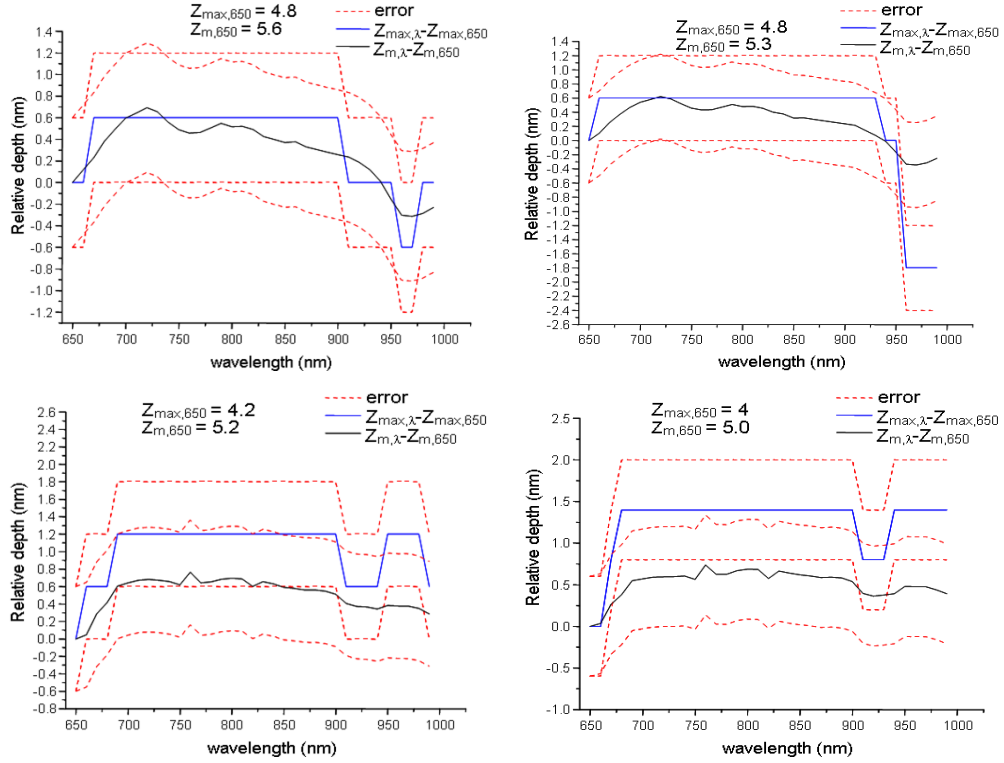


Figure 8.25: Maximum and mean sensitivity depths of the PMDF $_{\mu_a}$ (left) and PMDF $_{\kappa}$ (right) profiles for the homogeneous head model (top) and layered model (bottom).

the maximum and mean depths increase when μ_a decreases. For PMDF $_{\mu_a}$, the maximum z_{max} difference is 1.2 ± 1.2 mm, and occurs between wavelengths 720 nm and 970 nm. For PMDF $_{\kappa}$, for the same wavelengths, this value is 2.4 ± 1.2 mm. For z_m the maximum difference is 1 ± 1.2 mm for both PMDFs.

The inclusion of the external layers in the head model has a clear effect on the sensitivity depth. In particular, for z_m the values show smaller variations with wavelength. The high absorption and scattering coefficients of the brain tissue limits the sensitivity to the skull layer, in particular at 970 nm. The high μ'_s skin layer, compared to the skull layer, also confines the photons near the skin/skull boundary, and the low μ_a increases the sensitivity in this region. Hence, the region of high sensitivity is spread around the skull layer.

As a consequence of the pixel size, which only allows discrete spatial steps, the relative sensitivity depth and relative penetration depth spectra exhibit a step-like behaviour.

8.4.3.5 Maximum sensitivity

The maximum sensitivity (MS) was calculated for the PMDF profiles in the z direction, which were described for the PD analysis. Figure 8.26 shows the MS for the whole spectrum relative to the MS for the PMDF at 650 nm, MS_{650} . For the homogeneous model, the sensitivity of the PMDF $_{\mu_a}$ decreases with wavelength, which is a consequence of the decrease in μ'_s . When μ'_s increases the PMDF region with higher sensitivity is shifted towards the surface and the sensitivity to perturbations in this region increases. Similarly, but less obvious in the spectrum, the sensitivity at the surface is higher when μ_a is higher. The maximum changes in

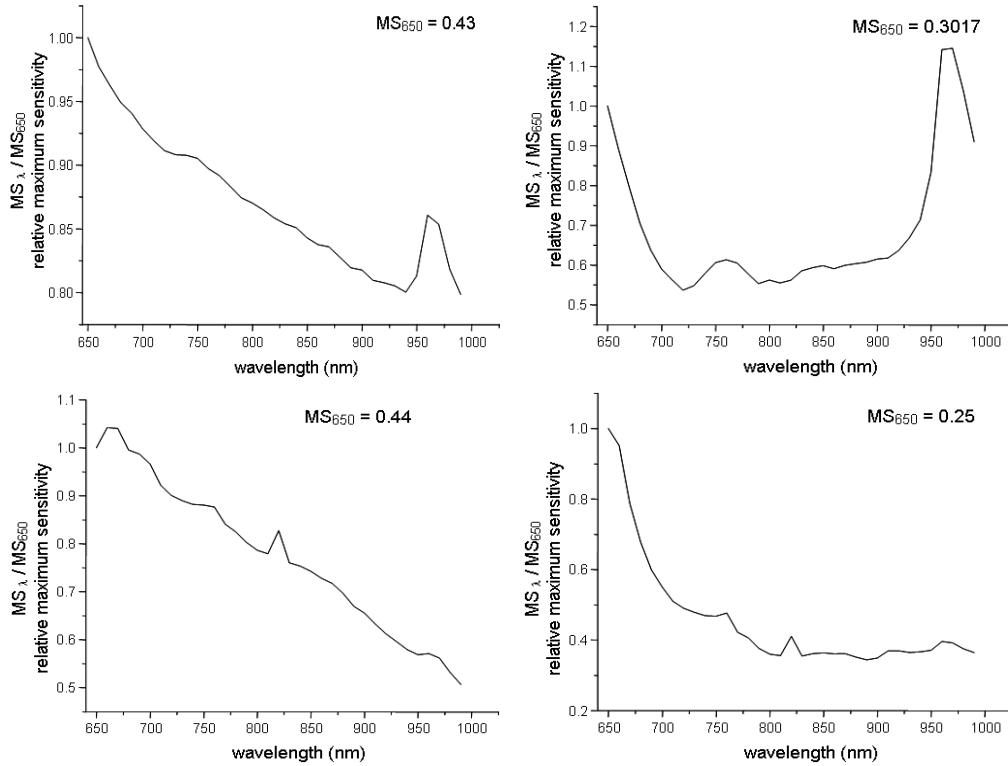


Figure 8.26: Maximum sensitivity of the PMDF $_{\mu_a}$ (left) and PMDF $_{\kappa}$ (right) profiles for the homogeneous head model (top) and layered model (bottom).

sensitivity are approximately 20% of the maximum sensitivity. The PMDF $_{\kappa}$ shows a sensitivity increase for large μ_a values, which corresponds to the extremes of the spectrum. For the central wavelengths the sensitivity is low and has small variations. The sensitivity increases when the PMDF is confined to a smaller region and closer to the surface. The sensitivity increase can be interpreted as an increase in the probability of a photon travelling through that region prior to being detected.

For the layered model, the spectrum for PMDF $_{\mu_a}$ is similar to the homogeneous case (see figure 8.26), except for the peak at 970nm, which is a result of the low μ_a and μ'_s of the external layers. The maximum changes in intensity are approximately 50% of the maximum intensity. For the PMDF $_{\kappa}$, the low μ'_s of the external layers, compared to the brain tissue, confines the PMDF to the superficial layers. The scattering coefficient decreases with wavelength, and so does the overall broadening and maximum sensitivity. The high μ_a brain tissue seems to reduce the sensitivity.

8.4.3.6 Discussion

It is important to quantify the main characteristics of the optically interrogated volumes, in order to determine how measurements at different NIR wavelengths may be combined to provide information about the volume. We compared the volume, FWHM, maximum sensitivity depth, penetration depth and maximum sensitivity of the PMDFs.

In general, for the homogeneous model, an increase in the absorption coefficient μ_a shifts the PMDF toward the surface. Meanwhile an increase in transport scattering coefficient μ'_s also

causes a narrowing in the PMDF_{μ_a} , whereas PMDF_{κ} becomes broader. For the three layer model, the higher optical properties of the brain tissue compared to the external layers largely restrains photons within the top layers and only a small amount penetrates into the brain. The PMDFs spread and broaden above the brain boundary.

Particular attention is given to the wavelengths selected using our wavelength selection method and Corlu's method [Corlu *et al.*, 2005]. For the homogeneous model our method returned the wavelengths with mean values (the values were rounded to one of the nearest wavelengths for which the PMDF was calculated) 680 nm, 720 nm and 830 nm, when three wavelengths were used, and wavelengths 680 nm, 710 nm, 730 and 830 nm, when four wavelengths were used. For the three layer model the three wavelengths found were 680 nm, 730 nm and 880 nm, and the set of four wavelengths was 680 nm, 710 nm, 730 nm and 880 nm. For the two condition criteria the wavelengths were 650 nm, 720 nm and 910 nm, for three wavelengths, and 650 nm, 710 nm, 730 nm and 910 nm, for four wavelengths.

For the optimal wavelengths selected using the high residual and low condition number criteria, for the homogeneous model, the maximum penetration depth for the PMDF_{μ_a} is 15.6 ± 0.6 mm, and occurs for the mid-range wavelengths. For the lowest wavelength, with mean value 650 nm, the penetration depth is 13.8 ± 0.6 mm. For the highest wavelength, with mean value at 910 nm, the PD is 14.4 ± 0.6 mm. The maximum PD difference is 1.8 ± 1.2 mm. If we consider the thickness of the head superficial layers to be approximately 11 mm, and the grey matter to be 4 mm, then a difference in the PD of about 3 mm could make a very significant difference in the information contained in the measurements. In theory, the lowest wavelength penetrates 2.8 ± 0.6 mm into the brain tissue, which corresponds to the grey matter, whereas the mid-range wavelengths penetrate 4.6 ± 0.6 mm, which means that some light could reach the white matter. Including our PMDF overlap condition in the selection of the optimal wavelengths reduces the difference in the PD to 0.6 ± 1.2 mm. The first and last wavelengths, with mean values 680 nm and 830 nm respectively, penetrate $4 \text{ mm} \pm 0.6 \text{ mm}$ into the brain tissue. This value is closer to that obtained for the central wavelengths, which is 4.6 ± 0.6 mm. The volumes interrogated are also larger and more similar for the wavelengths selected using our method. For the PMDF_{κ} , the changes in the PD are smaller and the maximum penetration depth is only 12.6 ± 0.6 mm, hence confined to the grey matter. However, the differences between penetrations depths are smaller when our method is used.

For the three layer adult head model the discrepancies between the interrogated volumes at different wavelengths are smaller. However, the volumes are more similar for the wavelengths selected using our method than using the two selection criteria. In terms of PD, for the PMDF_{μ_a} , the variations are smaller than for the homogeneous model. The maximum PD is 14.5 ± 0.6 mm and is attained for the central wavelengths of the selected wavelength sets. This means that light penetrates 3.5 ± 0.6 mm into the grey matter. For the first wavelengths obtained with our method the PD is 13.9 ± 0.6 mm, hence it penetrates 2.9 ± 0.6 mm into the brain. For the remaining wavelengths the PD is 13.2 ± 0.6 mm, or equivalently, a maximum depth of 2.2 ± 0.6 mm from the brain surface is probed. Although the differences between PD for the different wavelengths, for both selection methods, are identical, when using our method it optimises for the wavelength set with higher penetration depth possible, simultaneously with

the best separation between chromophores, absorption and scattering. The PMDF spread, (i.e. the FWHM) is also larger and more similar for our wavelength set. The maximum sensitivity depth z_{max} and mean sensitivity depth z_m show small variations for wavelengths from 680 nm to 900 nm, where these values are maximum. For the PMDF $_{\kappa}$, the maximum penetration depth is only 11.4 ± 0.6 mm, hence light hardly reaches the brain surface. The differences PMDF $_{\kappa}$ are much smaller than for the absorption PMDFs.

This method could be applied to other types of tissue. For example, for measurements of muscle, breast or neonate brain. For muscle, the highest absorption occurs at 650 nm and the lowest approximately at 850 nm [Xia *et al.*, 2006]. For these wavelengths the PD is 15 ± 0.6 mm and 21.6 ± 0.6 mm, respectively. This results in a PD maximum difference of 6.6 ± 1.2 mm. For studies of muscle exercise, of the forearm for example, which has a diameter of approximately 7 cm, a difference of about 8 mm is quite significant. The influence of the skin and fat layer on the light propagation is likely to increase the PD, and it should depend on the thickness of these layers. For breast tissue, using the optical properties from [Cubeddu *et al.*, 2000], the largest differences in the absorption coefficient occur at wavelengths 720 nm and 970 nm, with PD of 22 ± 0.6 mm and 16 ± 0.6 mm, respectively. Therefore, the maximum PD difference is 6 ± 1.2 mm. The main objective of breast diffuse optical imaging is tumour detection. A tumour is distinguished from surrounding healthy tissue on the basis of its optical properties, and subsequently in terms of apparent blood volume, and oxygenation. Overlapping sensitivities are required to provide accurate measures of these parameters and reduce the uncertainty in the measured spatial location of the tumour. Light propagation in the neonatal brain is different from the adult brain tissue, because the surrounding layers are thinner and the optical properties are different. The μ'_s is lower for the neonatal brain tissue [van der Zee, 1992]. Because of the high scattering of the neonatal brain, the sensitive regions of the PMDF should be confined to the grey matter. However, the cortical thickness is only about 2.5 mm, meaning that light is likely to penetrate into deeper regions of the white matter [Fukui *et al.*, 2003].

The PMDFs of other types of measurements, either frequency or time domain, can also be analysed using our method. For example, the region of higher sensitivity of the mean time or phase PMDF is located deeper in the tissue than that of the CW PMDFs [Arridge & Schweiger, 1995; Delpy & Cope, 1997].

It is clear that the adult head model had a large influence on the results. A more complete model of the adult head should include the cerebrospinal fluid (CSF) surrounding the brain tissue. However the diffusion equation-based model is only valid for high scattering tissues, and cannot be applied to the low scattering CSF. Monte Carlo simulations of light transport can be used, but they require long computational times. A popular method is the hybrid radiosity-diffusion method, where the light propagation in high scattering media is calculated using the diffusion equation and the low scattering region by radiosity theory [Arridge *et al.*, 2000; Dehghani *et al.*, 2000].

Ono *et al* [2000] studied a simple three layer model with skin/skull, CSF and brain tissue optical properties. They used the hybrid radiosity-diffusion method to model light propagation, and observed that the sensitive region shifts towards the CSF layer and spreads between the

source and detector. Light only penetrates the superficial parts of the brain.

A study by Fukui *et al* [2003] compared the influence of the CSF on the light propagation within both neonatal and adult heads. They arrived at similar conclusions, that the sensitivity profiles appear distorted around the CSF layer and only reach superficial regions of grey matter. For larger source-detector separations the sensitivity increases within the grey matter. However, for the neonate model, even though the influence of the CSF is significant, the penetration depth increases with the source-detector distance and can possibly reach the deep regions of the white matter. The scattering coefficient of the grey and white matter for the neonate are much smaller than for the adult brain, hence the CSF has a smaller effect.

Okada *et al* [1997] studied light propagation within head models with different levels of complexity. At a source-detector separation of 30 mm, for a simple model without CSF, they verified that the spatial sensitivity profile is confined to the superficial layers, and only a small fraction of detected photons actually reach the grey matter. Including CSF in the model shifts the sensitivity towards the CSF layer, which spreads into the grey matter. The inclusion of more complex structures in the model, for example, sulci filled with CSF, did not produce significant differences between the sensitivity profiles and those corresponding to the simple model with a CSF layer. In a more recent study, Okada *et al* [2003b] analysed the effect of the superficial layers' thickness on the spatial sensitivity profiles. The effect of skull or CSF thickness on the penetration into the brain tissue was found to be negligible. However, the maximum sensitivity decreases with increasing thickness. An increase in the CSF layer thickness was found to broaden the spatial sensitivity profile. This indicates that even though the thicknesses of the layers in our model were approximations, because the skull thickness varies between individuals and with position on the head, the PMDF penetration depth into the brain tissue is not likely to be highly influenced by these differences.

Another study that explored the influence of internal structures in the CSF layer was performed by Dehghani *et al* [2000]. Regions with the same optical properties as the grey matter were inserted into the CSF layer. The PMDFs calculated for this model show little difference to those calculated for a model with no intrusions in the CSF layer. However, for larger intrusions the sensitivity of the intrusion itself increases and also the regions immediately adjacent to it. Ogoshi *et al* [2005] used an even more complete model, which included the folded brain surface with sulci, and the space between the skull and the brain surface filled with CSF, called the subarachnoid space. They used a hybrid Monte Carlo-diffusion method to model the light propagation through this more realistic human head model, and found that photons tend to propagate in shallow regions of the brain. The PMDF distorts around the low scattering subarachnoid space and superficial brain tissue. The sensitivity region is broadened between the source and detector, and an increase in the source-detector separation does not significantly increase the penetration depth, but the sensitivity to the shallow region of the brain increases. They were able to detect changes in intensity caused by focal absorption changes in the white matter. The results were very similar to those found using a simpler model.

The inclusion of a CSF layer in our adult head model would not significantly affect our results, since the high sensitivity regions of the PMDF would still be confined to superficial regions of the brain, and because this layer has very little wavelength dependence. The

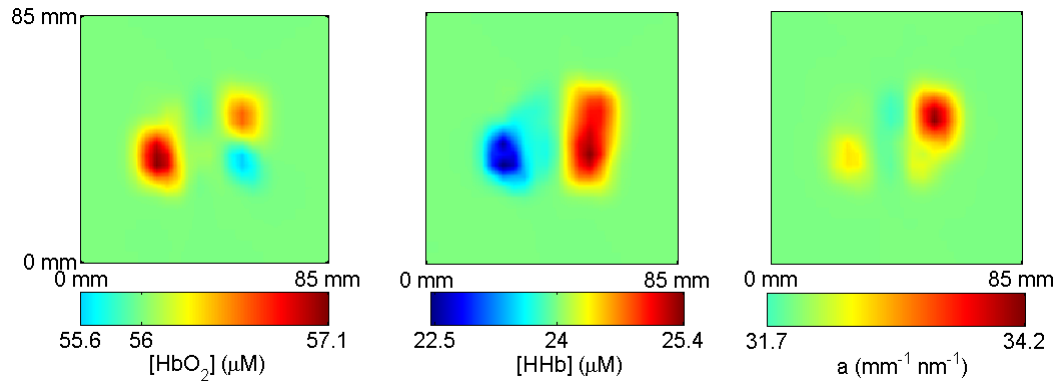


Figure 8.27: Reconstructed images of oxy-haemoglobin (left), deoxy-haemoglobin (centre) and scattering amplitude (right) from measurements at wavelengths 650 nm, 720 nm and 910 nm.

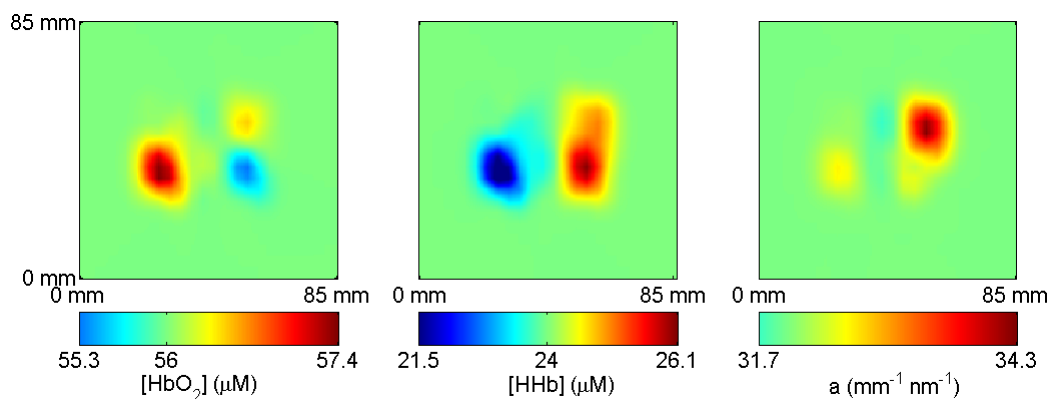


Figure 8.28: Reconstructed images of oxy-haemoglobin (left), deoxy-haemoglobin (centre) and scattering amplitude (right) from measurements at wavelengths 680 nm, 720 nm and 830 nm.

penetration depth of the light in models with and without CSF is almost the same, for source-detector separations of 30 mm and 40 mm [Okada & Delpy, 2003a].

An aspect to take into consideration is the error associated with the extinction coefficient for the different chromophores. The spectra found in the literature, for certain chromophores, show some variation. Based on Corlu's method [Corlu *et al.*, 2005] for wavelength optimisation, Brendel and Nielsen [2007] added a criterion that evaluates the reconstruction errors due to uncertainty in the extinction coefficients. From simulations and breast-like phantom measurements, they found a set of optimum wavelengths almost independent upon the uncertainties in the extinction coefficients spectra. Another potential source of error is the assumed values of the chromophore concentrations for each head layer. This influences the optical properties of the layers, which could possibly have an effect on the spatial sensitivity profiles.

8.4.4 Image reconstruction

Images were reconstructed from simulated measurements at wavelengths 680 nm, 720 nm and 830 nm (wavelengths selected using our method) and from measurements at 650 nm, 720 nm and 910 nm (wavelengths selected using the two criteria method). Figure 8.27 shows the xy cross-sectional images at $z=8.4$ mm, using measurement wavelengths at 650

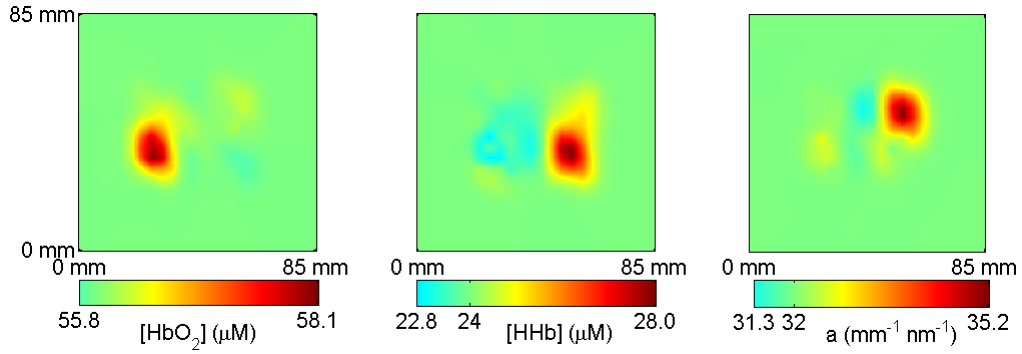


Figure 8.29: Reconstructed images of oxy-haemoglobin (left), deoxy-haemoglobin (centre) and scattering amplitude (right) from measurements at wavelengths 650 nm, 710 nm, 730 nm and 910 nm.

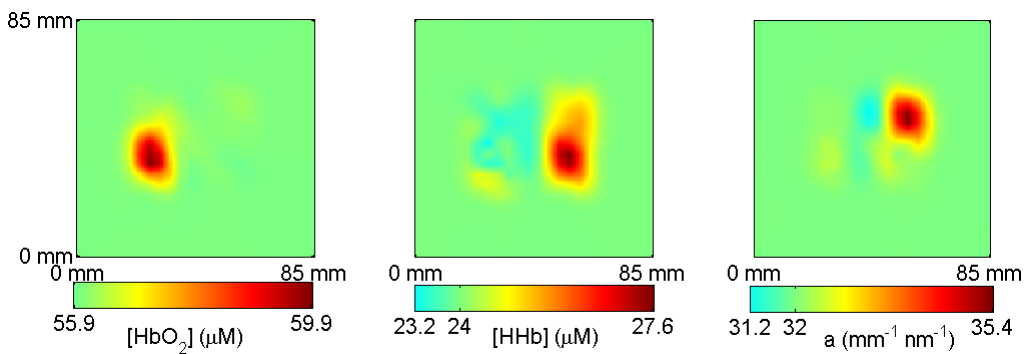


Figure 8.30: Reconstructed images of oxy-haemoglobin (left), deoxy-haemoglobin (centre) and scattering amplitude (right) from measurements at wavelengths 680 nm, 710 nm, 730 nm and 830 nm.

nm, 720 nm and 910 nm simultaneously. Figure 8.28 shows the reconstructed images using measurement wavelengths at 680 nm, 720 nm and 830 nm simultaneously. For both sets of measurement wavelengths, the images exhibit cross-talk between chromophores and also scattering. Cross-talk is slightly more evident in the images reconstructed using the wavelengths selected by our method. Cross-talk is particularly significant between oxy- and deoxy-haemoglobin concentrations in the deoxy-haemoglobin reconstruction. The $[\text{HbO}_2]$ is underestimated and $[\text{HHb}]$ overestimated in both cases.

The same simulation study was performed using the four wavelengths found using our method (680 nm, 710 nm, 730 nm and 830 nm) and the two criteria method (650 nm, 710 nm, 730 nm and 910 nm). Figure 8.29 shows the xy cross-sectional images at $z=8.4$ mm, using the wavelengths selected by the two criteria method. Figure 8.30 shows the images for the wavelengths selected using our method. Due to the limited sensitivity of the measurement to scattering to regions deeper than 12 mm below the surface, the reconstruction were found to have better contrast at $z=8.4$ mm than at $z=11.7$ mm. The reconstructions for both sets of wavelengths are similar and show very small cross-talk. However, the reconstructed values using the wavelengths selected by our method are closer to the true values. Nevertheless, the scattering amplitude and deoxy-haemoglobin concentration are overestimated. The slightly better reconstructions obtained using our wavelength set could be due to the higher sensitivity

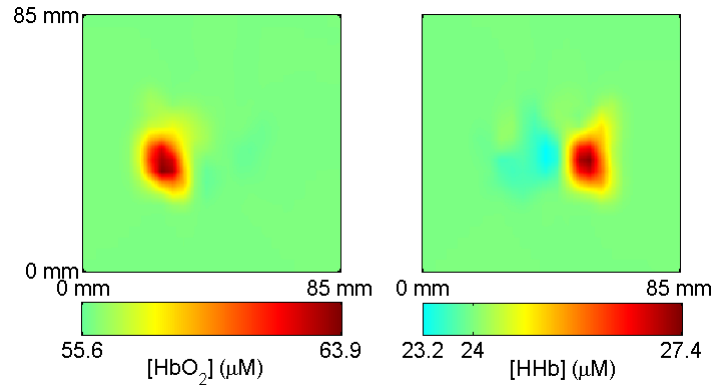


Figure 8.31: Reconstructed images of oxy-haemoglobin (left) and deoxy-haemoglobin (right) from measurements at wavelengths 690 nm and 880 nm.

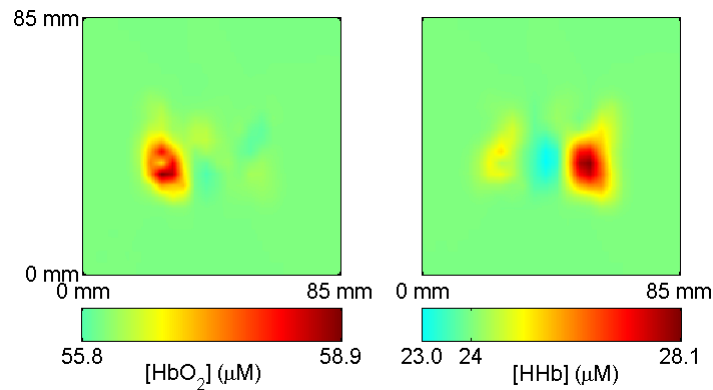


Figure 8.32: Reconstructed images of oxy-haemoglobin (left) and deoxy-haemoglobin (right) from measurements at wavelengths 710 nm and 970 nm.

of these wavelengths to the depth at which the targets are located. Furthermore, even though the sensitivity to scattering is quite limited in deeper regions, the penetration depth of the PMDF_κ has its maximum at these wavelengths. The simulation results indicate that four measurement wavelengths are required, in order to reconstruct images with low cross-talk, from reflection measurements and using linear multispectral reconstruction.

Finally, simulations were performed using the same medium as before, but without the scattering perturbation and for two wavelengths only. Figure 8.31 shows the reconstructions from measurements at wavelengths found using our method (690 nm and 880 nm). The images show a good separation between chromophores but with their concentrations overestimated. Figure 8.32 shows the reconstructions from measurements at wavelengths found using two criteria method (710 nm and 970 nm). The retrieved [HbO₂] is more accurate but there is cross-talk in the [HHb] reconstruction. The presence of cross-talk could be due to sensitivity differences between the two wavelengths. At 970 nm only the top of the target is interrogated, whereas the light at 720 nm is likely to travel through the whole target volume.

One of the wavelength pairs used by the UCL topography system is 770 nm and 850 nm. Simulations were also performed to evaluate the image reconstruction performance for this wavelength pair. The reconstructions are similar to the ones obtained using our optical

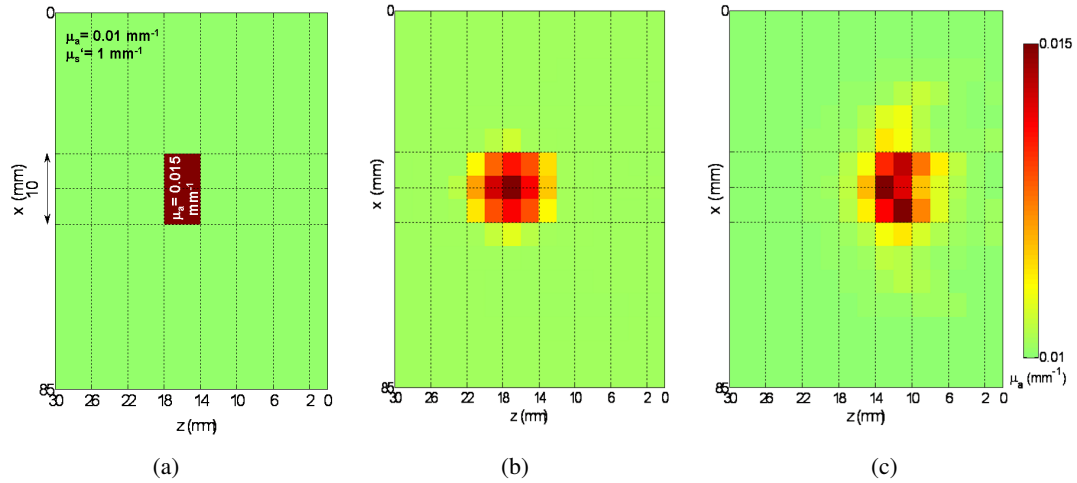


Figure 8.33: (a) The xz cross section of the medium used in the simulation. (b) Image reconstructed using the correct sensitivity matrix. (c) Image reconstructed using the erroneous sensitivity matrix .

wavelengths, but the perturbation [HbO₂] was a bit higher, 64.5 μM .

All the simulations were repeated for larger concentration and scattering amplitude changes, and the results showed a similar tendency.

The sensitivity matrix is usually calculated using an estimate of the average optical properties of the medium, and subsequently used to reconstruct images from measurements at different wavelengths. Given that the optical properties of the medium are wavelength dependent, we analyse how reconstructed images are affected by using sensitivity matrices generated for optical properties different from the true properties. Data were generated for a homogeneous medium with the same optical properties as those used to generate the PMDF_r ($\mu_a=0.01 \text{ mm}^{-1}$ and $\mu'_s=1 \text{ mm}^{-1}$, refer to §8.4.1), with an embedded absorption perturbation centrally placed at a distance of 16 mm below the imaging surface ($z=0$) with an absorption coefficient $\mu_a=0.015 \text{ mm}^{-1}$, as depicted in figure 8.33(a). The array of sources and detectors was described in §3.3.1. However, here the maximum source-detector separation was considered to be 41 mm. Figure 8.33(b) shows an image reconstructed using a sensitivity matrix generated for a medium with the true optical properties. To emphasise the effect of using inaccurate estimates of the optical properties to calculate the sensitivity matrix, we used optical properties that result in very different PMDFs, such as those used to generate PMDF₄ ($\mu_a=0.05 \text{ mm}^{-1}$ and $\mu'_s=1.5 \text{ mm}^{-1}$, refer to §8.4.1). Figure 8.33(c) shows the image reconstructed using this erroneous sensitivity matrix. The absorption perturbation appears misplaced in the z direction by approximately 4 mm.

It is common to consider the typical optical properties of biological tissue at NIR wavelengths to be $\mu_a=0.01 \text{ mm}^{-1}$ and $\mu'_s=1 \text{ mm}^{-1}$. To explore the consequences of using these general optical properties to reconstruct images from measurements performed using the wavelength sets obtained using our method, we compare the measurement sensitivity of the reference PMDF (PMDF_r) profile in the z direction (line at the midpoint between optodes) with that of the PMDF profiles at different wavelengths. Thus we can compare the magnitude

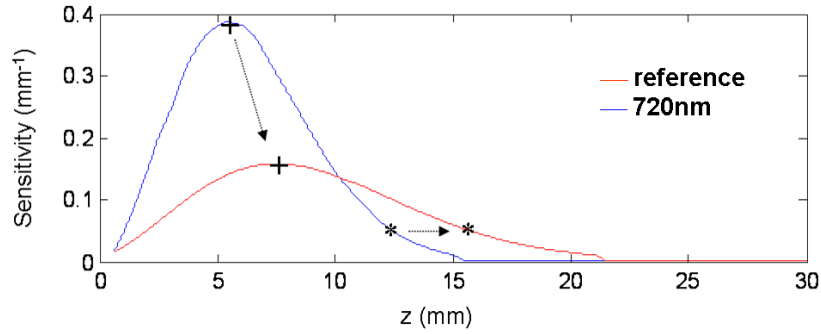


Figure 8.34: PMDF_r and PMDF₇₂₀ profiles. The marks on PMDF₇₂₀ represent a unit perturbation in the medium and the arrows indicate where it would occur if PMDF_r was used to locate the perturbation.

of the sensitivity between two PMDFs and extract the depth difference. This is a very simplified analysis and it should be noted, of course, that this analysis in terms of imaging is more complex and not so linear.

Figure 8.34 shows the PMDF_r and PMDF profile at 720nm (PMDF₇₂₀), which we use as an example. Consider that the measurements are performed at 720 nm, and the marks on the PMDF₇₂₀ to be unit perturbations in the medium, hence the measured intensity change would be equal to the sensitivity value. Then, if the optical properties of the medium are unknown and we use the PMDF_r to identify where the perturbation occurs, it would provide a position different from the true position by placing the perturbation in a region with equivalent sensitivity, as indicated by the arrows in figure 8.34. If the perturbation occurs in regions where the measurement PMDF profile, PMDF₇₂₀, has higher sensitivity than the imaging PMDF profile, PMDF_r, then it would be reconstructed in a region where the sensitivity of the PMDF_r is higher. Furthermore, the magnitude difference results in an overestimation of the optical properties.

The above analysis is performed throughout the whole PMDF depth and for the 3 wavelengths selected using our method and Corlu's method [Corlu *et al.*, 2005]. The depth differences are displayed in figure 8.35. This gives us a rough indication of the maximum depth differences that can be expected, if a sensitivity matrix generated based on typical optical properties is used to reconstruct images. For deeper regions the depth difference can be as large as 7.5 mm. If we consider an absorption perturbation localised in the grey matter, this means that it could appear in the reconstructed image misplaced between 2 mm and 7.5 mm. At approximately 10 mm depth the difference is zero because the PMDF intersect (have the same value). Another zero can be found closer to the surface, although here the magnitude of the sensitivities is different. The depth differences in the regions of highest sensitivity varies between 3 mm and 5.5 mm. The largest differences are observed when measurements are performed at 650 nm. If we now consider the reference PMDF to be PMDF₇₂₀ and perform the same analysis, the depth differences are smaller than 1.5 mm, as shown in figure 8.36. As expected, the wavelengths set selected using our method returns the smallest differences.

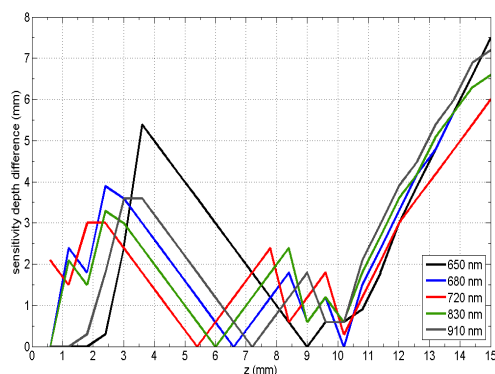


Figure 8.35: Depth difference between equal sensitivities of PMDF₇ and PMDFs at wavelengths selected using our method (680 nm, 720 nm and 830 nm) and wavelengths selected using the two criteria method (650 nm, 720 nm and 910 nm).

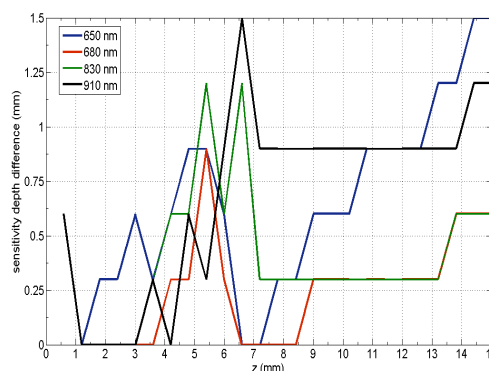


Figure 8.36: Depth difference between equal sensitivities of PMDF₇₂₀ and PMDFs at the other wavelengths selected using our method (680 nm and 830 nm) and wavelengths selected using the two criteria method (650 nm and 910 nm).

8.5 Conclusions

The method developed by Corlu *et al* [2005] to optimise the wavelengths for DOT systems, based on multispectral reconstruction, has proven to successfully overcome the non-uniqueness problem and to minimise cross-talk between chromophores. Based on this method a new wavelength optimisation method was introduced, which uses SSD to compare PMDFs at different wavelengths. By minimising this quantity one can guarantee that the sensitivity to changes in absorption and scattering changes in the studied medium are similar. The wavelength distributions obtained using this new method differ from those determined without comparing the similarity between sensitivity profiles. From our analysis, we conclude that the differences in the interrogated volume at NIR wavelengths are very small for the three layer adult head model and slightly more significant for the homogeneous model. This indicates that the results depend on the accuracy of the head model. However, increasing the complexity of our model, by including CSF, is expected to have a small influence on the results. Our new approach can be applied to other tissue types, like breast, muscle or neonatal brain, and also to other diffuse optical tomography/topography measurement types.

Measurements with overlapping sensitivities ensure that the same regions are being sampled, however, they do not significantly improve the image reconstruction. As expected, the quality of the reconstructions depends on the accuracy of the estimated optical properties of the medium used in the forward model. In CW optical topography using linear multispectral reconstruction, at least four measurement wavelengths are required to separate absorption from scattering and separate two chromophores from one another.

The selection of optimal wavelengths is limited by the commercial availability of laser diodes, although the error and consequences of using wavelengths with non-overlapping sensitivities can be quantified from our analysis.

This study could depend on the background chromophore concentrations, scattering amplitude and extinction coefficients used, which requires further analysis, but it is not likely that it will have a significant influence on the PMDFs differences. Future work could also include phantom experiments, to help validate the method. The results are not expected to

be exactly the same because there are many other factors involved, like differences between the signal to noise ratio of the laser sources at different wavelengths. Simulations could help evaluate the effect of noise on the reconstructed images, and indicate which wavelengths are more sensitive to noise.

The main conclusion is that for the adult human head it can be assumed that the volumes of interrogated tissue at different NIR wavelengths are identical. Hence, measurements performed at different NIR wavelengths provide information from the same brain regions.

Imaging of functional brain activation

The brain regions involved in social cognition have been extensively studied in adults. However, the relation between social stimuli and functional activation of the infant brain is less well known. One of the main reasons is the inadequacy of imaging systems, such as fMRI and PET, to image conscious infants. Optical topography is a potential imaging technique for this age-group, due to its portability, safety and low sensitivity to motion since the imaging probe is attached to the head of the subject. However, it can only image superficial regions of the brain.

Researchers at the Centre for Brain and Cognitive Development (CBCD at Birkbeck, University of London) have performed several studies using an optical topography system developed at UCL, to investigate regions of the infant brain areas involved in social cognitive processes. A recent study performed by Lloyd-Fox *et al* [2009c] showed an increase in [HbO₂] in the posterior superior temporal region bilaterally, when visual social dynamic stimuli were presented to twenty-four five-month-old infants. Grossmann *et al* [2008] studied the social perception of gaze in four-month-old infants and detected cerebral activation in areas of the temporal and prefrontal cortex. However, these and many other NIRS studies simply analyse single source-detector hemodynamic responses and do not attempt to reconstruct haemodynamic maps of the region of study. From a study performed at CBCD, 3D images are reconstructed representing the haemodynamic responses to different social related stimuli and the social brain regions are identified.

9.1 Methods

9.1.1 Stimuli

Thirteen five-month-old infants took part in the study. The infants sat on their parent's lap while the stimuli were displayed on a 46-inch plasma screen. The experiment consisted of three different social stimuli performed by female actors who either moved their eyes, mouth or hand. In the eye condition (E) the eyes were moved left or right (figure 9.1(c)), silent vowel mouth movements were used for the mouth condition (M) (figure 9.1(b)) and hand games were performed for the hand condition (H) (figure 9.1(a)). The baseline condition consisted

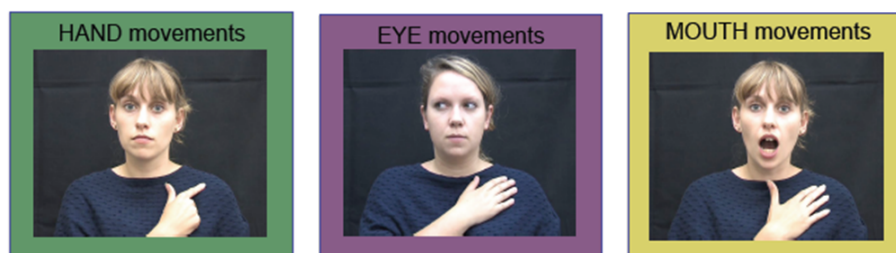


Figure 9.1: Social stimuli presented to 13 healthy five-month old infants, which consisted of hand, eye and mouth movements for 10 seconds per trial [Lloyd-Fox *et al.*, 2009b].

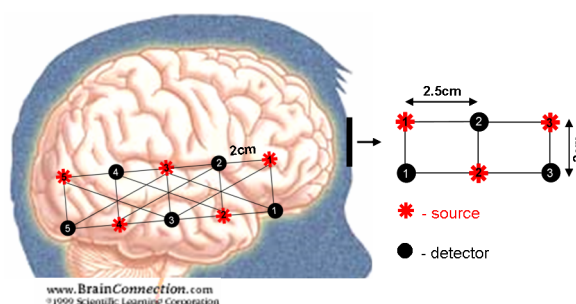


Figure 9.2: Lateral and frontal array, showing the approximate position of the array of the infant head.

of non-social dynamic motion (rotating toy).

Before the actual experiment starts, there is a rest period so that infants become comfortable and familiarised with the environment and experiment setup. Then, for each experimental trial, a stimulus with a duration of 10s was alternated with a 10s baseline measurement. The experimental stimulus and baseline are repeated at least 6 times. Trials where infants failed to look at the stimuli were excluded.

9.1.2 Data acquisition and processing

The UCL optical topography system described in § 3.2.1.1, with wavelengths at 770 nm and 850 nm, was used for the data acquisition. An array of sources and detectors appropriate for this study was designed by the CBCD. Two arrays were placed on the temporal lobe over each hemisphere, which consisted of 5 sources and 5 detectors each, making a total of 19 channels, with source-detector separations at 2 cm and 4.5 cm (figure 9.2). Another 3 source-detector pairs, at separations 2.5 cm, were placed over the frontal lobe (figure 9.2).

Once data were acquired for all the trials, each trial was separated into blocks consisting of 4s baseline, followed by a 10s trial and 10s baseline. Assuming that the initial and final magnitude of the measurements is identical, each block was detrended using a linear fit between the two baselines [Blasi *et al.*, 2007].

Trials for each infant and experimental condition were averaged. Subsequently, each experimental condition was averaged across all thirteen infants. This results in a set of data over time for each of the three social stimuli. Their respective baselines were averaged over time.

9.1.3 Image reconstruction

The TOAST forward model was used to calculate the Jacobian matrix, for each of the three arrays, using a FEM mesh with 32615 nodes, 21845 elements and dimensions 120 mm \times 60 mm \times 40 mm. The medium is considered to be homogeneous with $\eta = 1.4$ and optical properties $\mu_{a,brain}(\lambda)$ and $1/3 \times \mu_{s',brain}(\lambda)$ (neonate $\mu_{s'}$ is approximately 1/3 of that of adult, refer to §2.2) for each wavelength λ (770 nm and 850 nm), as described in §2.2 and appendix A. Three-dimensional multispectral reconstruction with NOSER type regularisation was used to reconstruct images of oxy-haemoglobin concentration changes [HbO₂] and deoxy-haemoglobin concentration changes [HHb], for each experimental condition and probe.

9.2 Results

For each stimulus, data were averaged over 1s time intervals centred at 13s, which corresponds in general to the point of maximum activation. The L-curve method was used to find the constant regularisation parameter. However, the L-curve method fails to find a regularisation parameter when the problem is normalised by the standard deviation of the data and row normalisation is applied, which possibly occurs since the problem becomes less ill-conditioned. Nevertheless, the GCV method successfully finds a regularisation parameter despite normalisation. Without normalisation the L-curve method selects a regularisation parameter that provides images with higher accuracy than those generated using the regularisation parameter acquired via the GCV method. However, we find that the images improve when normalisation are use.

Figure 9.3 shows xy cross-sections at $z=8$ mm of reconstructed images representing [HbO₂], corresponding to the probe placed on the left temporal lobe, while infants were presented with a hand movement stimulus. Note that for both hemispheres, images are displayed with the posterior temporal area on the left and anterior area on the right. Recall that the stimulus was initiated at 4s and ends at 14s, and is followed by a baseline condition. An increase in [HbO₂] is visible in the posterior site between 10s and 16s. A small region in the anterior site also shows an increase in oxy-haemoglobin. The increase appears to be preceded by a global decrease in oxy-haemoglobin. Figure 9.4 shows xy cross-sections at $z=8$ mm of reconstructed images of [HHb] for the hand condition. An increase in HHb is visible as a response to the stimulus and is followed by a decrease, which occurs 3s after the increase in [HbO₂]. The change in [HHb] occurs in the central and anterior regions, which are adjacent positions to sites of increase in [HbO₂]. The images of [HbO₂] and [HHb] for the mouth condition are displayed in figure 9.5 and figure 9.6, respectively. An increase in [HbO₂] occurs 1s post stimulus onset in the central region, and moves to the posterior site once the control condition begins. Figure 9.6 shows that during the stimulus the [HHb] increases. A decrease in deoxy-haemoglobin is visible between 14s and 24s in the anterior region. Figure 9.7 shows oxy-haemoglobin images for the eye condition. A slight decrease in [HbO₂] is observed followed by an increase. The maximum change in [HbO₂] occurs 8-12s post stimulus onset. Figure 9.8 shows a decrease in [HHb] that starts just as the stimulus is ending, preceded by an increase. The decrease in [HHb] is located at the central and anterior sites.

Figure 9.9 shows xy cross-sections at $z=6$ mm of reconstructed images of [HbO₂],

measured by the probe placed on the right temporal lobe, while infants visualised hand movements. The maximum change in $[\text{HbO}_2]$ occurs 5-13s post stimulus onset and mainly at a posterior location. Some activity is also visible at the central/inferior site and anterior region. The regions of maximum deoxy-haemoglobin decrease appear in figure 9.10 adjacent to the regions of maximum change in $[\text{HbO}_2]$. The response latency is similar for both chromophores. Figure 9.11 shows xy cross-sections at $z=6$ mm of images of $[\text{HbO}_2]$ for the mouth condition. The maximum $[\text{HbO}_2]$ increase occurs at the anterior site, which becomes visible in the images 1s after the stimulus onset. There is also an increase in $[\text{HbO}_2]$ at the posterior and central regions. The haemodynamic activity predominantly influences data collected using the short separation source-detector pairs, which is also reflected in the images. Figure 9.12 shows xy cross-sections at $z=6$ mm of images of $[\text{HHb}]$ for the same condition, where a decrease in deoxy-haemoglobin can be seen in the central region, with a response latency of 6s. Figure 9.13 shows xy cross-sections at $z=8$ mm of images of $[\text{HbO}_2]$ for the eye condition. An increase in oxy-haemoglobin is visible 2s post stimulus onset, which is located in the posterior/central region. Figure 9.14 shows a large increase in deoxy-haemoglobin during the stimulus, followed by a centrally located decrease.

The images reconstructed of oxy- and deoxy-haemoglobin concentration changes at depth $z=6$ mm from the measurements performed on the frontal cortex while hand stimuli were presented are displayed in figure 9.15 and figure 9.16, respectively. A strong change in oxy-haemoglobin occurs in the centre. The $[\text{HHb}]$ images show a decrease in the chromophore concentration at the extremities, although it also exhibits regions of concentration increase. The images for the mouth condition show a similar pattern, as displayed in figure 9.17 and figure 9.18, for $[\text{HbO}_2]$ and $[\text{HHb}]$ respectively. The maximum $[\text{HbO}_2]$ occurs at the centre and just to the right of centre. The maximum $[\text{HHb}]$ is observed on the right side, but a smaller change in deoxy-haemoglobin can also be observed on the left. Figure 9.19 shows the oxy-haemoglobin concentration change for the eye condition. Here, the maximum $[\text{HbO}_2]$ occurs in the centre at 6s and again at 11s from the stimulus onset. Figure 9.20 shows the deoxy-haemoglobin concentration change for the eye condition. Besides the decrease in deoxy-haemoglobin at the lateral sites, a decrease is also observed at the top right 2s post stimulus onset.

Figure 9.21 shows the xz cross-section at $y=30$ mm (section through the probe length) of the image corresponding to the oxy-haemoglobin concentration change in the right hemisphere evoked by eye movements. The maximum activity occurs at depths between 5 mm and 10 mm. Nevertheless, regions down to 15 mm still reveal activity. A similar image can be seen in figure 9.22 for the mouth condition. Differences between interrogated regions by different source-detector separations become evident in this image. Small source-detector separation can only measure activity down to approximately 8 mm depth.

Using the reconstructed images, the regions where the increases in oxy-haemoglobin and decreases in deoxy-haemoglobin occurred were estimated for all the three conditions and probes, and are depicted in figure 9.23.

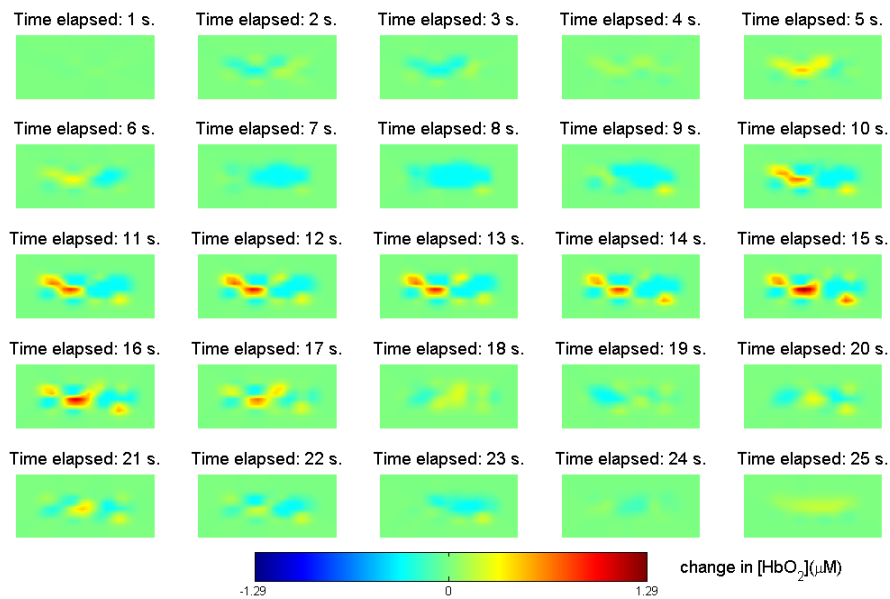


Figure 9.3: Reconstructed images of changes in $[HbO_2]$ occurring in the left hemisphere over time, showing the response to hand movements.

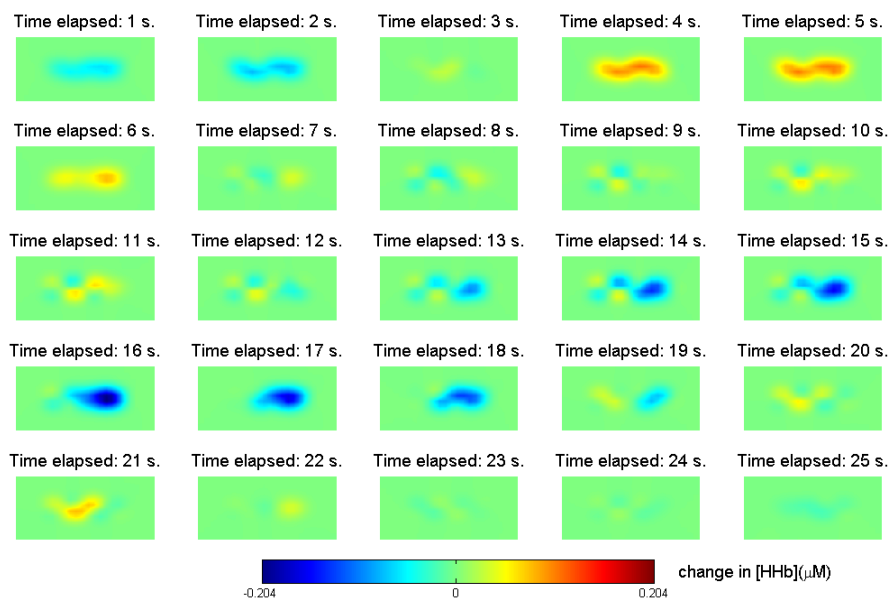


Figure 9.4: Reconstructed images of changes in $[HHb]$ occurring in the left hemisphere over time, showing the response to hand movements.

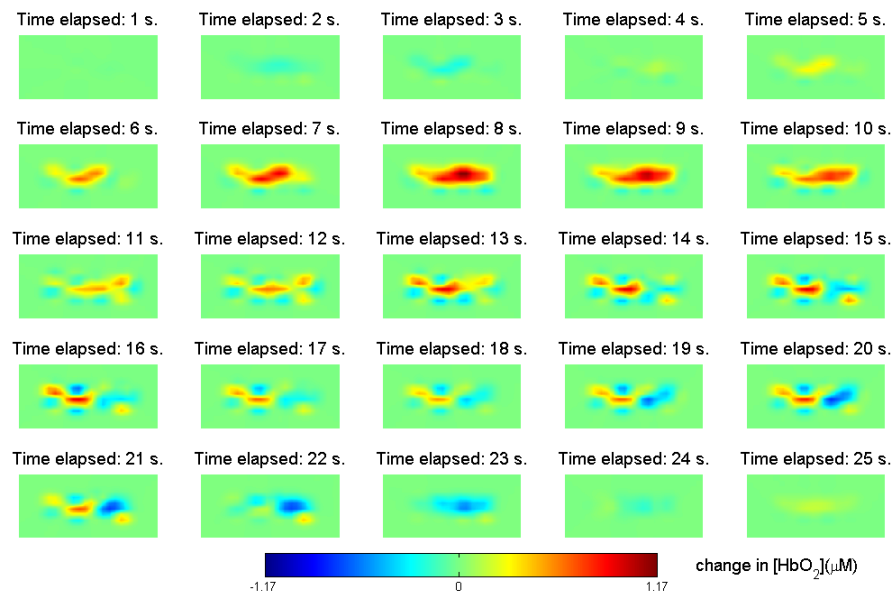


Figure 9.5: Reconstructed images of changes in $[\text{HbO}_2]$ occurring in the left hemisphere over time, showing the response to mouth movements.

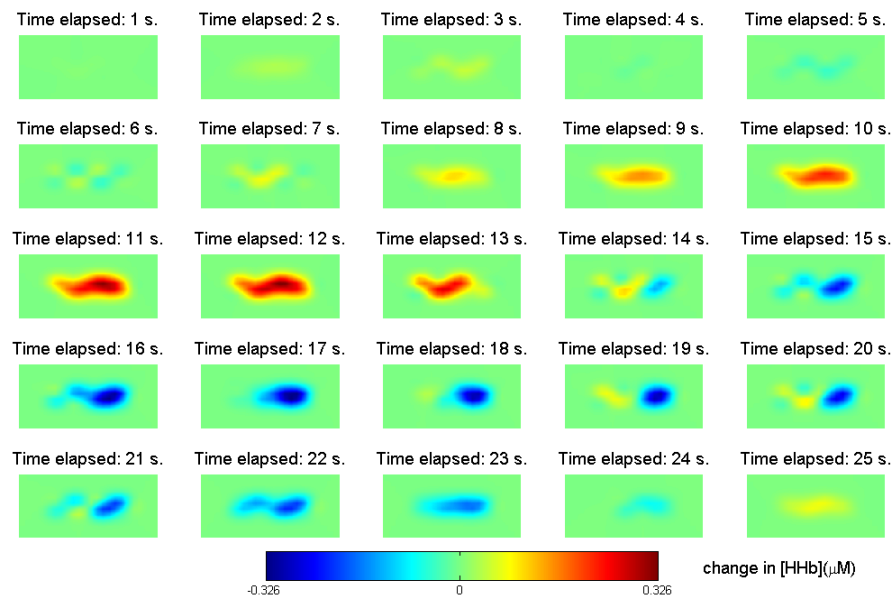


Figure 9.6: Reconstructed images of changes in $[\text{HHb}]$ occurring in the left hemisphere over time, showing the response to mouth movements.

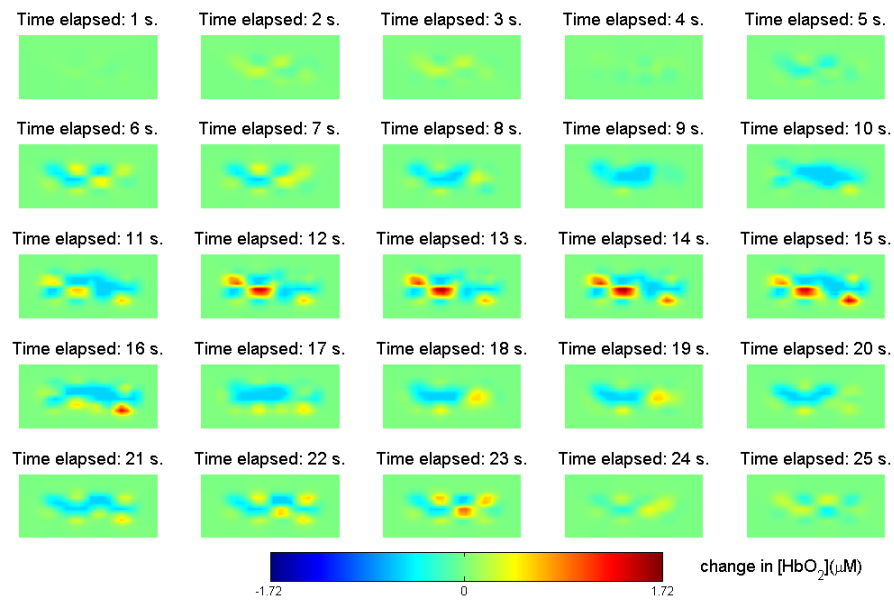


Figure 9.7: Reconstructed images of changes in $[HbO_2]$ occurring in the left hemisphere over time, showing the response to eye movements.

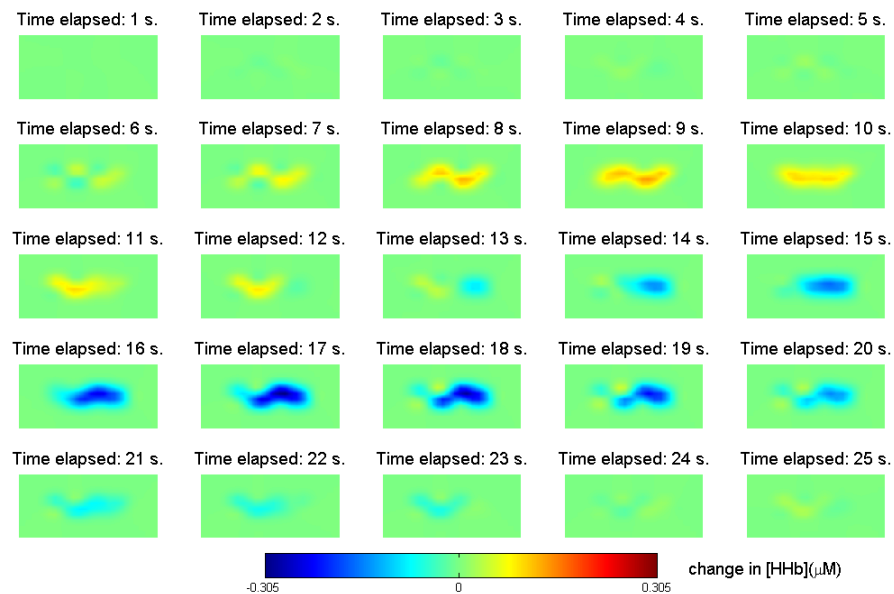


Figure 9.8: Reconstructed images of changes in $[HHb]$ occurring in the left hemisphere over time, showing the response to eye movements.

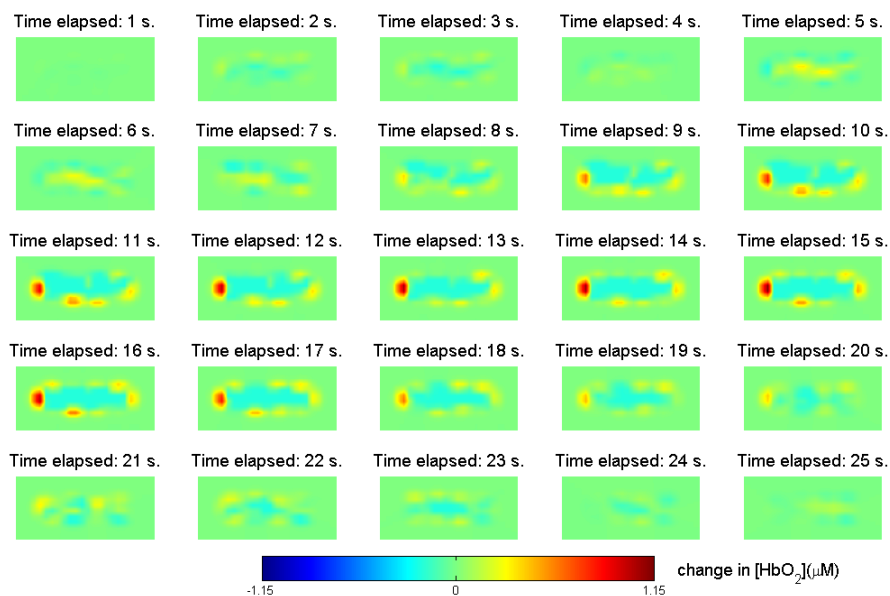


Figure 9.9: Reconstructed images of changes in $[\text{HbO}_2]$ occurring in the right hemisphere over time, showing the response to hand movements.

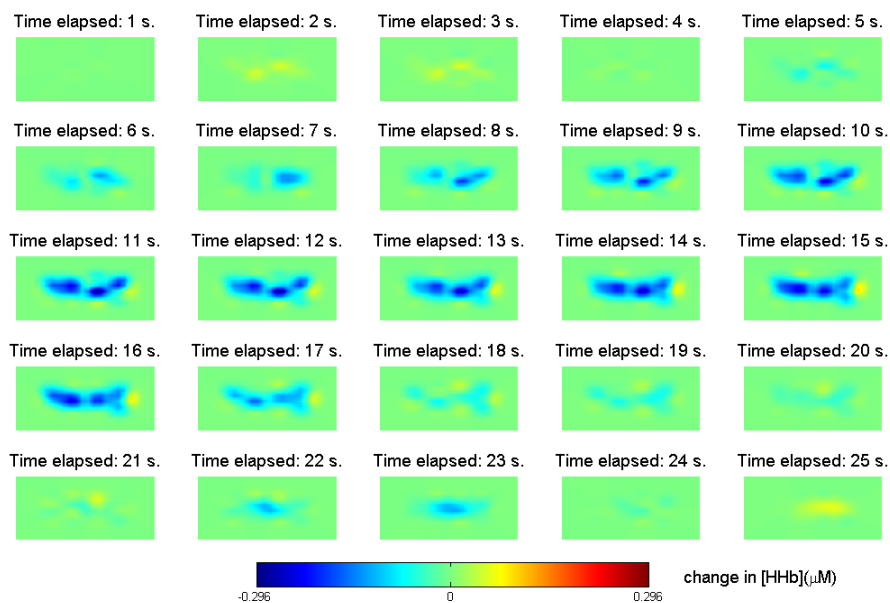


Figure 9.10: Reconstructed images of changes in $[\text{HHb}]$ occurring in the right hemisphere over time, showing the response to hand movements.

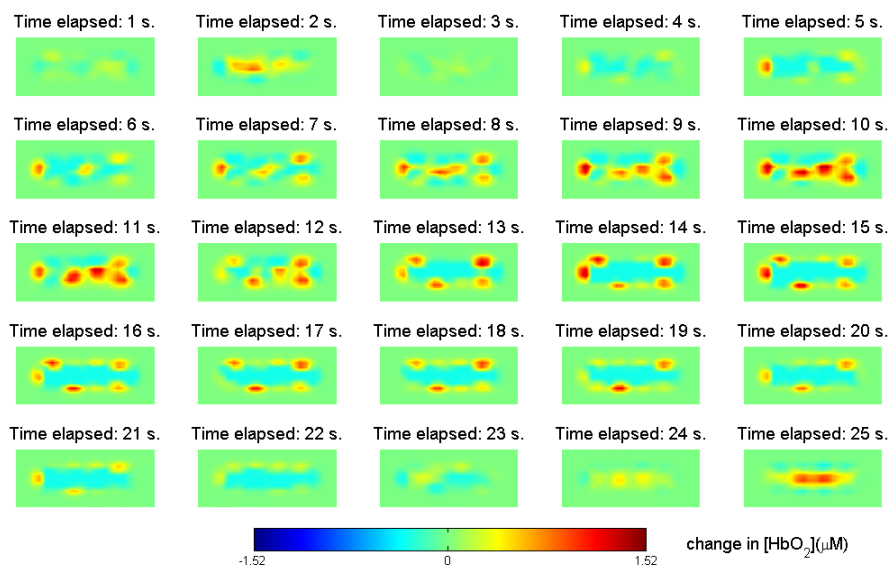


Figure 9.11: Reconstructed images of changes in $[HbO_2]$ occurring in the right hemisphere over time, showing the response to mouth movements.

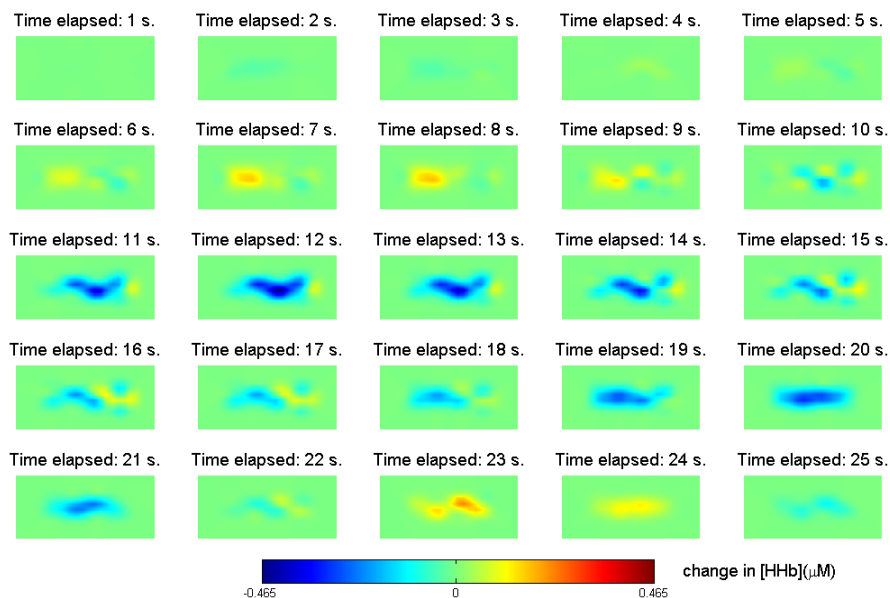


Figure 9.12: Reconstructed images of changes in $[HHb]$ occurring in the right hemisphere over time, showing the response to mouth movements.

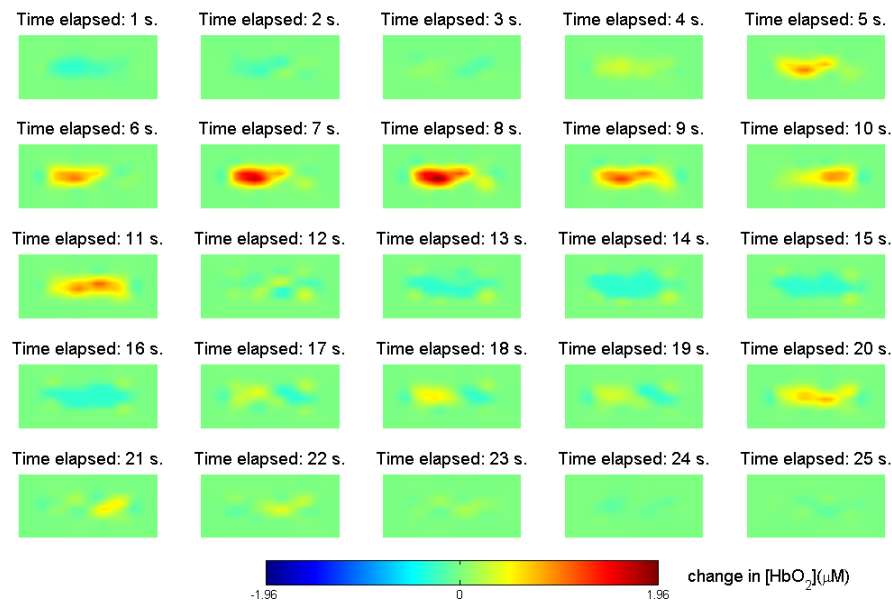


Figure 9.13: Reconstructed images of changes in $[HbO_2]$ occurring in the right hemisphere over time, showing the response to eye movements.

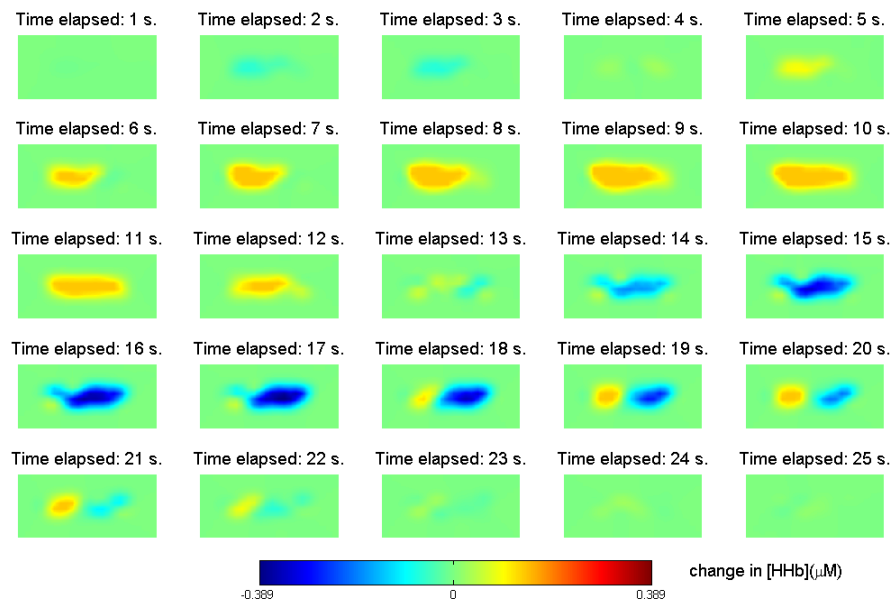


Figure 9.14: Reconstructed images of changes in $[HHb]$ occurring in the right hemisphere over time, showing the response to eye movements.

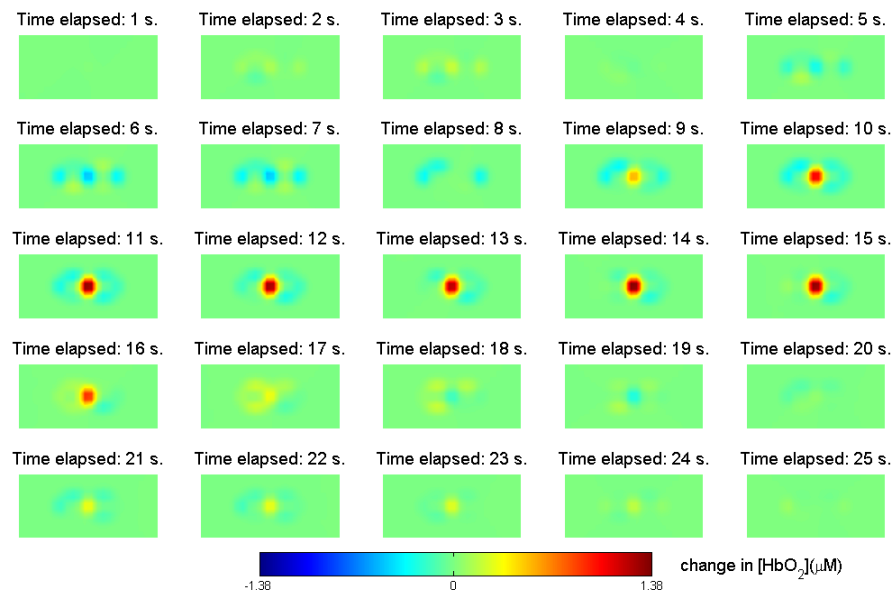


Figure 9.15: Reconstructed images of changes in $[HbO_2]$ occurring in the frontal lobe over time, showing the response to hand movements.

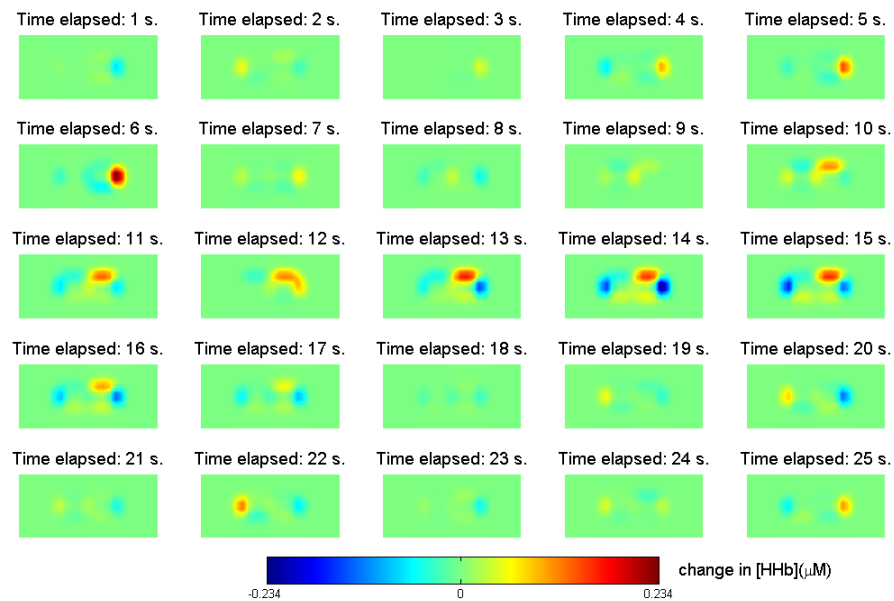


Figure 9.16: Reconstructed images of changes in $[HHb]$ occurring in the frontal lobe over time, showing the response to hand movements.

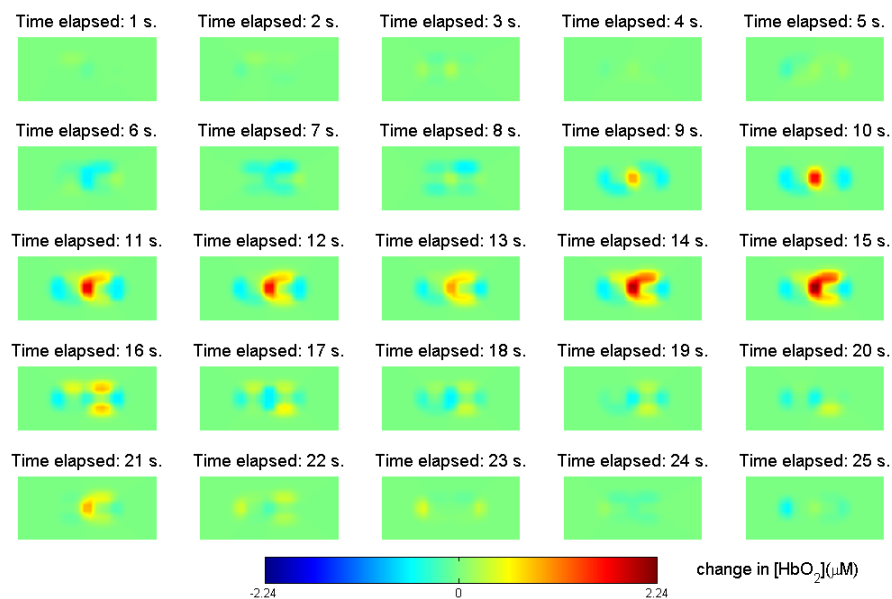


Figure 9.17: Reconstructed images of changes in $[HbO_2]$ occurring in the frontal lobe over time, showing the response to mouth movements.

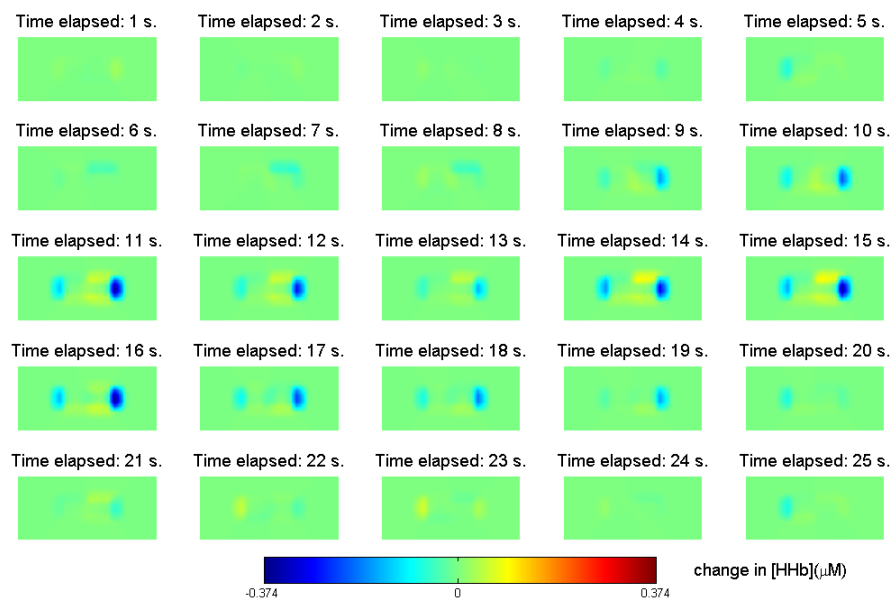


Figure 9.18: Reconstructed images of changes in $[HHb]$ occurring in the frontal lobe over time, showing the response to mouth movements.

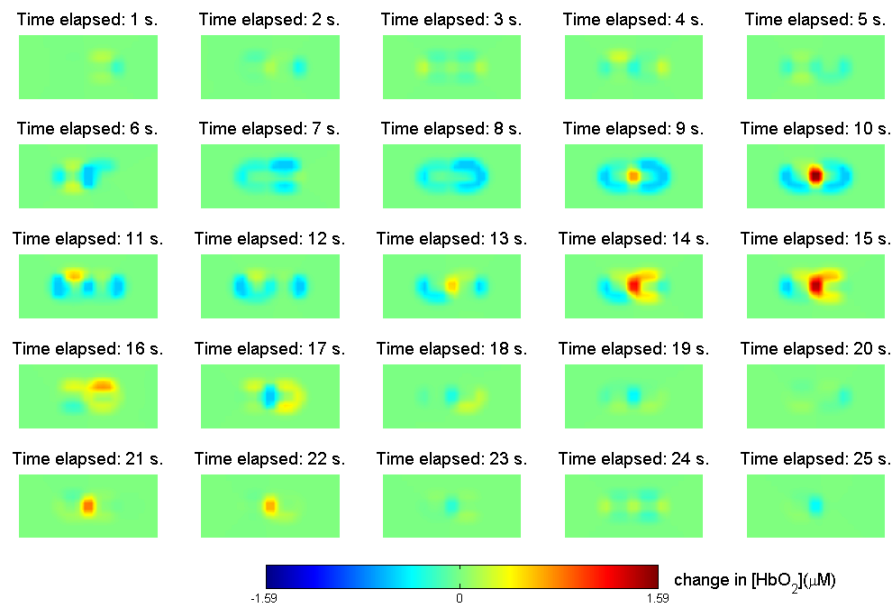


Figure 9.19: Reconstructed images of changes in $[HbO_2]$ occurring in the frontal lobe over time, showing the response to eye movements.

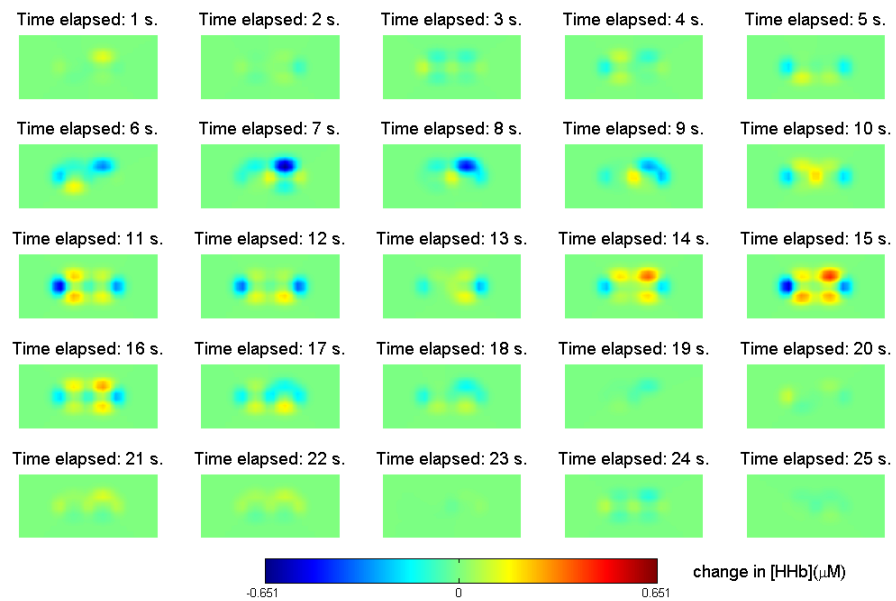


Figure 9.20: Reconstructed images of changes in $[HHb]$ occurring in the frontal lobe over time, showing the response to eye movements.

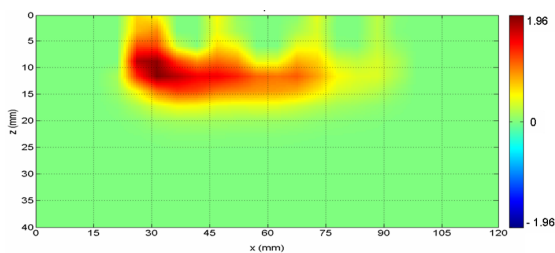


Figure 9.21: xz cross-sectional images of changes in $[HbO_2]$ in the right hemisphere evoked by eye movements.

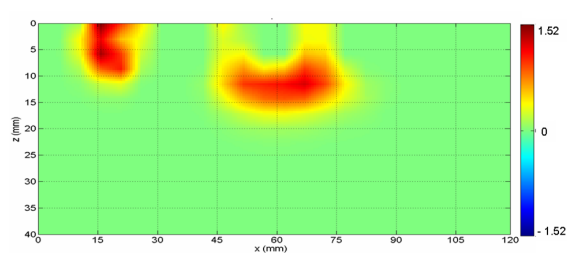


Figure 9.22: xz cross-sectional images of changes in $[HbO_2]$ in the right hemisphere evoked by mouth movements.

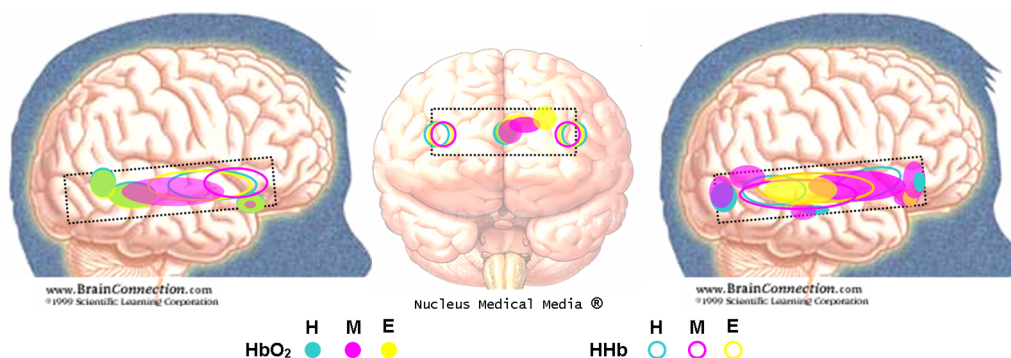


Figure 9.23: Schematic illustration of the regions of maximum haemodynamic activity for the three experimental conditions, for the left pad (left), frontal pad (centre) and right pad (right). For both hemispheres the posterior temporal area is on the left and the anterior area is on the right.

9.3 Discussion

Optical topography can clearly measure visually induced haemodynamic activity in regions of the social brain. The array of optodes was placed on both temporal lobes and the frontal lobe. Images of oxy-haemoglobin and deoxy-haemoglobin concentration changes were reconstructed for all three imaged regions, for three different experimental conditions: hand, mouth and eye movements.

The images obtained from the left pad representing oxy-haemoglobin change are quite similar for the three conditions, showing a maximum change in the posterior region, which is likely to be associated with activity in the posterior superior temporal sulcus (pSTS). Strong activity is also detected in the anterior area of the image, which is possibly reflecting activity taking place in the inferior frontal gyrus (IFG). For the mouth movement there is also an increase in oxy-haemoglobin in the regions corresponding to the mid-STs. For the hand and eye conditions the maximum $[HbO_2]$ has a latency from the stimulus onset of 7-12s and is preceded by a decrease in its concentration. This is reasonable, since the great metabolic demand in the activated brain regions causes an initial small decrease in local $[HbO_2]$, to which the brain responds by increasing the blood flow that results in an increase in local $[HbO_2]$. The mouth condition seems to trigger an almost instantaneous response with longer duration. In general, the images of changes in deoxy-haemoglobin, occurring in the left hemisphere, show an increase in $[HHb]$ followed by a decrease. This decrease has a latency of approximately 4s from the time at which the $[HbO_2]$ is maximum, and appears adjacent to regions of increase in

[HbO₂].

The activity in the right hemisphere induced by hand and mouth movements is mostly measured by source-detector pairs at small separations, and therefore the images show activity in regions immediately between optodes with small separations. Activity in the anterior region can be seen in the images of the right hemisphere for all conditions. As before, this may correspond to activity in the IFG. The images reconstructed from measurements performed while infants observed hand movements show a maximum increase in [HbO₂] in the posterior region, which may correspond to activity measured in the pSTS. For the mouth condition activity is observed for regions corresponding to the pSTS and mid-STS/anterior temporal region. Whereas for the eye condition activity is seen for regions corresponding to posterior/mid STS. In general, the latency period from the stimulus onset is shorter for the right hemisphere than for the left hemisphere. Once more, decreases in [HHb] occur in regions next of maximum increase in [HbO₂]. A longer latency from the stimulus onset is observed for [HHb] than for [HbO₂], except for the hand condition where they both respond at approximately 4s after the stimulus starts.

As mentioned previously, a study by Lloyd-Fox et al [2009c] to determine activity in regions of the social brain of five-month-old infants using the same conditions as those described above, showed an increase in [HbO₂] in posterior regions of the temporal lobes. Activation was also found in the anterior temporal/inferior frontal region, which is in agreement with our findings.

Puce *et al* [1998] performed a fMRI study to determine which regions of the temporal cortex became activated in adult humans viewing eye and mouth movements, and found strong bilateral activation in the pSTS. They also found that changes in gaze direction activate the right STS more than the left. Studies of neural activity involved in the perception of eye, hand and mouth movements suggest that the main focus of activation is the pSTS [Allison *et al.*, 2000]. Responses to hand and mouth movements have been also measured in more anterior regions [Allison *et al.*, 2000].

Pelphrey *et al* [2005] conducted a fMRI study to compare cerebral activity during viewing of eye, mouth and hand movements. They found bilateral activation of the posterior temporal-occipital cortex. The right hemisphere showed a greater response to the stimuli, and activity evoked by mouth movements dominated the left hemisphere. Activity induced by eye movements was greatest in the pSTS region. Mouth movements activated more anterior regions of the STS, whereas the most posterior activation was evoked by hand movements. A similar posterior to anterior structural organisation of the STS can be seen in our oxy-haemoglobin concentration change images of activity in the right hemisphere: the hand condition evokes activity in the most posterior region, followed by the eye activity region and mouth activity in more anterior regions. Activity evoked by mouth movements in the left hemisphere also occurs in more anterior regions relative to other stimuli activity. It has been reported that the STS on the right hemisphere of newborns is larger and presents a degree of complexity higher than that on the left hemisphere [Dubois *et al.*, 2008], which could explain the smaller structural organisation level observed for the optical topography images of the left hemisphere. Pelphrey *et al* [2005] also observed activation in the IFG, predominantly induced by eye and mouth

movements. Thompson *et al* [2007] observed activity in the left IFG of subjects viewing hand and mouth movements. This reinforces the hypothesis that the optically measured activity in the most anterior site has its origin in the IFG.

The images reconstructed from measurements performed on the frontal lobe show a strong increase in $[\text{HbO}_2]$ mainly in the centre and slightly deviated to the right. A decrease in $[\text{HHb}]$ is observed in the surrounding regions. PET and MRI studies of spatial working memory, i.e., short-term storage and manipulation of information regarding spatial positions of stimuli, observed activity lateralised to the right in the frontal cortex [Wager & Smith, 2003]. Hand actions were visualised by children during a fMRI study and activity was found in the right hemisphere of the medial prefrontal cortex (mPFC) [Ohnishi *et al.*, 2004]. The optical probe placed over the frontal cortex contains few channels, which results in images with low spatial accuracy and resolution. For example, if activity occurs in a regions slightly to the right of the central region, then an intensity change measured by the central source-detector pair will still cause a reconstructed image to display the activated region in the exact centre of the image.

In the previous chapter, we analysed the similarity between sensitivity profiles for wavelengths in the NIR range. We observed that for 770 nm and 850 nm the interrogated volume is similar. Therefore, the information measured at these wavelengths should not differ significantly. Also, simulations were performed to study the cross-talk between oxy- and deoxy-haemoglobin present in the images using measurements acquired at these wavelengths. The image reconstruction was able to separate both chromophores, but $[\text{HbO}_2]$ was overestimated by 15 % and $[\text{HHb}]$ by 10 %. Therefore we expect our concentration change values to be overestimated. Nevertheless these values are of the same order of magnitude as those found in NIRS studies [Lloyd-Fox *et al.*, 2009c; Grossmann *et al.*, 2008; Carlsson *et al.*, 2008]. As expected, the decreases in $[\text{HHb}]$ are smaller than the increases in $[\text{HbO}_2]$.

The optical properties of the brain are not exactly known, and if the sensitivity matrix is generated for optical properties very different from the true values it leads to cross-talk in the reconstructions. Some of the decreases in the $[\text{HbO}_2]$ and increases in $[\text{HHb}]$ present in the images (refer for example to figure 9.13 and figure 9.14) could be due to this uncertainty in the optical properties of the brain. Nevertheless, an increase followed by a decrease and vice-versa can be observed in the intensity data (not shown), indicating that there actually is an increase (decrease) in $[\text{HHb}]$ ($[\text{HbO}_2]$) before a decrease (increase).

Haemodynamic activity is observed in the reconstructed images down to a depth of 15 mm. Assuming the scalp thickness in infants to be about 3 mm, skull thickness 5 mm (may be thinner for the temporal region), CSF layer 1 mm and grey matter 3 mm, then the whole cerebral cortex depth and superficial regions of the white matter are being imaged. Short source-detector separations can only measure activity at the surface of the cerebral cortex. As shown in the previous chapter, the spatial accuracy of the haemodynamic activity depends on the accuracy of the optical properties used to generate the sensitivity matrix.

As future work, optical topography should be validated by performing it simultaneously with an imaging modality that has already proven to be successful at identifying activity in regions of the social brain, such as fMRI and PET. The quality of the optical topography images would be greatly improved if more optodes and source-detector separations were used, which

is something to consider for future studies. However, the CGCD and UCL have already tested probes with higher optode density and found that data were more robust and reliable when a smaller number of optodes were used [Lloyd-Fox *et al.*, 2009a]. Also, data were less sensitive to head movements for the type of probe used in this study.

9.4 Conclusions

Three-dimensional images of regions of the social brain were successfully reconstructed, using the multispectral method, from optical topography measurements on five-month-old infants. In general, images of oxy-haemoglobin concentration changes show a decrease in concentration followed by a larger increase. The opposite is seen in images of deoxy-haemoglobin concentration changes. There is a few seconds latency between the stimulus onset and the maximum increase in $[HbO_2]$, and the $[HHb]$ peak tends to occur later. The results are in agreement with adult fMRI studies, where activity is visible in the pSTS, IFG and mPFC. Activity was observed in both hemispheres showing a certain degree of structural organisation, where specific regions of the temporal cortex, possibly STS, respond to certain stimuli. The right hemisphere exhibits a posterior to anterior response in the temporal lobe, possibly within the STS, to viewing hand, eye and mouth movements. The left hemisphere shows less structural organisation, but mouth movements evoke activity in a more anterior region than eye and hand movements. These results suggest that five-month-old infants already have specialised regions for processing social stimuli. Optical topography appears to be a suitable technique for imaging responses to social stimuli in infants.

Summary and future developments

Optical topography is an emerging imaging technique, which has proven to be suitable for monitoring cerebral activity of awake infants. One of the main research applications is the study of brain development. Although other imaging modalities exist, such as fMRI and PET, which have successfully been used to study the adult brain, these techniques have factors that restrict their use in infants. Moreover, the information provided by optical topography, i.e., oxy- and deoxy-haemoglobin concentration changes, cannot be obtained by any other imaging modality. Note that measurements at an appropriate combination of wavelengths enables other chromophore concentrations to be monitored. Most optical topography functional activation studies, in both adults and infants, determine chromophore concentration changes for single source-detector pairs. However, maps of cortical activity, providing volumetric haemodynamic information, can be obtained by inverting the forward problem, which models the light interaction with tissue, and by using overlapping measurements at different source-detector separations. In this thesis different means of improving the quality of three-dimensional optical topography images have been explored.

Image reconstruction in optical topography is an ill-posed and under-determined problem. Tikhonov regularisation can be used to stabilise the solution of the inverse problem, which is achieved by filtering out small singular values of the sensitivity matrix. The amount of filtering is controlled by the regularisation parameter, which has a large influence on the image quality. In this thesis we reviewed some of the most common methods used to select the regularisation parameter. We were aiming to find a method that returned a single regularisation parameter value, did not require a subjective input from the user or any prior information about the feature being examined. Alternative methods were tested on a relatively simple ill-posed problem with a known solution, the deblurring problem, where different levels of noise were added to the blurred data. Finally, the methods were applied to experimental optical topography data. We found that the L-curve method can find a reasonable regularisation parameter for both deblurring and optical topography problems. However, the L-curve method fails when the

multispectral reconstruction method is used to directly obtain the chromophore concentrations, instead of the conventional approach where the absorption and scattering coefficients are reconstructed, and normalisations are applied to the sensitivity matrix. The GCV method has proven to be an effective alternative to the L-curve method.

Another approach used to optimise the linear image reconstruction algorithm in order to provide more qualitatively accurate images, particularly in the depth direction, was the inclusion of a spatially variant regularisation parameter. We have shown using simulations and phantoms, with targets located at different depths from the imaging surface, that better depth discriminations is obtained when NOSER type regularisation is used instead of a constant regularisation parameter. The performance of the UCL topography system and image reconstruction algorithm was quantified in terms of contrast, spatial resolution and spatial accuracy.

The published literature includes several previous descriptions of methods designed to find the optimal set of wavelengths for optical topography imaging of the adult brain. None of these studies examined the similarities between interrogated volumes of measurements performed at different wavelengths. Based on the multispectral reconstruction we introduced a method to select the optimal wavelengths, which not only gives good separation between chromophores, absorption and scattering, but also minimises the differences between interrogated volumes. This method uses the SSD to compare PMDFs, which were generated for wavelengths in the NIR range. The image reconstructions generated from simulated data slightly improved when measurements with overlapping sensitivities were used. Interestingly, we found that in CW optical topography using linear multispectral reconstruction that at least four measurement wavelengths are required in order to separate absorption from scattering and separate two chromophores from one another. Our method can be applied to other types of tissue.

Throughout this project we studied the limitations and potential of the UCL system and linear image reconstruction, which is essential to evaluate the reliability of images reconstructed from *in vivo* data. Finally, the new three-dimensional imaging tools were employed to reconstruct images from measurements of visually evoked response in the infant brain. To our knowledge these were the first optical topography images of regions of the social brain network, showing decreases in [HHb] and increases in the [HbO₂] in the temporal and prefrontal cortex. More precisely, we believe that activity occurred in regions of the STS, IFG and mPFC. These results are in agreement with PET and fMRI studies of the adult social brain. Optical topography appears to be the ideal imaging modality for infant brain imaging. Many advantages can be enumerated, such as its non-invasiveness, low-cost, short scanning times, portability, infants can be monitored while they are awake, relatively small sensitivity to motion, and the number of optodes and their configuration can be easily adjusted according to the types of study. Nevertheless, these are preliminary studies and we hope that in the future we can obtain better contrast, resolution and spatial accuracy. So far we have reconstructed images from optical topography data using the zero-order Tikhonov regularisation, which is known to produce smooth solutions. In the future, different penalty functions will be included in the reconstruction algorithm and their effect on the reconstructed images analysed. These functions include higher order Tikhonov functions, Total Variation (TV) and possibly a hybrid method

like the one described by Douiri *et al* [2007]. The inclusion of priors in the image reconstruction is briefly described below.

10.1 Use of priors in linear reconstruction

The generalised Tikhonov regularisation is given by [Vogel, 2002]

$$\|Ax - b\|_2^2 + \lambda^2 F(x) = \min \quad (10.1)$$

where the function $F(x)$ is a penalty function, which can incorporate *a priori* information. In the zero-order Tikhonov regularisation $F(x) = \|Dx\|_2^2$ with $D = I$ and the above becomes equation (5.87). The matrix D can have other forms, for example it can be a covariance matrix, as described further in §7, or it can be a derivative operator.

The special case of the Tikhonov regularisation described in equation (5.87), with $F(x) = \|x\|_2^2$, results in very smooth solutions. As mentioned previously the regularisation parameter λ has a major effect on the solution. If λ is too large the inverse problem is only slightly related to the original unregularised problem, whereas if λ is chosen to be small the solution will be dominated by amplified errors.

If some information is known about the solution and its behaviour, then $F(x)$ can incorporate this *a priori* information [Pedersen, 2005; Aster *et al.*, 2004]. If the solution is known to have a small first derivative then selecting $D = \nabla$, where ∇ is the gradient operator, should penalise any rapid changes in the solution x . It is possible to approximate the first derivative of x (in one dimension) by the multiplication D_1x , where D_1 is the matrix

$$D_1 = \begin{pmatrix} -1 & 1 & & & & & \\ & -1 & 1 & & & & \\ & & \ddots & \ddots & & & \\ & & & \ddots & \ddots & & \\ & & & & -1 & 1 & \end{pmatrix}.$$

In the second-order Tikhonov regularisation $D = \Delta$, where Δ is the Laplacian operator. The inclusion of this penalty function imposes a solution with a small second derivative, meaning that it penalises curvature. The second derivative of x can be approximated by D_2x , where D_2 is

$$D_2 = \begin{pmatrix} 1 & -2 & 1 & & & & \\ & 1 & -2 & 1 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & \ddots & \ddots & \ddots & \\ & & & & 1 & -2 & 1 \end{pmatrix}.$$

The Total Variation (TV) penalty function is used when there are discontinuous jumps in the solution. Like Tikhonov regularisation it also penalizes highly oscillating solutions [Vogel, 2002]. The TV function is defined as

$$TV(x) = \int_{\Omega} |\nabla x|, \quad (10.2)$$

where Ω is the solution domain. In the one dimensional case $F(x) = \|\nabla x\|_1$. The main difference between the Tikhonov regularisation penalty functions and the TV function is the norm; Tikhonov regularisation functions are L2 norms and TV is an L1 norm. The

Tikhonov regularisation penalises discontinuities in favour of smooth transitions, whereas TV regularisation allows discontinuous transitions, thereby preserving the edge information in the reconstructed images [Aster *et al.*, 2004].

A one-dimensional example is used to illustrate the advantage of the TV function. Consider a function $\Gamma = \{f : \mathbb{R} \rightarrow \mathbb{R}, f(0) = 0, f(b) = a\}$ and also the Ln norm which is defined as follows

$$\|f(x)\|_n = \left(\int |f(x)|^n dx \right)^{1/n}. \quad (10.3)$$

Solving for the n norm raised to the power n we obtain $\|f'(x)\|_n = a^n b^{1-n}$, hence $\|f'(x)\|_n^n = ab^{\frac{1-n}{n}}$. If $n = 1$ then $\|f'(x)\|_1 = a$, and for $n = 2$ then $\|f'(x)\|_2 = a/\sqrt{b}$ [Pedersen, 2005; Borsic, 2002]. From these results it can be concluded that the L1 norm is only influenced by variations, whereas for the L2 norm variations are penalised but also rapid changes (steep gradients) are penalised; it can be seen that the norm becomes large for a small b . In this example $\|f'(x)\|_1$ can be minimised by an infinite number of functions, and only a straight line that connects the origin to the point (b,a) can minimise $\|f'(x)\|_2$.

Hybrid methods that take advantage of both features of the Tikhonov regularisation and TV functions have been developed, and have proven to enhance the quality of optical images [Douiri *et al.*, 2007; Paulsen & Jiang, 1996]. Optical topography images exhibit poor spatial resolution due to the diffuse nature of the photon migration, but the quality of the images can be improved by the use of prior information. One possible example is to use MRI images to provide anatomical information [Ntziachristos *et al.*, 2000; Davis *et al.*, 2007; Brooksby *et al.*, 2003; Carpenter *et al.*, 2008].

The methods used to select the regularisation parameter described in §5 could be adapted so that we are able to find the regularisation parameters for each prior type, followed by an analysis of these methods similar to the one done for the zero-order Tikhonov regularisation. A study would then determine if the selection methods return valid regularisation parameters when priors are included in the inverse problem. The deblurring problem could be used again, as it represents a simpler problem than optical imaging and it is always easier to obtain reasonable results using generated data. Thereafter, the method could be assessed using experimental optical topography data.

Absorption and reduced scattering coefficients table

Table A.1 contains the absorption coefficients μ_a and reduced scattering coefficients μ'_s used in this thesis for the brain, skull and skin spectra.

$\lambda(\text{nm})$	BRAIN		SKIN		SKULL	
	absorption $\mu_a(\text{mm}^{-1})$	scattering $\mu_s(\text{mm}^{-1})$	absorption $\mu_a(\text{mm}^{-1})$	scattering $\mu_s(\text{mm}^{-1})$	absorption $\mu_a(\text{mm}^{-1})$	scattering $\mu_s(\text{mm}^{-1})$
650	0.0507	2.4559	0.0293	2.3889	0.0400	2.7000
660	0.0474	2.4361	0.0270	2.3464	0.0350	2.6000
670	0.0434	2.4168	0.0248	2.3068	0.0300	2.5500
680	0.0392	2.3979	0.0223	2.2696	0.0280	2.4500
690	0.0357	2.3794	0.0201	2.2348	0.0250	2.4000
700	0.0336	2.3613	0.0186	2.2021	0.0240	2.3500
710	0.0326	2.3437	0.0177	2.1714	0.0245	2.2500
720	0.0323	2.3263	0.0173	2.1425	0.0247	2.2000
730	0.0329	2.3094	0.0175	2.1153	0.0248	2.1600
740	0.0346	2.2928	0.0185	2.0896	0.0249	2.1100
750	0.0366	2.2766	0.0198	2.0654	0.0250	2.0800
760	0.0382	2.2606	0.0207	2.0424	0.0255	2.0500
770	0.0376	2.2450	0.0200	2.0207	0.0250	1.9500
780	0.0366	2.2297	0.0190	2.0001	0.0250	1.9200
790	0.0364	2.2147	0.0185	1.9806	0.0245	1.8500
800	0.0366	2.2000	0.0185	1.9621	0.0245	1.8000
810	0.0372	2.1856	0.0188	1.9444	0.0247	1.7800
820	0.0381	2.1714	0.0194	1.9276	0.0250	1.7600
830	0.0396	2.1575	0.0204	1.9116	0.0260	1.7300
840	0.0409	2.1438	0.0214	1.8964	0.0270	1.7200
850	0.0417	2.1304	0.0220	1.8818	0.0280	1.7000
860	0.0424	2.1173	0.0225	1.8679	0.0290	1.6700
870	0.0432	2.1043	0.0230	1.8546	0.0300	1.6500
880	0.0440	2.0916	0.0236	1.8418	0.0310	1.6000
890	0.0451	2.0791	0.0244	1.8296	0.0330	1.5400
900	0.0460	2.0669	0.0250	1.8178	0.0350	1.5200
910	0.0469	2.0548	0.0255	1.8066	0.0400	1.5000
920	0.0485	2.0429	0.0265	1.7958	0.0430	1.4500
930	0.0517	2.0313	0.0287	1.7854	0.0450	1.4000
940	0.0561	2.0198	0.0325	1.7754	0.0480	1.3500
950	0.0642	2.0085	0.0391	1.7657	0.0500	1.3000
960	0.0750	1.9974	0.0474	1.7564	0.0530	1.2800
970	0.0765	1.9864	0.0486	1.7475	0.0550	1.2500
980	0.0745	1.9757	0.0471	1.7388	0.0580	1.2000
990	0.0701	1.9651	0.0437	1.7305	0.0630	1.1500

Table A.1: Absorption and reduced scattering coefficients for skin, skull and brain tissue.

B.1 Spatially variant regularisation - summarised results and reconstructed images

The following tables summarise the results of §7.4.1. Tables B.1-B.4 show the results of experiment 1, for the simulations and liquid phantom, where the absorption perturbation was located at different depths. Images were reconstructed using the different regularisation types (described in §7.2) and the apparent position of the target was calculated using the weighted mean position (described in §7.3). Table B.1 shows the real target position and the apparent position of the target in the images. The root squared error (RSE) was used to calculate the global depth accuracy. Table B.2 shows the mean PSF width (x and z directions) and the RSE for each regularisation type. Table B.3 summarises the normalised FVHM and the RSE. Table B.4 shows the contrast values and the RSE for each regularisation type. The results in table B.5 correspond to experiment 2, where the target was moved along the x and y axis in the liquid phantom. Images were reconstructed using a constant regularisation parameter λ_c and a variable λ (NOSER regularisation). This table shows the true location of the target (x, y), the apparent position in the image and the lateral spatial accuracy SA_{xy} , estimated from the RSE. The PSF width, FVHM and contrast values are also shown.

λ type	DEPTH OF TARGET IN THE IMAGE											
	simulation (mm)					phantom T1 (mm)						
target depth	10	15	20	25	30	RSE	12	17	22	27	32	RSE
E	10.1	18.5	24.7	26.9	26.4	7.1	7.2	17.4	26.4	27.7	27.9	7.7
S	11.8	18.0	23.5	25.9	26.3	6.3	9.3	18.8	24.5	25.7	26.0	7.4
NOSER	10.8	18.5	22.7	24.1	25.3	6.4	10.9	17.4	22.3	25.9	25.2	7.2
NOSER & C	12.9	18.7	22.8	24.9	25.6	7.1	9.1	18.1	21.8	25.8	26.3	6.6
VU	12.2	18.9	22.7	23.8	24.7	7.8	9.9	18.1	23.5	24.6	25.0	7.8
R & C	10.2	18.5	22.8	24.5	25.7	6.3	12.2	17.7	20.4	25.0	25.4	7.1
R	10.9	17.9	22.6	24.1	25.3	6.9	7.1	14.6	21.2	24.2	24.3	9.9
λ_c	8.7	13.4	16.2	18.8	20.9	11.9	6.6	10.9	13.1	12.3	12.5	27.3

Table B.1: Weighted average depth for the simulation and liquid phantom. The uncertainty in depth corresponds to the pixel size, which is 3.3 mm. The RSE values are the global depth accuracy.

λ type	PSF WIDTH											
	simulation (mm)						phantom T1 (mm)					
target depth	10	15	20	25	30	RSE	12	17	22	27	32	RSE
E	2.4	11.0	14.3	17.1	18.3	30.9	3.3	14.3	15.2	16.1	14.8	30.5
S	4.5	6.5	9.5	7.4	10.7	17.9	6.2	12.5	8.9	10.0	10.7	22.2
NOSER	5.5	17.0	14.1	11.2	9.8	27.2	3.7	15.9	11.5	11.1	12.0	25.9
NOSER & C	3.4	11.4	12.5	10.9	11.1	23.2	8.4	18.5	16.2	18.9	20.7	38.2
VU	8.1	9.9	10.8	10.9	12.4	23.5	9.3	8.1	4.6	6.1	6.6	15.9
R & C	4.7	14.9	20.1	22.7	25.3	42.5	1.8	15.1	8.8	6.7	7.3	20.1
R	7.1	14.5	21.9	28.0	27.5	47.8	6.4	17.0	18.9	13.7	13.5	32.6
λ_c	4.2	8.8	10.9	12.8	14.8	24.5	7.7	2.9	3.5	5.6	6.1	12.2

Table B.2: PSF width obtained from the reconstructed images of the simulation and liquid phantom. The uncertainty is given by the pixel size, which has a mean value of 3.4 mm. The RSE values represent the global PSF width.

λ type	NORMALISED FVHM											
	simulation (mm)						phantom T1 (mm)					
target depth	10	15	20	25	30	RSE	12	17	22	27	32	RSE
E	0.009	0.022	0.033	0.044	0.050	0.071	0.016	0.058	0.030	0.038	0.043	0.081
S	0.009	0.014	0.025	0.022	0.028	0.040	0.019	0.033	0.022	0.030	0.030	0.053
NOSER	0.016	0.035	0.038	0.038	0.038	0.068	0.016	0.038	0.030	0.038	0.043	0.081
NOSER & C	0.009	0.028	0.032	0.030	0.028	0.052	0.003	0.052	0.047	0.074	0.073	0.118
VU	0.032	0.032	0.027	0.030	0.033	0.061	0.019	0.035	0.011	0.019	0.017	0.041
R & C	0.013	0.041	0.091	0.104	0.101	0.170	0.013	0.046	0.017	0.027	0.025	0.055
R	0.009	0.036	0.095	0.106	0.106	0.174	0.005	0.035	0.050	0.047	0.052	0.086
λ_c	0.013	0.014	0.022	0.032	0.039	0.051	0.008	0.014	0.017	0.009	0.009	0.020

Table B.3: Normalised FVHM. The uncertainty is 0.0001. The RSE values represent the FVHM for all depths.

λ type	CONTRAST											
	simulation (mm)						phantom T1 (mm)					
target depth	10	15	20	25	30	RSE	12	17	22	27	32	RSE
E	3.1	1.2	1.0	0.33	0.21	2.8	0.066	0.019	0.014	0.011	0.022	1.6
S	± 0.3	± 0.1	± 0.1	± 0.01	± 0.01	± 0.5	± 0.005	± 0.004	± 0.001	± 0.001	± 0.003	± 0.01
NOSER	4.1	2.0	1.6	0.63	0.43	4.3	0.066	0.045	0.042	0.047	0.062	1.6
NOSER & C	± 0.4	± 0.2	± 0.1	± 0.04	± 0.02	± 0.7	± 0.004	± 0.005	± 0.004	± 0.005	± 0.006	± 0.02
VU	1.9	0.8	0.67	0.27	0.11	1.4	0.064	0.032	0.022	0.026	0.028	1.6
R & C	± 0.1	± 0.1	± 0.02	± 0.01	± 0.02	± 0.3	± 0.006	± 0.004	± 0.002	± 0.003	± 0.005	± 0.02
R	2.6	1.9	0.67	0.15	0.11	2.2	1.2	1.0	1.0	0.59	0.25	0.8
λ_c	± 0.2	± 0.1	± 0.03	± 0.01	± 0.03	± 0.4	± 0.1	± 0.1	± 0.1	± 0.03	± 0.07	± 0.4
E	1.9	0.9	0.67	0.19	0.11	1.6	0.027	0.012	0.0069	0.0052	0.013	1.6
S	± 0.01	± 0.1	± 0.03	± 0.01	± 0.02	± 0.2	± 0.001	± 0.008	± 0.0005	± 0.0002	± 0.009	± 0.03
NOSER	1.4	0.35	0.15	0.062	0.021	0.8	0.063	0.008	0.008	0.0073	0.0087	1.6
NOSER & C	± 0.01	± 0.03	± 0.01	± 0.004	± 0.004	± 0.1	± 0.003	± 0.007	± 0.006	± 0.0005	± 0.0005	± 0.02
VU	1.8	0.33	0.014	0.050	0.019	1.1	0.083	0.041	0.024	0.025	0.029	1.6
R & C	± 0.01	± 0.02	± 0.003	± 0.002	± 0.01	± 0.1	± 0.005	± 0.008	± 0.001	± 0.004	± 0.001	± 0.02
R	3.1	1.4	1.1	0.78	0.16	3.1	0.32	0.78	0.68	0.83	1.6	0.9
λ_c	± 0.1	± 0.1	± 0.1	± 0.01	± 0.01	± 0.3	0.02	± 0.05	0.05	± 0.06	± 0.3	± 0.5

Table B.4: Contrast of the target in the reconstructed images from simulation and phantom data. The RSE values are the global contrast.

(x,y)	constant λ						variable λ					
	x (mm)	y (mm)	SA_{xy} (mm)	PSF (mm)	FVHM	Contrast	x (mm)	y (mm)	SA_{xy} (mm)	PSF (mm)	FVHM	Contrast
(-10,-10)	-16.7	-13.5	7.6	4.3	0.003	0.48 ± 0.04	-11.3	-11.2	1.8	7.3	0.003	0.40 ± 0.05
(-10,0)	-16.3	-3.6	7.2	6.4	0.007	0.32 ± 0.04	-14.7	-3.8	6.1	6.6	0.005	0.30 ± 0.03
(-10,10)	-15.9	3.3	8.9	3.4	0.002	0.99 ± 0.18	-12.9	7.9	3.7	16.7	0.005	0.40 ± 0.04
(0,-10)	-5.1	-11.4	5.3	6.8	0.005	0.23 ± 0.02	-3.2	-7.6	4.0	4.3	0.003	0.32 ± 0.03
(0,0)	-3.8	-1.3	4.0	15.7	0.011	0.19 ± 0.02	-2.8	0.3	2.8	20.6	0.006	0.29 ± 0.03
(0,10)	-5.4	-7.12	6.0	16.7	0.006	0.38 ± 0.06	-4.6	9.1	4.6	16.1	0.006	0.43 ± 0.05
(10,-10)	5.2	-9.8	4.8	13.8	0.010	0.11 ± 0.01	4.9	-7.8	5.5	17.7	0.011	0.25 ± 0.01
(10,0)	-0.8	4.2	5.9	13.8	0.008	0.28 ± 0.03	4.2	0.3	5.8	14.9	0.007	0.24 ± 0.03
(10,10)	3.7	9.2	6.4	7.5	0.003	0.30 ± 0.03	4.6	10.1	5.4	21.0	0.011	0.34 ± 0.05

Table B.5: Apparent position (x and y) of the target in the images, spatial accuracy SA_{xy} , PSF width, FVHM and contrast.

B.2 Reconstructed images

The following images were reconstructed using the different regularisation types from simulated and liquid phantom data. First, images reconstructed from simulated data are shown, followed by the phantom images. The images are zx cross-sections, the distances from the surface are indicated for the z direction. Pixel dimensions are $3.3 \text{ mm} \times 3.5 \text{ mm}$. Figure captions are self-explanatory.

B.2.1 Simulations

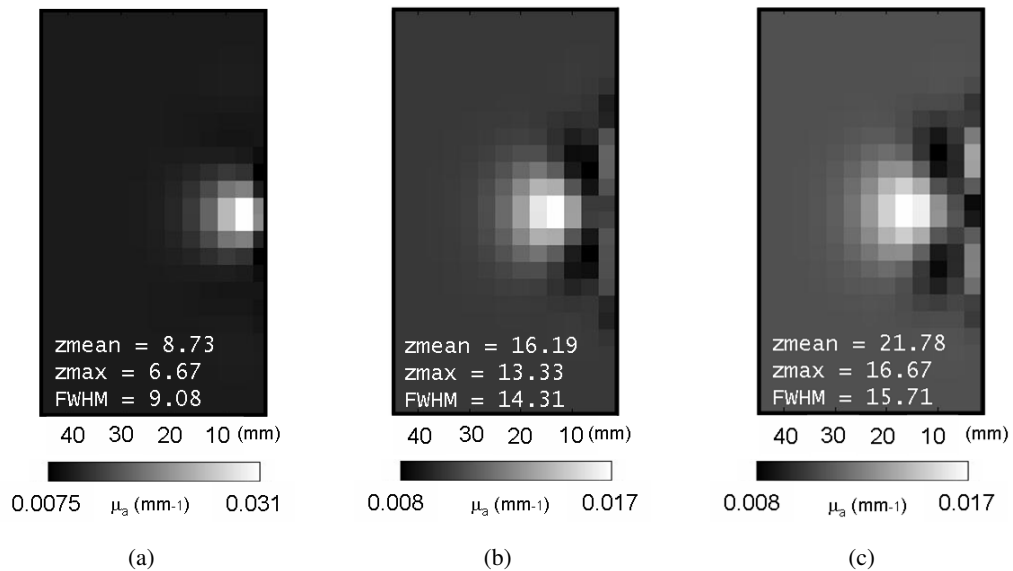


Figure B.1: Images reconstructed from simulated data using a constant regularisation parameter λ_c , for absorption perturbations located at (a) 10 mm, (b) 20 mm and (c) 30 mm below the surface.

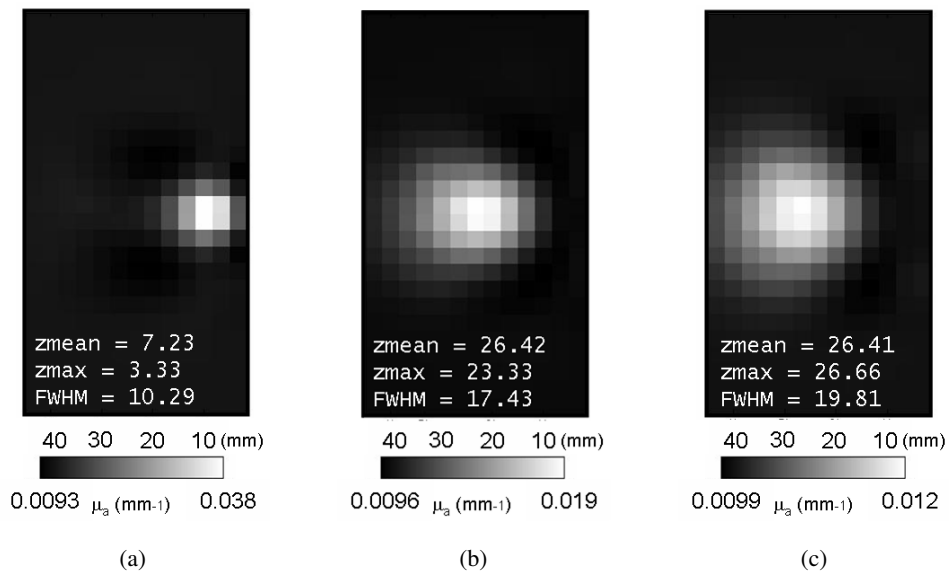


Figure B.2: Images reconstructed from simulated data using an exponential decay function (E) as regularisation parameter, for absorption perturbations located at (a) 10 mm, (b) 20 mm and (c) 30 mm below the surface.

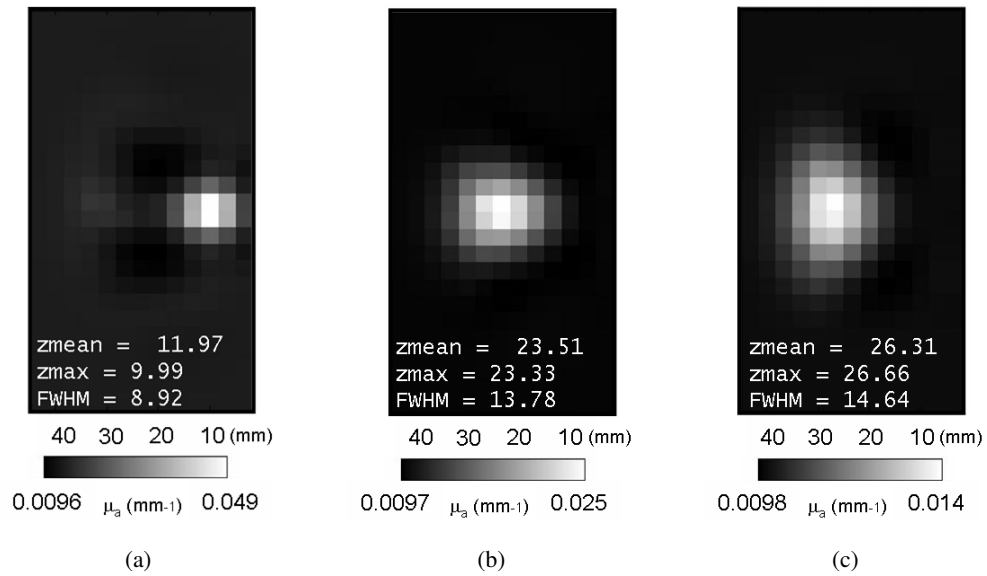


Figure B.3: Images reconstructed from simulated data using a sigmoid decay function (S) as regularisation parameter, for absorption perturbations located at (a) 10 mm, (b) 20 mm and (c) 30 mm below the surface.

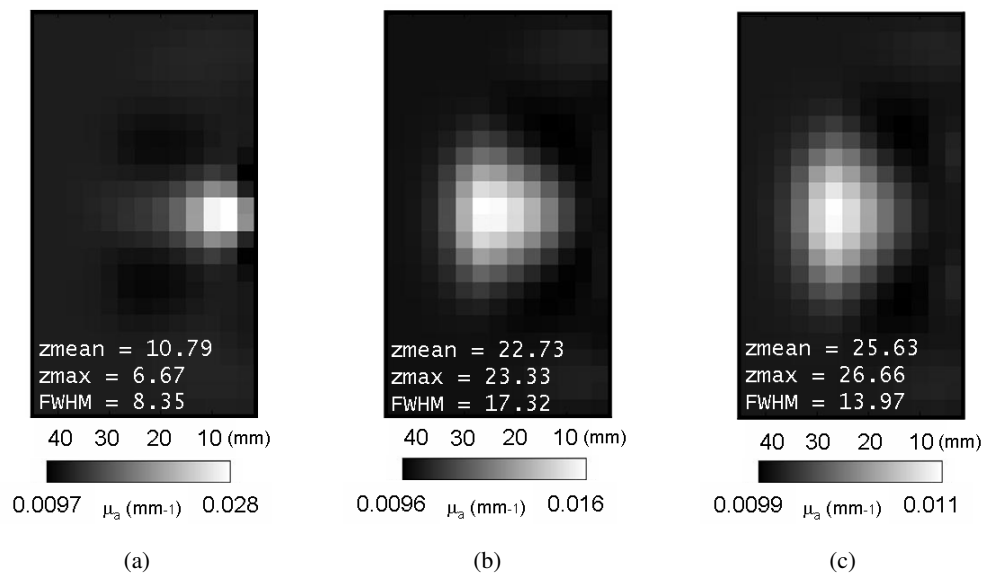


Figure B.4: Images reconstructed from simulated data using the NOSER type regularisation parameter, for absorption perturbations located at (a) 10 mm, (b) 20 mm and (c) 30 mm below the surface.

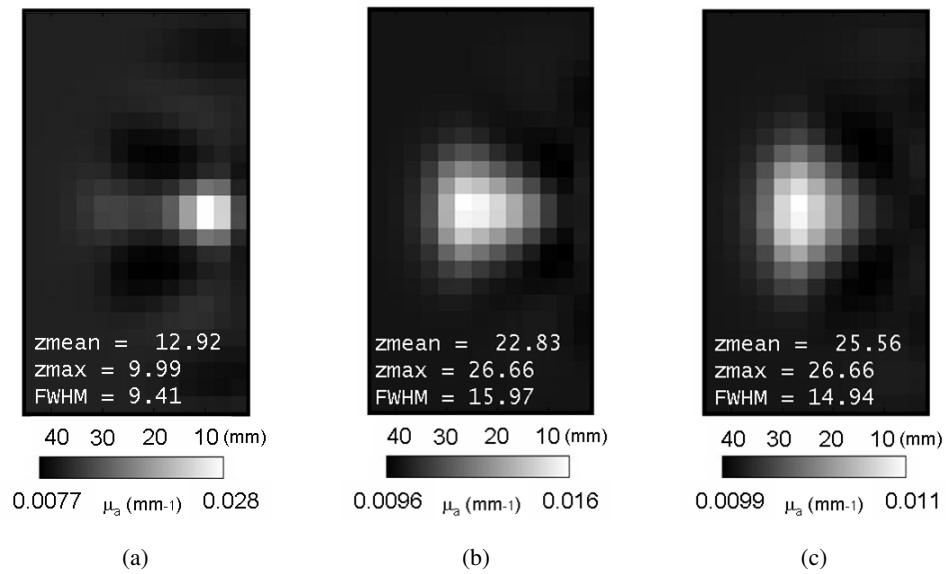


Figure B.5: Images reconstructed from simulated data using the NOSER type regularisation parameter and data covariance matrix C , for absorption perturbations located at (a) 10 mm, (b) 20 mm and (c) 30 mm below the surface.

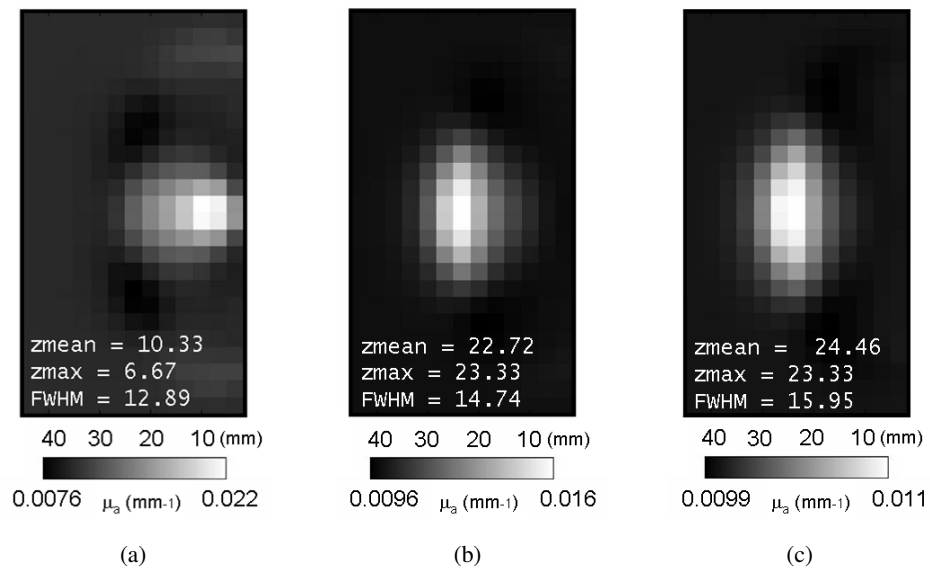


Figure B.6: Images reconstructed from simulated data using the variance uniformisation (VU) regularisation parameter, for absorption perturbations located at (a) 10 mm, (b) 20 mm and (c) 30 mm below the surface.

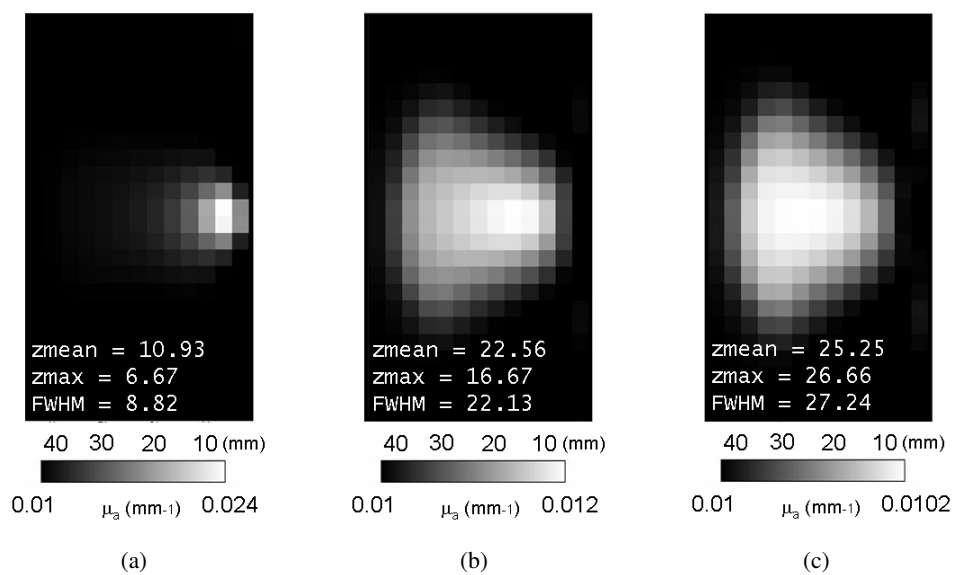


Figure B.7: Images reconstructed from simulated data using the covariance matrix R as regularisation parameter, for absorption perturbations located at (a) 10 mm, (b) 20 mm and (c) 30 mm below the surface.

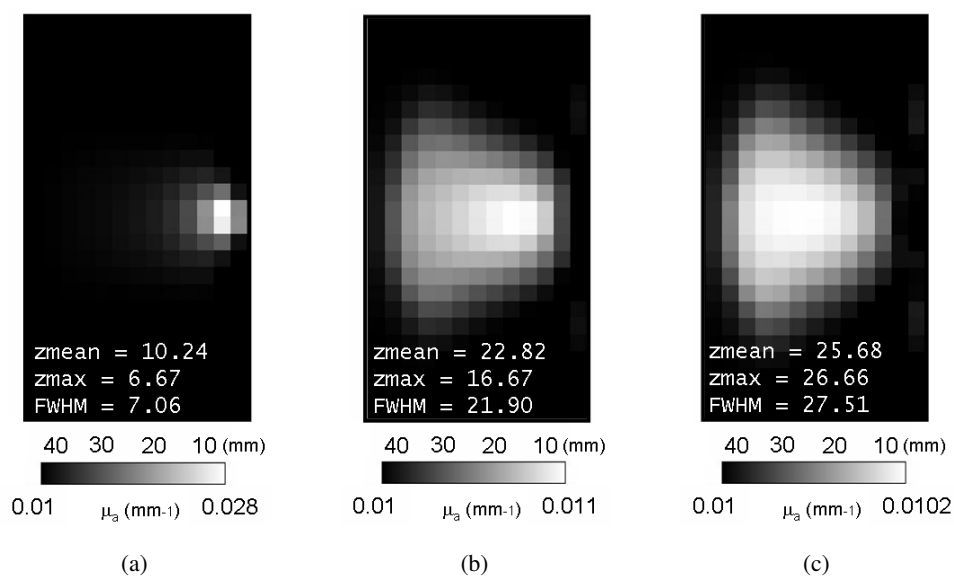


Figure B.8: Images reconstructed from simulated data using the covariance matrices R and C as regularisation parameter, for absorption perturbations located at (a) 10 mm, (b) 20 mm and (c) 30 mm below the surface.

B.2.2 Liquid phantom

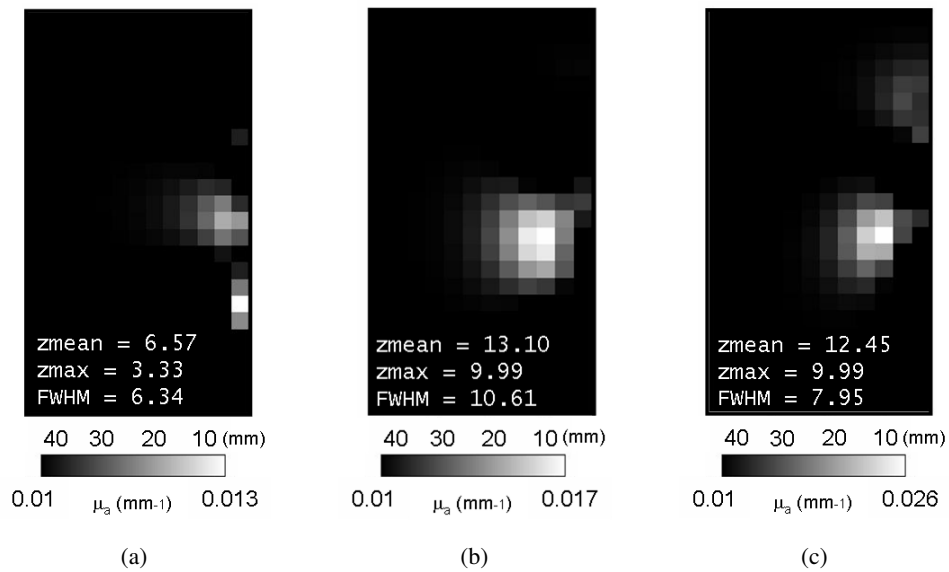


Figure B.9: Images reconstructed from phantom data using a constant regularisation parameter λ_c , for absorption perturbations located at (a) 12 mm, (b) 22 mm and (c) 32 mm below the surface.

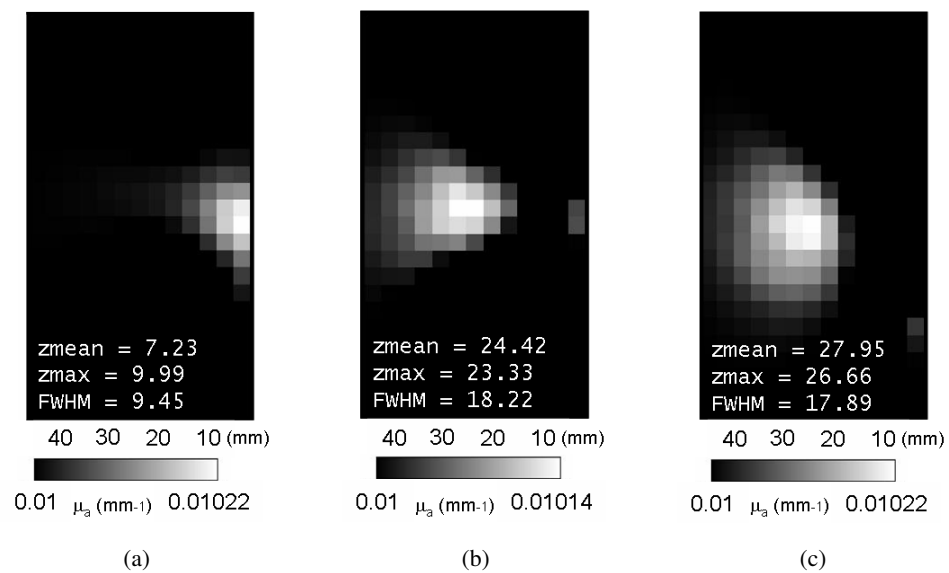


Figure B.10: Images reconstructed from phantom data using an exponential decay function (E) as regularisation parameter, for absorption perturbations located at (a) 12 mm, (b) 22 mm and (c) 32 mm below the surface.

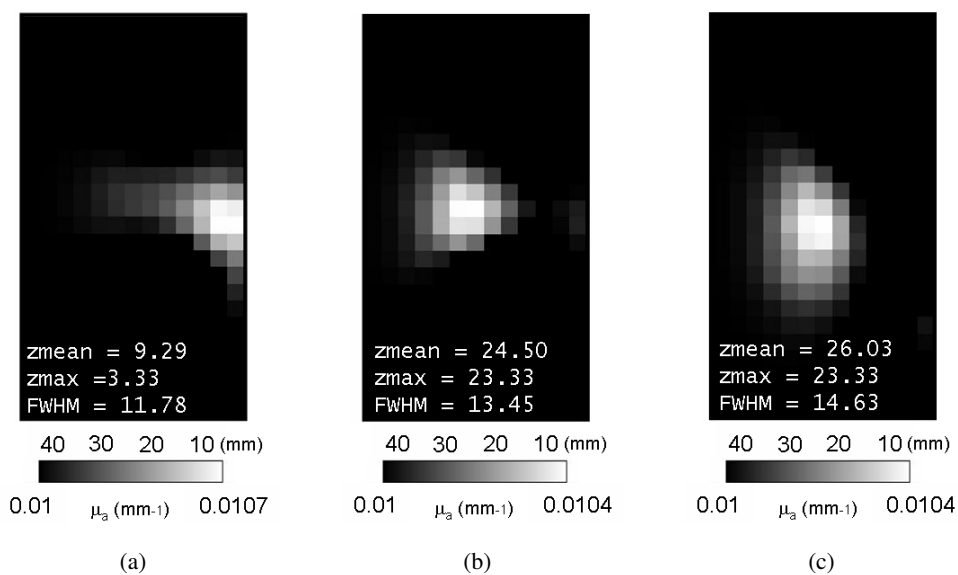


Figure B.11: Images reconstructed from phantom data using a sigmoid decay function (S) as regularisation parameter, for absorption perturbations located at (a) 12 mm, (b) 22 mm and (c) 32 mm below the surface.

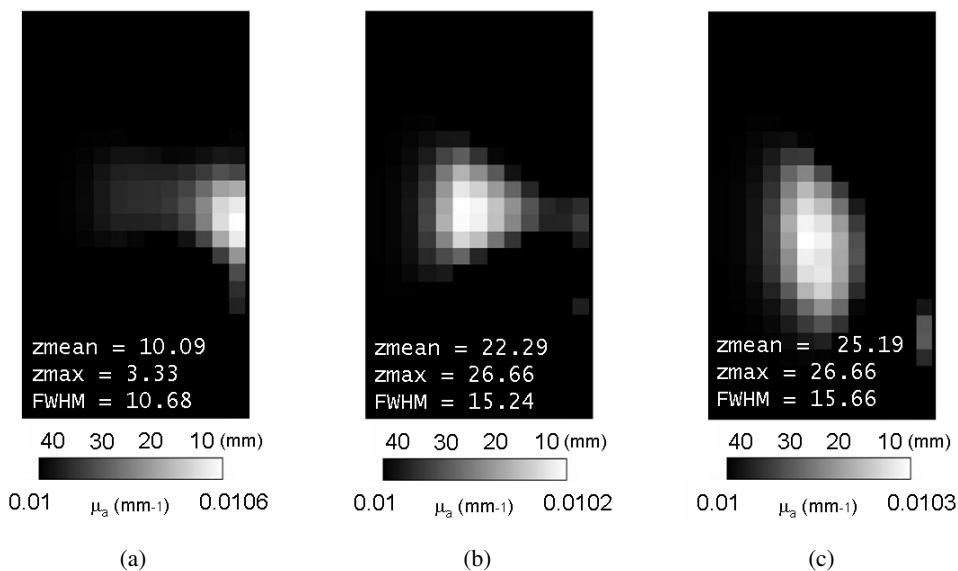


Figure B.12: Images reconstructed from phantom data using the NOSER type regularisation parameter, for absorption perturbations located at (a) 12 mm, (b) 22 mm and (c) 32 mm below the surface.

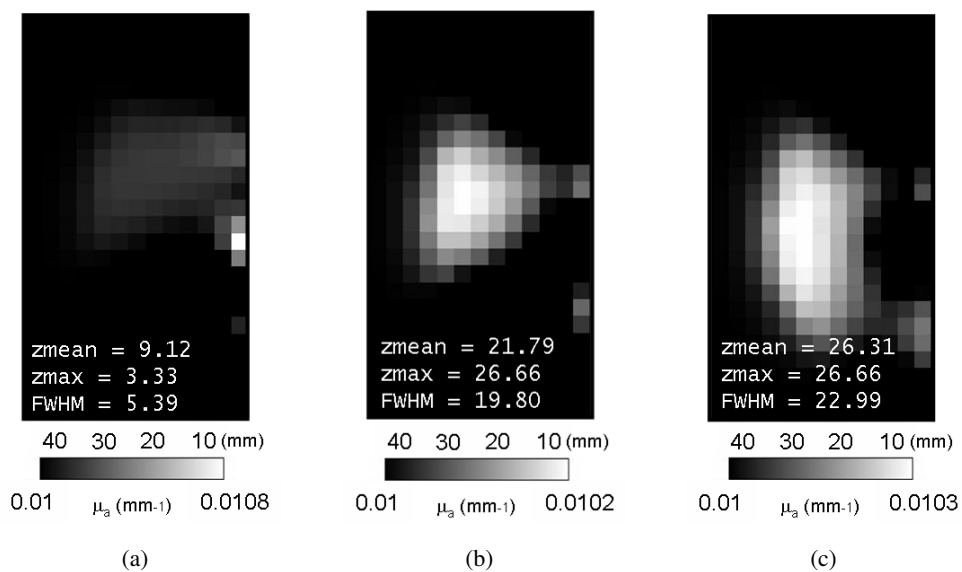


Figure B.13: Images reconstructed from phantom data using the NOSER type regularisation parameter and data covariance matrix C , for absorption perturbations located at (a) 12 mm, (b) 22 mm and (c) 32 mm below the surface.

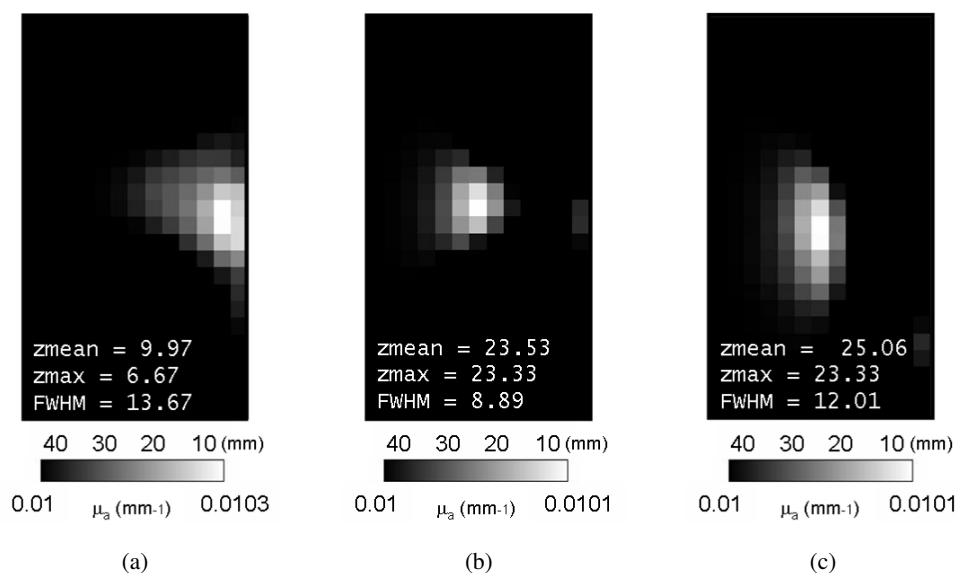


Figure B.14: Images reconstructed from phantom data using the variance uniformisation (VU) regularisation parameter, for absorption perturbations located at (a) 12 mm, (b) 22 mm and (c) 32 mm below the surface.

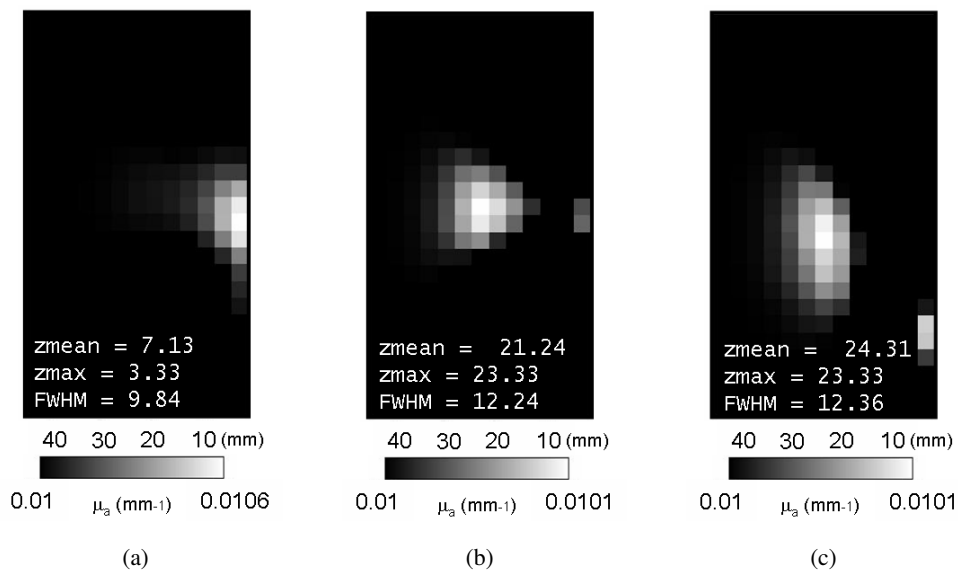


Figure B.15: Images reconstructed from phantom data using the covariance matrix R as regularisation parameter, for absorption perturbations located at (a) 12 mm, (b) 22 mm and (c) 32 mm below the surface.

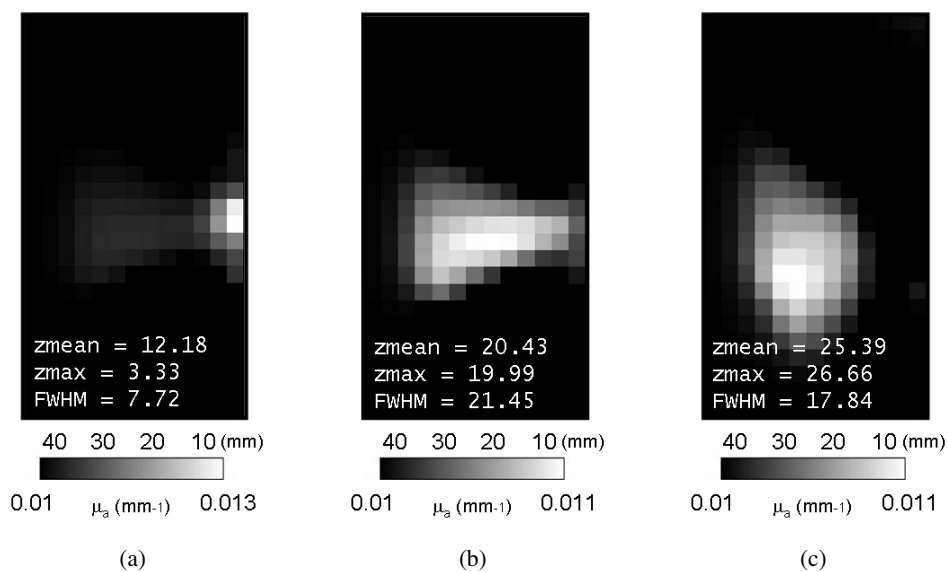


Figure B.16: Images reconstructed from phantom data using the covariance matrices R and C as regularisation parameter, for absorption perturbations located at (a) 12 mm, (b) 22 mm and (c) 32 mm below the surface.

Bibliography

- Abascal, J P, Arridge, S, Bayford, R, & Holder, D. 2008. Comparison of methods for optimal choice of the regularization parameter for linear electrical impedance tomography of brain function. *Physiological Measurement*, **29**(11), 1319–1334.
- Adler, A. 2004. Accounting for erroneous electrode data in electrical impedance tomography. *Physiological Measurement*, **25**, 227–238.
- Adler, A, & Guardo, R. 1996. Electrical Impedance Tomography: Regularized imaging and contrast detection. *IEEE Transactions on Medical Imaging*, **15**(2), 170–179.
- Adler, A, Dai, T, & Lionheart, WRB. 2007. Temporal image reconstruction in electrical impedance tomography. *Physiological Measurement*, **28**, S1–S11.
- Allison, T, Puce, A, & McCarthy, G. 2000. Social perception from visual cues: role of the STS region. *Trends in Cognitive Science*, **4**(7), 251–191.
- Ancel, P, Livinec, F, Larroque, B, Marret, S, Arnaud, C, Pierrat, V, Dehan, M, N'Guyen, S, Escande, B, Burguet, A, Thiriez, G, Picaud, J, Andre, M, Breart, G, Kaminski, M, & the EPIPAGE Study Group. 2006. Cerebral Palsy Among Very Preterm Children in Relation to Gestational Age and Neonatal Ultrasound Abnormalities: The EPIPAGE Cohort Study. *Pediatrics*, **117**(3), 828–835.
- Arridge, S R. 1997. Image reconstruction in optical tomography. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **352**, 717–726.
- Arridge, S R, & Schweiger, M. 1993. Inverse Methods for Optical Tomography. *Chap. Inverse methods for optical tomography, pages 259–277 of: Proceedings of the 13th International Conference on Information Processing in Medical Imaging*, vol. 687. Springer Berlin / Heidelberg.
- Arridge, S R, Cope, M, & Delpy, D T. 1992. The theoretical basis for the determination of optical pathlengths in tissue: temporal and frequency analysis. *Physics in Medicine and Biology*, **37**(7), 1531–1560.
- Arridge, S R, Dehghani, H, Schweiger, Martin, & Okada, E. 2000. The finite element model for the propagation of light in scattering media: A direct method for domains with nonscattering regions. *Medical Physics*, **27**(1), 251–264.
- Arridge, SR. 1995. Photon measurement density functions. Part1: Analytical forms. *Applied Optics*, **34**, 7395–7409.

- Arridge, SR. 1999. Optical tomography in medical imaging. *Inverse Problems*, **15**(2), R41–R93.
- Arridge, SR, & Hebden, JC. 1997. Optical imaging in medicine: II. Modelling and reconstruction. *Physics in Medicine and Biology*, **42**, 80712–80719.
- Arridge, SR, & Lionheart, WRB. 1998. Nonuniqueness in diffusion-based optical tomography. *Optics Letters*, **23**, 882–884.
- Arridge, SR, & Schotland, JC. 2009. Optical tomography: forward and inverse problems. *Inverse Problems*, **25**(12), 123010 (59pp).
- Arridge, SR, & Schweiger, M. 1995. Photon measurement density functions. Part2: Finite element method calculations. *Applied Optics*, **34**, 8026–8037.
- Aster, R, Borchers, B, & Thurber, C. 2004. *Parameter Estimation and Inverse Problems*. Academic Press.
- Avis, NJ, & Barber, DC. 1994. Image reconstruction using non-adjacent drive configurations (Electric Impedance Tomography). *Physiological Measurement*, **15**(2A), A153–A160.
- Bagshaw, A P, Liston, A D, Bayford, R H, Tizzard, A, Gibson, A P, Tidswell, A T, Sparkes, M K, Dehghani, H, Binnie, C D, & Holdera, D S. 2003. Electrical impedance tomography of human brain function using reconstruction algorithms based on the finite element method. *NeuroImage*, **20**, 752–764.
- Bashkatov, A N, Genina, EA, Kochubey, VI, & Tuchin, VV. 2005. Optical properties of human skin, subcutaneous and mucous tissues in the wavelength range from 400 to 2000 nm. *Journal of Physics D: Applied Physics*, **38**, 2543–2555.
- Bevilacqua, F, Piguët, D, Marquet, P, Gross, JD, Tromberg, BJ, & Depeursinge, C. 1999. *In Vivo* Local Determination of Tissue Optical Properties: Applications to Human Brain. *Applied Optics*, **38**(22), 4939–4950.
- Blakemore, S J. 2008. The social brain in adolescence. *Nature Reviews Neuroscience*, **9**, 267–277.
- Blasi, A, Fox, S, Everdell, N, Volein, A, Tucker, L, Csibra, G, Gibson, AP, Hebden, JC, Johnson, MH, & Elwell, CE. 2007. Investigation of depth dependent changes in cerebral haemodynamics during face perception in infants. *Physics in Medicine and Biology*, **53**(27), 6849–6864.
- Bluestone, AY, Abdoulaev, G, Schmitz, CH, Barbour, RL, & Hielscher, AH. 2001. Three-dimensional optical tomography of hemodynamics in the human head. *Optics Express*, **9**(6), 272–286.
- Bluestone, AY, Stewart, M, Lasker, J, Abdoulaev, GS, & Hielscher, AH. 2004. Three-dimensional optical tomographic brain imaging in small animals, part 1: hypercapnia. *Journal of Biomedical Optics*, **9**(5), 1046–1062.
- Boas, DA, & Dale, AM. 2005. Simulation study of magnetic resonance imaging-guided cortically constrained diffuse optical tomography of human brain function. *Applied Optics*, **44**(10), 1957–1968.
- Boas, DA, Gaudette, T, Strangman, G, Cheng, X, Marota, JJA, & Mandeville, JB. 2001a. The Accuracy of Near Infrared Spectroscopy and Imaging during Focal Changes in Cerebral Hemodynamics. *NeuroImage*, **13**, 76–90.

- Boas, DA, Brooks, DA, Miller, EL, DiMarzio, CA, Kilmer, M, Gaudette, JR, & Zang, Q. 2001b. Imaging the body with diffuse optical tomography. *IEEE Signal Processing Magazine*, **18**(6), 57–75.
- Boas, DA, Gaudette, T, & Arridge, SR. 2001c. Simultaneous imaging and optode calibration with diffuse optical tomography. *Optics Express*, **8**(5), 263–270.
- Boas, DA, Dale, AM, & Franceschini, MA. 2004a. Diffuse optical imaging of brain activation: approaches to optimizing image sensitivity, resolution, and accuracy. *NeuroImage*, **23**, S275–S288.
- Boas, DA, Chen, K, Grebert, D, & Franceschini, MA. 2004b. Improving the diffuse optical imaging spatial resolution of the cerebral hemodynamic response to brain activation in humans. *Optics Letters*, **29**(13), 1506–1508.
- Bolin, FP, Preuss, LE, Taylor, RC, & Ference, RJ. 1989. Refractive index of some mammalian tissues using a fiber optic cladding method. *Applied Optics*, **28**(12), 2297–2303.
- Borsic, A. 2002. *Regularization Methods for Imaging from Electrical Measurements*. Ph.D. thesis, Oxford Brookes University.
- Branco, G. 2007. *The development and evaluation of head probes for optical imaging of the infant head*. Ph.D. thesis, University College London.
- Brendel, B, & Nielsen, T. 2007. Wavelengths optimization in multi spectral diffuse optical tomography considering uncertainties in absorption spectra. *Proceeding of the SPIE - Diffuse Optical Imaging of Tissue*, **6629**, 66290A.
- Brockwell, P J, & Davis, R A. 1998. *Time Series: Theory and Methods*. 2nd edition edn. New York, USA: Springer Series in Statistics.
- Bronzino, JD. 2000. *The biomedical engineering handbook*. 2nd edition edn. Boca Raton, FL: CRC Press.
- Brooksby, BA, Dehghani, H, Pogue, BW, & Paulsen, KD. 2003. Near-Infrared (NIR) tomography breast image reconstruction with *a priori* structural information from MRI: algorithm development for reconstructing heterogeneities. *IEEE Journal of Selected Topics in Quantum Electronics*, **9**(2), 199–209.
- Brott, T, Broderick, J, Kothari, R, Barsan, W, Tomsick, T, Sauerbeck, L, Spilker, J, Duldner, J, & Khoury, J. 1997. Early hemorrhage growth in patients with intracerebral hemorrhage. *Stroke*, **28**(1), 1–5.
- Buxton, RB. 2002. *Introduction to functional magnetic resonance imaging: principles and techniques*. Cambridge University Press.
- Carlsson, J, Lagercrantz, H, Olson, L, Printz, G, & Bartocci, M. 2008. Activation of the right fronto-temporal cortex during maternal facial recognition in young infants. *Acta Paediatrica*, **97**, 1221–1225.
- Carpenter, CM, Srinivasan, S, Pogue, BW, & Paulsen, KD. 2008. Methodology development for three-dimensional MR-guided near infrared spectroscopy of breast tumors. *Optics Express*, **16**(22), 17903–17914.
- Chance, B, Anday, E, Nioka, S, Zhou, S, Hong, L, Worden, K, Li, C, Murray, T, Ovetsky, Y, & Thomas, R. 1998. A novel method for fast imaging of brain function, non-invasively, with light. *Optics Express*, **10**(2), 411–423.

- Cheney, M, Isaacson, D, Newell, JC, Simske, S, & Goble, JC. 1990. NOSER: an algorithm for solving the inverse conductivity problem. *International Journal of Imaging Systems and Technology*, **2**, 66–75.
- Cheng, X, & Boas, DA. 1998. Diffuse optical reflection tomography with continuous-wave illumination. *Optics Express*, **3**(3), 118–123.
- Cheong, WF, Prah, SA, & Welch, AJ. 1990. A review of the optical properties of biological tissues. *IEEE Journal of Quantum Electronics*, **26**, 2166–2185.
- Cherry, SM, Sorenson, JA, & Phelps, M. 2003. *Physics in Nuclear Medicine*. 3rd edition edn. Philadelphia, Pa: Saunders.
- Chiras, D D. 2005. *Human Biology*. 5th edition edn. Sudbury, MA, USA: Jones & Bartlett Publishers.
- Choi, J, Wolf, M, Toronov, V, Wolf, U, CPolzonetti, Hueber, D, Safonova, LP, Gupta, R, Michalos, A, Mantulin, W, & Gratton, E. 2004. Noninvasive determination of the optical properties of adult brain: near-infrared spectroscopy approach. *Journal of Biomedical Optics*, **9**(1), 221–229.
- Cohen-Bacrie, C, Goussard, Y, & Guardo, R. 1997. Regularized reconstruction in electrical impedance tomography using a variance uniformization constraint. *IEEE Transactions on Medical Imaging*, **16**(5), 562–571.
- Conry, BG, & Reznick, RH. 1987. The clinical application of computed x-ray tomography in tumor identification. *Cancer Surveys*, **6**(2), 379–396.
- Cooper, CE, & Springett, R. 1997. Measurement of cytochrome oxidase and mitochondrial energies by near-infrared spectroscopy. *Philosophical Transactions of the Royal Society of London. Series B*, **352**, 669–676.
- Cope, M. 1991. *The application of near infrared spectroscopy to non invasive monitoring of cerebral oxygenation in the newborn infant*. Ph.D. thesis, University College London, Department of Medical Physics and Bioengineering, London.
- Cope, M, & Delpy, D T. 1988. System for long-term measurement of cerebral blood and tissue oxygenation on newborn infants by near infra-red transillumination. *Medical and Biological Engineering and Computing*, **26**(3), 289–294.
- Corlu, A, Choe, R, Durduran, T, Lee, K, Schweiger, M, Arridge, SR, Hillman, EMC, & Yodh, AG. 2005. Diffuse optical tomography with spectral constraints and wavelength optimization. *Applied Optics*, **44**(11), 2082–2093.
- Correia, T, Banga, A, Everdell, NL, Gibson, AP, & Hebden, JC. 2009a. A quantitative assessment of the depth sensitivity of an optical topography system using a solid dynamic tissue-phantom. *Physics in Medicine and Biology*, **54**, 6277–6286.
- Correia, T, Gibson, A, Schweiger, M, & Hebden, J. 2009b. Selection of regularization parameter for optical topography. *Journal of Biomedical Optics*, **14**, 034044.
- Craven, P, & Wahba, G. 1979. Smoothing noisy data with spline functions (Estimating the correct degree of smoothing by the method of Generalized Cross-Validation). *Numerische Mathematik*, **24**(5), 377–403.

- Cubeddu, R, D'Andrea, C, Pifferi, A, Taroni, P, Torricelli, A, & Valentini, G. 2000. Effects of the menstrual cycle on the red and near-infrared optical properties of the human breast. *Photochemistry and Photobiology*, **72**(3), 383–391.
- Culver, JP, Choe, R, Holboke, MJ, Zubkov, L, Durduran, T, Slemph, A, Ntziachristos, V, Chance, B, & Yodh, AG. 2003. Three-dimensional diffuse optical tomography in the parallel plane transmission geometry: Evaluation of a hybrid frequency domain/continuous wave clinical system for breast imaging. *Medical Physics*, **30**(2), 235–247.
- Culver, JP, Siegel, AM, Franceschini, MA, Mandeville, JB, & Boas, DA. 2005. Evidence that cerebral blood volume can provide brain activation maps with better spatial resolution than deoxygenated hemoglobin. *NeuroImage*, **27**, 947–959.
- Custo, A, Boas, DA, Tsuzuki, D, Dan, I, Mesquita, R, Fischl, B, Grimson, WEL, & Wells, W. 2010. Anatomical atlas-guided diffuse optical tomography of brain activation. *NeuroImage*, **49**, 561–567.
- Davis, SC, Dehghani, H, Wang, H, Jiang, S, Pogue, BW, & Paulsen, KD. 2007. Image-guided diffuse optical fluorescence tomography implemented with Laplacian-type regularization. *Optics Express*, **15**(7), 4066–4082.
- Dehghani, H, & Delpy, DT. 2000. Near-infrared spectroscopy of the adult head: effect of scattering and absorbing obstructions in the cerebrospinal fluid layer on light distribution in the tissue. *Applied Optics*, **39**(25), 4721–4729.
- Dehghani, H, Arridge, S R, Schweiger, Martin, & Delpy, D T. 2000. Optical tomography in the presence of void regions. *Journal of Optical the Optical Society of America A*, **17**(9), 1659–1670.
- Dehghani, H, White, B R, Zeff, B W, Tizzard, A, & Culver, J P. 2009. Depth sensitivity and image reconstruction analysis of dense imaging arrays for mapping brain function with diffuse optical tomography. *Applied Optics*, **48**(10), D137–D143.
- Delpy, DT, & Cope, M. 1997. Quantification in tissue near-infrared spectroscopy. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **352**, 649–659.
- Douiri, A, Schweiger, M, Riley, J, & Arridge, SR. 2007. Anisotropic diffusion regularization methods for diffuse optical tomography using edge prior information. *Measurement Science and Technology*, **18**, 87–95.
- Dubois, J, Benders, M, Cachia, A, Lazeyras, F, Leuchter, R, Sizonenko, SV, Borradori-Tolsa, C, Mangin, JF, & Hüppi, PS. 2008. Mapping the early cortical folding process in the preterm newborn brain. *Cerebral Cortex*, **18**, 1444–1454.
- Eames, M, Wang, J, Pogue, B W, & Dehghani, H. 2008. Wavelength band optimization in spectral near-infrared optical tomography improves accuracy while reducing data acquisition and computational burden. *Journal of Biomedical Optics*, **13**, 054037.
- Eames, ME, Pogue, BW, Yalavarthy, PK, & Dehghani, H. 2007. An efficient Jacobian reduction method for diffuse optical image reconstruction. *Optics Express*, **15**(24), 15908–15919.
- Egan, WG, & Hilgeman, TW. 1979. *Optical properties of inhomogeneous materials*. New York: Academic Press.
- Elwell, CE. 1995. *A practical users guide to near infrared spectroscopy*. London: Hamamatsu Photonics KK.

- Endoh, R, Fujii, M, & Nakayama, K. 2005. The Simulated Detection of the Objects at Variant Depth Using Depth-Adaptive Voxel Size Scheme in Diffuse Reflective Optical Tomography (DROT). *Optical and Quantum Electronics*, **37**(13), 1409–1413.
- Endoh, R, Fujii, M, & Nakayama, K. 2008. Depth-adaptive regularized reconstruction for reflection diffuse optical tomography. *Optical Review*, **15**(1), 51–56.
- Everdell, NL, Gibson, AP, Tullis, IDC, Hebden, JC, & Delpy, DT. 2005. A frequency multiplexed near infrared topography system for imaging functional activation in the brain. *Review of Scientific Instruments*, **76**, 093705.
- Fabrizi, L, Sparkes, M, Horesh, L, abascal, JF, McEwan, A, Bayford, RH, Elwes, R, Binnie, CD, & Holder, DS. 2006. Factors limiting the application of electrical impedance tomography for identification of regional conductivity changes using scalp electrodes during epileptic seizures in humans. *Physiological Measurement*, **27**(5), S163–S174.
- Firbank, M, Hiraoka, M, Essenpreis, M, & Delpy, DT. 1993. Measurement of the optical properties of the skull in the wavelength range 650-950 nm. *Physics in Medicine and Biology*, **38**, 503–510.
- Fox, P T, & Raichle, M E. 1986. Focal physiological uncoupling of cerebral blood flow and oxidative metabolism during somatosensory stimulation in human subjects. *Proceedings of the National Academy of Sciences of the United States of America*, **83**(4), 1140–1144.
- Franceschini, MA, & Boas, DA. 2004. Noninvasive measurement of neuronal activity with near-infrared optical imaging. *NeuroImage*, **21**, 372–386.
- Franceschini, MA, Toronov, V, Filiaci, M, Gratton, E, & Fantini, S. 2000. On-line optical imaging of the human brain with 160-ms temporal resolution. *Optics Express*, **6**(3), 49–57.
- Franceschini, MA, Fantini, A, Thompson, JH, Culver, JP, & Boas, DA. 2003. Hemodynamic evoked response of the sensorimotor cortex measured noninvasively with near-infrared optical imaging. *Psychophysiology*, **40**, 548–560.
- Franceschini, MA, Joseph, DK, & Boas, TJ Huppert SG Diamond DA. 2006. Diffuse optical imaging of the whole head. *Journal of Biomedical Optics*, **11**(5), 054007.
- Fukui, Y, Ajichi, Y, & Okada, E. 2003. Monte Carlo prediction of near-infrared propagation in realistic adult and neonatal head models. *Applied Optics*, **42**(16), 2881–2887.
- Gaudette, R J, Boas, D A, Brooks, D H, DiMarzio, C A, Kilmer, Misha E, & Miller, Eric L. 2000. A comparison study of linear reconstruction techniques for diffuse optical tomographic imaging of absorption coefficient. *Physics in Medicine and Biology*, **45**(4), 1051–1070.
- Gebhart, SC. 2006. *In vitro* determination of normal and neoplastic human brain tissue optical properties using inverse adding-doubling. *Physics in Medicine and Biology*, **51**, 2011–2027.
- Gibson, AP, Hebden, JC, & Arridge, SR. 2005. Recent advances in diffuse optical imaging. *Physics in Medicine and Biology*, **50**, R1–R43.
- Gibson, AP, Austin, T, Everdell, NL, Schweiger, M, Arridge, SR, Meek, JH, Wyatt, JS, Delpy, DT, & Hebden, JC. 2006. Three-dimensional whole-head optical tomography of passive motor evoked responses in the neonate. *NeuroImage*, **30**(2), 521–528.
- Golub, G, Heath, M, & Wahba, G. 1979. Generalized-cross validation as a method for choosing a good ridge parameter. *Technometrics*, **21**(2), 215–223.

- Golub, GH. 1965. Numerical methods for solving least squares problems. *Numerische Mathematik*, **7**, 206–216.
- Graham, BM, & Adler, A. 2006. Objective selection of hyperparameter for EIT. *Physiological Measurement*, **27**, S65–S79.
- Graham, BM, & Adler, A. 2007. Electrode placement configuration for 3D EIT. *Physiological Measurement*, **28**, S1–S11.
- Grossmann, T, Johnson, MH, Lloyd-Fox, S, Blasi, A, Deligianni, F, Elwell, C, & Csibra, G. 2008. Early cortical specialization for face-to-face communication in human infants. *Proceedings of the Royal Society B: Biological Sciences*, **275**, 2803–2811.
- Guven, M, Yazici, B, Intes, X, & Chance, B. 2005. Diffuse optical tomography with *a priori* anatomical information. *Physics in Medicine and Biology*, **50**, 2837–2858.
- Hadamard, J. 1902. Sur les Problèmes aux Dérivées Partielles et Leur Signification Physique. *Princeton University Bulletin*, **13**, 49–52.
- Hanke, M. 1996. Limitations of the L-curve method in ill-posed problems. *BIT Numerical Mathematics*, **36**(2), 287–301.
- Hanke, M, & Hansen, PC. 1993. Regularization methods for large-scale problems. *Survey on Mathematics for Industry*, **3**(4), 253–315.
- Hanke, M, & Raus, T. 1996. A General heuristic for choosing the regularization parameter in ill-posed problems. *SIAM Journal on Scientific Computing*, **17**(4), 956–972.
- Hansen, PC. 1990. The discrete Picard condition of discrete ill-posed problems. *BIT Numerical Mathematics*, **30**(4), 658–672.
- Hansen, PC. 1992. Analysis of discrete ill-posed problems by means of the L-curve. *SIAM Rev.*, **34**(4), 561–580.
- Hansen, PC. 1994. REGULARIZATION TOOLS: A Matlab package for analysis and solution of discrete ill-posed problems. *Numerical Algorithms*, **6**(1–2), 1–35.
- Hansen, PC. 1998. *Rank-deficient and discrete ill-posed problems: numerical aspects of linear inversion*. Philadelphia, USA: Society for Industrial and Applied Mathematics.
- Hansen, PC. 2001. The L-Curve and its use in the numerical treatment of inverse problems. *Pages 119–142 of: Computational Inverse Problems in Electrocardiology*. Advances in Computational Bioengineering. WIT Press.
- Hansen, PC. 2002. Deconvolution and regularization with Toeplitz matrices. *Numerical Algorithms*, **29**, 323–378(56).
- Hansen, PC. 2007. Regularization Tools Version 4.1 (for Matlab Version 7.3). *Numerical Algorithms*, **46**, 189–194.
- Hansen, PC, & O’Leary, DP. 1993. The use of the L-curve in the regularization of discrete ill-posed problems. *SIAM J. Sci. Comput.*, **14**(6), 1487–1503.
- Hansen, PC, Kilmer, ME, & Kjeldsen, RH. 2006. Exploiting residual information in the parameter choice for discrete ill-posed problems. *BIT Numerical Mathematics*, **46**, 41–59.

- Harrer, JU, Mayfrank, L, Mull, M, & Kltzsch, C. 2003. Second harmonic imaging: a new ultrasound technique to assess human brain tumour perfusion. *Journal of Neurology, Neurosurgery and Psychiatry*, **74**, 333–342.
- Hayashi, T, Kashio, Y, & Okada, E. 2003. Hybrid Monte Carlo-Diffusion Method For Light Propagation in Tissue With a Low-Scattering Region. *Applied Optics*, **42**(16), 2888–2896.
- Hebden, J C, Arridge, S R, & Delpy, D T. 1997. Optical imaging in medicine: I. Experimental techniques. *Physics in Medicine and Biology*, **42**, 825–840.
- Hebden, JC. 2003. Advances in optical imaging of the newborn infant brain. *Psychophysiology*, **40**, 501–510.
- Hebden, JC, Correia, T, Khakoo, I, Gibson, AP, & Everdell, NL. 2008a. A dynamic optical imaging phantom based on an array of semiconductor diodes. *Physics in Medicine and Biology*, **53**, N407–N413.
- Hebden, JC, Brunner, J, Correia, T, Price, BD, Gibson, AP, & Everdell, NL. 2008b. An electrically-activated dynamic tissue-equivalent phantom for assessment of diffuse optical imaging systems. *Physics in Medicine and Biology*, **53**, 329–337.
- Hedrick, WR, & Metzger, L. 2005. Tissue Harmonic Imaging: A Review. *Journal of Diagnostic Medical Sonography*, **21**, 183–189.
- Heiskala, J, Hiltunen, P, & Nissilä, I. 2009. Significance of background optical properties, time-resolved information and optode arrangement in diffuse optical imaging of term neonates. *Physics in Medicine and Biology*, **54**(3), 535–554.
- Hill, DLG, Batchelor, PG, Holden, M, & Hawkes, DJ. 2001. Medical Image registration. *Physics in Medicine and Biology*, **46**, R1–R45.
- Hofmann, B. 2005. On multiplication operators occurring in inverse problems of natural sciences and stochastic finance. *TU Chemnitz, Fakultt fr Mathematik*, 107–120.
- Homae, F, Watanabe, H, Nakano, T, Asakawa, K, & Taga, G. 2006. The right hemisphere of sleeping infant perceives sentential prosody. *Neuroscience Research*, **54**, 276–280.
- Hoshino, T, Sakatani, K, Katayama, Y, Fujiwara, N, Murata, Y, Kobayashi, K, Fukaya, C, & Yamamoto, T. 2005. Application of multichannel near-infrared spectroscopic topography to physiological monitoring of the cortex during cortical mapping: technical case report. *Surgical Neurology*, **64**(3), 272–275.
- Igawa, M, Atsumi, Y, Takahashi, K, Shiotsuka, S, Hirasama, H, Yamamoto, R, Maki, A, Yamashita, Y, & Koizumi, H. 2001. Activation of visual cortex in REM sleep measured by 24-channel NIRS imaging. *Psychiatry and Clinical Neurosciences*, **55**, 187–188.
- Ishimaru, A. 1999. *Wave propagation and scattering in random media*. John Wiley and Sons.
- Jacques, SL. 1998. *Skin Optics*. <http://omlc.ogi.edu/news/jan98/skinoptics.html>.
- Jöbsis, FF. 1977. Noninvasive, infrared monitoring of cerebral and myocardial oxygen sufficiency and circulatory parameters. *Science*, **198**, 1264–1267.
- Kaltenbach, JP, & Kaschke, M. 1993. Frequency- and time-domain modelling of light transport in random media. *Medical Optical Tomography: Functional Imaging and Monitoring, SPIE Optical Engineering Press*, **IS11**, 65–86.

- Kawaguchi, F, Ichikawa, N, Fujiwara, N, Yamashita, Y, & Kawasaki, S. 2001. Clinically Available Optical Topography System. *Hitachi Review*, **50**(1), 18–22.
- Kawaguchi, H, Koyama, T, & Okada, E. 2007. Effect of probe arrangement on reproducibility of images by near-infrared topography evaluated by a virtual head phantom. *Applied Optics*, **46**(10), 1658–1668.
- Kepshire, DS, Davis, SC, Dehdhani, H, Paulsen, KD, & Pogue, BW. 2007. Subsurface diffuse optical tomography can localize absorber and fluorescent objects but recovered image sensitivity is nonlinear with depth. *Applied Optics*, **46**(10), 1669–1678.
- Kitagawa, T, Nakata, S, & Hosoda, Y. 2001. Regularization using QR factorization and the estimation of the optimal parameter. *BIT Numerical Mathematics*, **41**(5), 1049–1058.
- Klose, AD, & Hielscher, AH. 1999. Iterative reconstruction scheme for optical tomography based on que equation of radiative transfer. *Medical Physics*, **26**(8), 1698–1707.
- Kohl-Bareis, M, Obrig, H, Steinbrink, J, Malak, J, Uludag, K, & Villringer, A. 2002. Noninvasive monitoring of cerebral blood flow by a dye bolus method: Separation of brain from skin and skull signals. *Journal of Biomedical Optics*, **7**(3), 464–470.
- Koizumi, H, Yamamoto, T, Maki, A, Yamashita, Y, Sato, H, Kawaguchi, H, & Ichikawa, N. 2003. Optical topography: practical problems and new applications. *Applied Optics*, **42**(16), 3054–3062.
- Koizumi, H, Maki, A, Yamamoto, T, Sato, H, Yamamoto, Y, & Kawaguchi, H. 2005. Non-invasive brain-function imaging by optical topography. *Trends in Analytical Chemistry*, **24**(2), 147–156.
- Kubelka, P. 1948. New contributions to the optics of intensely light-scattering materials, Part I. *Journal of the Optical Society of America*, **38**, 448–457.
- Kubelka, P. 1954. New contributions to the optics of intensely light-scattering materials, Part II: Nonhomogeneous layers. *Journal of the Optical Society of America*, **44**, 330–335.
- Leung, TS, Elwell, CE, & Delpy, ST. 2005. Estimation of cerebral oxy- and deoxy-haemoglobin concentration changes in a layered adult head model using near-infrared spectroscopy and multivariate statistical analysis. *Physics in Medicine and Biology*, **50**, 5783–5798.
- Leung, TS, Tachtsidis, I, Tisdall, M, Smith, M, Delpy, DT, & Elwell, CE. 2007. Theoretical investigation of measuring cerebral blood flow in the adult human head using bolus indocyanine green injection and near-infrared spectroscopy. *Applied Optics*, **46**(10), 1604–1614.
- Li, A, Miller, EL, Kilmer, ME, Brukilacchio, TJ, Chaves, T, Stott, J, Zhang, Q, Wu, T, Chorlton, M, Moore, RH, Kopans, DB, & Boas, DA. 2003. Tomographic optical breast imaging guided by three-dimensional mammography. *Applied Optics*, **42**(25), 5181–5190.
- Lloyd-Fox, S, Blasi, A, & Elwell, C. 2009a. Illuminating the developing brain: the past, present and future of functional near infrared spectroscopy. *Neuroscience and Biobehavioral reviews*.
- Lloyd-Fox, S, Blasi, A, Elwell, C, & Johnson, M. 2009b. Illuminating the infant brain: the use of near infrared spectroscopy (NIRS) for studying brain function in young infants. *In: Workshop Developmental Neuroimaging*.

- Lloyd-Fox, S, Blasi, A, Volein, A, Everdell, N, Elwell, CE, & Johnson, MH. 2009c. Social Perception in Infancy: A Near Infrared Spectroscopy Study. *Child Development*, **80**(4), 986–999.
- Mansfield, P, & Morris, PG. 1982. *NMR imaging in biomedicine*. Advances in Magnetic Resonance. London: Academic Press.
- Matcher, SJ, Cope, M, & Delpy, DT. 1994. Use of the water absorption spectrum to quantify tissue chromophore concentration changes in near-infrared spectroscopy. *Physics in Medicine and Biology*, **39**, 177–196.
- Matcher, SJ, Cope, M, & Delpy, DT. 1997. *In vivo* measurements of the wavelength dependence of tissue-scattering coefficients between 760 and 900 nm measured with time-resolved spectroscopy. *Applied Optics*, **36**(1), 386–396.
- Matsuda, G, & Hiraki, K. 2006. Sustained decrease in oxygenated hemoglobin during video games in the dorsal prefrontal cortex: A NIRS study of children. *NeuroImage*, **29**, 706–711.
- Matthews, PM, ET, GD Honey, & Bullmore. 1999. Applications of fMRI in translational medicine and clinical practice. *Nature Reviews Neuroscience*, **7**(9), 732–744.
- Mayer, TE, Hamanna, GF, Baranczyk, J, Rosengarten, B, Klotz, E, Wiesmann, M, Missler, U, Schulte-Altdorneburg, G, & Brueckmann, HJ. 2000. Dynamic CT Perfusion Imaging of Acute Stroke. *American Journal of Neuroradiology*, **21**, 1441–1449.
- McGahan, JP, & Goldberg, BB. 2008. *Diagnostic Ultrasound*. New York: Informa Health Care.
- McRobbie, D W, & E A Moore, M J Graves. 2007. *MRI from picture to proton*. Cambridge, UK: Cambridge University Press.
- Meglinski, IV, & Matcher, SJ. 2002. Quantitative assessment of skin layers absorption and skin reflectance spectra simulation in the visible and near-infrared spectral regions. *Physiological Measurement*, **23**, 741–753.
- Michael, G. 2001. X-ray computed tomography. *Physics Education*, **36**, 442–451.
- Mitsuda, T, & Yoshida, R. 2005. Application of near-infrared spectroscopy to measuring of attractiveness of opposite-sex faces. *Proceedings of the 2005 IEEE. 27th Annual International Conference of Engineering in Medicine and Biology Society*, 5900–5903.
- Montcel, B, Chabrier, R, & Poulet, P. 2005. Detection of cortical activation with time-resolved diffuse optical methods. *Applied Optics*, **44**(10), 1942–1947.
- Nakamura, K, Kawashima, R, Ito, K, Sugiura, M, Kato, T, Nakamura, A, Hatano, K, Nagumo, S, Kubota, K, Fukuda, H, & Kojima, S. 1999. Activation of the Right Inferior Frontal Cortex During Assessment of Facial Emotion. *Journal of Neurophysiology*, **82**, 1610–1614.
- Nissila, I, Kotilahti, K, Huottilainen, M, Makela, R, Lipiainen, L, Nojonen, T, Gavrielides, N, Naatanen, R, Fellman, V, & Katila, T. 2004. Auditory hemodynamic studies of newborn infants using near-infrared spectroscopic imaging. *Proceedings of the 2004 IEEE. 26th Annual International Conference of Engineering in Medicine and Biology Society*, 1244–1247.
- Niu, HJ, Guo, P, & Jiang, TZ. 2008. Improving Depth Resolution of Diffuse Optical Tomography with Intelligent Method. *Pages 514 – 520 of: Proceedings of the 4th international conference on Intelligent Computing: Advanced Intelligent Computing*

- Theories and Applications - with Aspects of Theoretical and Methodological Issues*, vol. 5526.
- Ntziachristos, V, Yodh, AG, Schnall, M, & Chance, B. 2000. Concurrent MRI and diffuse optical tomography of breast after indocyanine green enhancement. *Proceedings of the National Academy of Science*, **97**(6), 2767–2772.
- Obata, A, andH Sato, K Morimoto, Maki, A, & Koizumi, H. 2003. Acute effects of alcohol on hemodynamic changes during visual stimulation assessed using 24-channel near-infrared spectroscopy. *Psychiatry Research: Neuroimaging*, **123**, 145–152.
- Ogoshi, Y, & Okada, E. 2005. Analysis of light propagation in a realistic head model by a hybrid method for optical brain function measurement. *Optical Review*, **12**(3), 264–269.
- Ohnishi, T, Moriguchi, Y, Matsuda, H, Mori, T, Hirakata, M, Imabayashi, E, Hirao, K, Nemoto, K, Kaga, M, Inagaki, M, Yamada, M, & Uno, A. 2004. The neural network for the mirror system and mentalizing in normally developed children: an fMRI study. *Brain Imaging*, **15**(9), 1483–1487.
- Okada, E, & Delpy, DT. 2003a. Near-infrared light propagation in an adult head model. I. Modeling of low-level scattering in the cerebrospinal fluid layer. *Applied Optics*, **42**(16), 2906–2914.
- Okada, E, & Delpy, DT. 2003b. Near-infrared light propagation in an adult head model. II. Effect of superficial thickness on the sensitivity of the near-infrared spectroscopy signal. *Applied Optics*, **42**(16), 2915–2922.
- Okada, E, Firbank, M, Schweiger, M, Arridge, SR, Cope, M, & Delpy, DT. 1997. Theoretical and experimental investigation of near-infrared light propagation in a model of the adult head. *Applied Optics*, **36**(1), 21–31.
- Okui, N, & Okada, E. 2005. Wavelength dependence of crosstalk in dual-wavelength measurement of oxy- and deoxy-hemoglobin. *Journal of Biomedical Optics*, **10**(1), 011015.
- Ono, M, Kashio, Y, Schweiger, M, Dehghani, H, Arridge, SR, Firbank, M, & Okada, E. 2000. Topographic distribution of photon measurement density functions on the brain surface by hybrid radiosity-diffusion method. *Optical Review*, **7**(5), 426–431.
- Otsuka, Y, Nakato, E, Kanazawa, S, Yamaguchi, MK, Watanabe, S, & Kakigie, R. 2007. Neural activation to upright and inverted faces in infants measured by near infrared spectroscopy. *NeuroImage*, **34**, 399–406.
- Patterson, MS, Chance, B, & Wilson, BC. 1989. Time resolved reflectance and transmittance for the noninvasive measurement of tissue optical properties. *Applied Optics*, **28**(12), 2331–2336.
- Paulsen, KD, & Jiang, H. 1996. Enhanced frequency-domain optical image reconstruction in tissues through total-variation minimization. *Applied Optics*, **35**(19), 3447–3458.
- Pea, M, Maki, A, Kovačić, D, Dehaene-Lambertz, G, Koizumi, H, Bouquet, F, & Mehler, J. 2003. Sounds and silence: An optical topography study of language recognition at birth. *Proceedings of the National Academy of Sciences*, **100**(20), 11702–11705.
- Pedersen, J. 2005. *Modular algorithms for large-scale total variation image deblurring*. Ph.D. thesis, DTU, Lyngby.

- Pelphrey, KA, Morris, JP, Michelich, CR, Allison, T, & McCarthy, G. 2005. Functional Anatomy of biological motion perception in posterior temporal cortex: A fMRI study of eye, mouth and hand movements. *Cerebral Cortex*, **15**, 1866–1876.
- Pogue, BW, McBride, TO, Prewitt, J, sterberg, UL, & Paulsen, KD. 1999. Spatially variant regularization improves diffuse optical tomography. *Applied Optics*, **38**(13), 2950–2961.
- Pope, JA. 1999. *Medical Physics: Imaging*. Oxford: Heinemann.
- Powsner, RA, & Powsner, ER. 2006. *Essential Nuclear Medicine Physics*. 2nd edn. Blackwell.
- Pruce, A, Allison, T, Bentin, S, Core, JC, & McCarthy, G. 1998. Temporal cortex activation in humans viewing eye and mouth movements. *Journal of Neuroscience*, **18**(6), 2188–2199.
- Regińska, T. 1996. A regularization parameter in discrete ill-posed problems. *SIAM Journal of Scientific Computing*, **17**(3), 740–749.
- Sato, H, Kiguchi, M, Kawaguchi, F, & Makia, A. 2004. Practicality of wavelength selection to improve signal-to-noise ratio in near-infrared spectroscopy. *NeuroImage*, **21**, 1554–1562.
- Sato, H, Kiguchi, M, & Maki, A. 2006. Wavelength dependence of effective pathlength factor in noninvasive optical measurements of human brain functions. *Japanese Journal of Applied Physics*, **45**(12), L361–L363.
- Schenk, J, Herweh, C, Günther, P, Rohrschneider, W, Zieger, B, & Tröger, J. 2006. Imaging of congenital anomalies and variations of the caudal spine and back in neonates and small infants. *European Journal of Radiology*, **58**(1), 3–14.
- Schmitz, CH, Graber, HL, Luo, H, Arif, I, Hira, J., Pei, Y, Bluestone, AY, Zhong, S, Andronica, R, Soller, I, Ramirez, N, Barbour, SLS, & Barbour, RL. 2000. Instrumentation and calibration protocol for imaging dynamic features in dense-scattering media by optical tomography. *Applied Optics*, **39**, 6466–6486.
- Schmitz, CH, Lcker, M, Lasker, JM, Hielscher, AH, & Barbour, RL. 2002. Instrumentation for fast functional optical tomography. *Review of Scientific Instruments*, **73**(2), 429–439.
- Schmitz, CH, Klemer, DP, Hardin, R, Katz, MS, Pei, Y, Graber, HL, Levin, MB, Levina, RD, NA Franco, WB Solomon, & Barbour, RL. 2005. Design and implementation of dynamic near-infrared optical tomographic imaging instrumentation for simultaneous dual-breast measurements. *Applied Optics*, **44**(11), 2140–2153.
- Schöberl, J. 1997. NETGEN - An advancing front 2D/3D-mesh generator based on abstract rules. *Computing and Visualization in Science*, **1**, 41–52.
- Schünke, M, Ross, LM, Lamperti, ED, Schulte, E, Schumacher, U, Voll, M, & Wesker, K. 2007. *Thieme atlas of anatomy*. Vol. 3. New York: Thieme.
- Schweiger, M, Arridge, S R, Hiraoka, M, & Delpy, D T. 1995. A finite element method for the propagation of light in scattering media: Boundary and source conditions. *Medical Physics*, **22**(11), 1779–1792.
- Seidel, G, Algermissen, C, Christoph, A, Claassen, L, Vidal-Langwasser, M, & Katzer, T. 2000. Harmonic imaging of the human brain. Visualization of brain perfusion with ultrasound. *Stroke*, **31**(1), 151–154.

- Seidel, G, Meyer-Wiethe, K, Berdien, G, Hollstein, D, Toth, D, & Aach, T. 2004. Ultrasound Perfusion Imaging in Acute Middle Cerebral Artery Infarction Predicts Outcome. *Stroke*, **35**(5), 1107–1111.
- Selb, J, Stott, JJ, Franceschini, MA, Sorensen, AG, & Boas, DA. 2005. Improved sensitivity to cerebral hemodynamics during brain activation with a time-gated optical system: analytical model and experimental validation. *Journal of Biomedical Optics*, **10**(1), 011013.
- Selb, J, Joseph, DK, & Boas, DA. 2006. Time-gated optical system for depth-resolved functional brain imaging. *Journal of Biomedical Optics*, **11**(4), 044008.
- Siegel, A, & Sapru, HN. 2005. *Essential neuroscience*. Philadelphia and Baltimore, USA: Lippincott Williams & Wilkins.
- Simpson, C Rebecca, Kohl, Matthias, Essenpreis, Matthias, & Cope, Mark. 1998. Near-infrared optical properties of *ex vivo* human skin and subcutaneous tissues measured using the Monte Carlo inversion technique. *Physics in Medicine and Biology*, **43**, 2465–2478.
- Srinivasan, S, Pogue, BW, Dehghani, H, Jiang, S, Song, X, & Paulsen, KD. 2004. Improved quantification of small objects in near-infrared diffuse optical tomography. *Journal of Biomedical Optics*, **9**(6), 1161–1171.
- Steinbrink, J, Wabnitz, H, Obrig, H, Villringer, A, & Rinneberg, H. 2001. Determining changes in NIR absorption using a layered model of the human head. *Physics in Medicine and Biology*, **46**(3), 879–96.
- Strangman, G, Boas, DA, & Sutton, JP. 2002. Non-invasive neuroimaging using near-infrared light. *Society of Biological Psychiatry*, **52**, 679–693.
- Strangman, G, Franceschini, MA, & Boas, DA. 2003. Factors affecting the accuracy of near-infrared spectroscopy concentration calculations for focal changes in oxygenation parameters. *NeuroImage*, **18**, 865–879.
- Suto, T, Fukuda, M, Ito, M, Uehara, T, & Mikuni, M. 2004. Multichannel near-Infrared spectroscopy in depression and schizophrenia: cognitive brain activation study. *Biological Psychiatry*, **55**, 501–511.
- Taga, G, & Asakawa, K. 2007. Selectivity and localization of cortical response to auditory and visual stimulation in awake infants aged 2 to 4 months. *NeuroImage*, **36**, 1246–1252.
- Tarvainen, T, Vauhkonen, M, Kolehmainen, V, Arridge, S R, & Kaipio, J P. 2005. Coupled radiative transfer equation and diffusion approximation model for photon migration in turbid medium with low-scattering and non-scattering regions. *Physics in Medicine and Biology*, **50**, 4913–4930.
- Thompson, AM, Brown, JC, Kay, JW, & Titterington, DM. 1991. A study of methods of choosing the smoothing parameter in image restoration by regularization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **13**, 4.
- Thompson, JC, Hardee, JE, Panayiotou, A, Crewther, D, & Puce, A. 2007. Common and distinct brain activation to viewing dynamic sequences of face and hand movements. *NeuroImage*, **37**, 966–973.
- Thomson, AM, Kay, JW., & Titterington, DM. 1989. A cautionary note about crossvalidatory choice. *Journal of Statistical Computation and Simulation*, **33**(4), 199–216.

- Tidswell, AT, Gibson, A, Bayford, RH, & Holder, DS. 2001. Electrical impedance tomography of human brain activity with a two-dimensional ring of scalp electrodes. *Physiological Measurement*, **22**, 167–175.
- Tikhonov, AN, & Arsenin, VY. 1977. *Solutions of ill-posed problems*. Winston.
- Toronov, V, Franceschini, MA, Filiaci, M, Fantini, S, Wolf, M, Michalos, A, & Gratton, E. 2000. Near-infrared study of fluctuations in cerebral hemodynamics during rest and motor stimulation: temporal analysis and spatial mapping. *Medical Physics*, **27**, 801–815.
- Toronov, VY, Zhang, X, & Webb, AG. 2007. A spatial and temporal comparison of hemodynamic signals measured using optical and functional magnetic resonance imaging during activation in the human primary visual cortex. *NeuroImage*, **34**, 1136–1148.
- UCL, Spectra. 2005. http://www.medphys.ucl.ac.uk/research/borl/research/NIR_topics/spectra/spectra.htm.
- Uludağ, K, Steinbrink, J, Villringer, & Obrig, H. 2004. Separability and cross talk: optimizing dual wavelength combinations for near-infrared spectroscopy of the adult head. *NeuroImage*, **22**, 583–589.
- Umeyama, S, & Yamada, T. 2009. A new crosstalk measure of near-infrared spectroscopy and its application to wavelength combination optimization. *Journal of Biomedical Optics*, **14**(3), 034017.
- van der Zee, P. 1992. *Measurement and modelling of the optical properties of human tissue in the near-infrared*. Ph.D. thesis, University College London (University of London).
- van Veen, RLP, Sterenborg, HJCM, Pifferi, A, Torricelli, A, & Cubeddu, R. 2004. Determination of VIS- NIR absorption coefficients of mammalian fat, with time- and spatially resolved diffuse reflectance and transmission spectroscopy. In: *OSA Annual BIOMED Topical Meeting*.
- Vogel, C. 1996. Nonconvergence of the L-curve regularization parameter selection method. *Inverse Problems*, **12**(4), 535–547.
- Vogel, CR. 2002. *Computational Methods for Inverse Problems*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics.
- Wager, TD, & Smith, EE. 2003. Neuroimaging studies of working memory: A meta-analysis. *Cognitive, Affective, & Behavioral Neuroscience*, **3**(4), 255–274.
- Wahba, G. 1977. Practical approximate solutions to linear operator equations when the data are noisy. *SIAM Journal on Numerical Analysis*, **14**(4), 651–667.
- Wahba, G. 1990. *Spline models for observational data*. Philadelphia: Society for Industrial and Applied Mathematics.
- Wang, L V, & Wu, H. 2007. *Biomedical optics: principles and imaging*. New Jersey, USA: H Wu.
- Watanabe, E, Nagahori, Y, & Mayanagi, Y. 2002. Focus diagnosis of epilepsy using near-infrared spectroscopy. *Epilepsia*, **43**(9), 50–55.
- Wheeler, JL, Wang, W, & Tang, M. 2002. A comparison of methods for measurement of spatial resolution in two-dimensional circular EIT images. *Physiological Measurement*, **23**, 169–176.

- White, B R, & Culver, J P. 2009. Phase-encoded retinotopy as an evaluation of diffuse optical neuroimaging. *NeuroImage*, doi:10.1016/j.neuroimage.2009.07.023.
- White, DR, Widdowson, EM, Woodard, HQ, & Dickerson, JW. 1991. The composition of body tissues (II). Fetus to young adult. *The British Journal of Radiology*, **64**(758), 149–159.
- Wolbarst, AB, & Hendee, WR. 2006. Evolving and Experimental Technologies in Medical Imaging. *Radiology*, **238**, 16–39.
- Wolf, M, Ferrari, M, & Quaresima, V. 2007. Progress of near-infrared spectroscopy and topography for brain and muscle clinical applications. *Journal of Biomedical Optics*, **12**(6), 062104.
- Woods, RP, Cherry, SR, & Mazziotta, JC. 1992. Rapid automated algorithm for aligning and reslicing PET images. *Journal of Computer Assisted Tomography*, **16**, 620–633.
- Wu, L. 2003. A parameter choice method for Tikhonov regularization. *Electronic Transactions on Numerical Analysis*, **16**, 107–128.
- Xia, J, Weaver, A, Gerrard, DE, & Yao, G. 2006. Monitoring sarcomere structure changes in whole muscle using diffuse light reflectance. *Journal of Biomedical Optics*, **11**(4), 040504.
- Yalavarthy, PK, Pogue, BW, Dehghani, H, & Paulsen, KD. 2007. Weight-matrix structured regularization provides optimal generalized least-squares estimate in diffuse optical tomography. *Medical Physics*, **34**(6), 2085–2098.
- Yamamoto, T, Maki, A, Kadoya, T, Tanikawa, Y, Yamada, Y, Okada, E, & Koizumi, H. 2002. Arranging optical fibres for the spatial resolution improvement of topographical images. *Physics in Medicine and Biology*, **47**, 3429–3440.
- Yamashita, Y, Maki, A, & Koizumi, H. 1999. Measurement system for noninvasive dynamic optical topography. *Journal of Biomedical Optics*, **4**, 414–417.
- Yamashita, Y, Maki, A, & Koizumi, H. 2001. Wavelength dependence of the precision of noninvasive optical measurement of oxy-, deoxy-, and total-hemoglobin concentration. *Medical Physics*, **28**(6), 1108–1114.
- Yaroslavsky, AN, Schulze, PC, Yaroslavsky, IV, Schober, R, Ulrich, F, & Schwarzmaier, HJ. 2002. Optical properties of selected native and coagulated human brain tissues in vitro in the visible and near infrared spectral range. *Physics in Medicine and Biology*, **47**, 2059–2073.
- Yikilmaz, A, & Taylor, G A. 2008. Sonographic findings in bacterial meningitis in neonates and young infants. *Pediatric Radiology*, **38**(2), 129.127.
- Zeff, BW, White, BR, Dehghani, H, Schlaggar, BL, & Culver, JP. 2007. Retinotopic mapping of adult human visual cortex with high-density diffuse optical tomography. *Proceedings of the National Academy of Sciences*, **104**, 12169–12174.
- Zhao, J, Ding, HS, Hou, XL, Zhou, C, & Chance, B. 2005. In vivo determination of the optical properties of infant brain using frequency-domain near-infrared spectroscopy. *Journal of Biomedical Optics*, **10**(2), 024028.
- Zhao, Q, Ji, L, & Jiang, T. 2006a. Improving performance of reflectance diffuse optical imaging using a multicentered mode. *Journal of Biomedical Optics*, **11**(6), 064019.
- Zhao, Q, Ji, L, & Jiang, T. 2006b. Improving performance of reflectance diffuse optical imaging using a multicentered mode. *Journal of Biomedical Optics*, **11**(6), 064019.

- Zhao, Q, Ji, L, & Jiang, T. 2007. Improving depth resolution of diffuse optical tomography with a layer-based sigmoid adjustment method. *Optics Express*, **15**(7), 4018–4029.
- Zhou, C, Yu, G, Daisuke, F, Greenberg, J H, Yodh, A G, & Durdurang, T. 2006. Diffuse optical correlation tomography of cerebral blood flow during cortical spreading depression in rat brain. *Optics Express*, **14**, 1125–1144.
- Zimmerman, RD, Maldjian, JA, Brunc, NC, Horvath, B, & Skolnicke, BE. 2006. Radiologic Estimation of Hematoma Volume in Intracerebral Hemorrhage Trial by CT Scan. *American Journal of Neuroradiology*, **27**, 666–670.
- Zonios, George, Bykowski, Julie, & Kollias, Nikiforos. 2001. Skin Melanin, Hemoglobin, and Light Scattering Properties can be Quantitatively Assessed *in vivo* Using Diffuse Reflectance Spectroscopy. *Journal of Investigative Dermatology*, **117**, 1452–1457.